Learning from texts : extending and revising knowledge

Beker, K.

C H A P T E R

4

Children's Integration of Information Across Texts: Reading Processes and Knowledge Representations

Under revision

Katinka Beker, Paul van den Broek and Dietsje D. Jolles

Abstract

Constructing a knowledge representation from multiple texts requires the integration of information across texts. The aim of the current study was to determine whether children are able to integrate information across texts during reading and whether information from different texts is integrated in their knowledge representation. A sample of 105 children in Grade 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 analyze integration of information from different texts in the resulting knowledge representations. Individual differences in reading comprehension ability and working memory were also taken into account. The results indicate that children in both grades integrate information across texts during reading and that they integrate information across texts in their knowledge representations. Reading comprehension ability and working memory were unrelated to the process of integrating information across texts and the integration of information across texts in knowledge representations. The current study extends previous research by showing that already in fourth grade, children spontaneously integrate information across multiple texts during reading. Results will be discussed in relation to the different mechanisms that may be involved in integration of information across texts.

Introduction

Textbooks are one of the most important sources of information in education. However, 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. This development poses challenges on learning that were previously restricted to expert readers only (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 constructing a knowledge representation from multiple texts. The aim of the current study was to determine whether children in primary school can learn from multiple texts and if so, whether learning is influenced by individual and developmental differences in reading comprehension ability and working memory.

Integration Processes and Integration in Memory

Learning is a broad concept that can refer to a variety of processes (Alexander, Schallert, & Hare, 1991; Beker, Jolles, & van den Broek, in press; Shuell, 1986). In the current article this focus is on one aspect of learning from multiple texts: The integration of information across texts. The integration process roughly consists of two phases: A) the process of integrating information across texts during reading B) integrating text information in memory (including relations across texts). Both of these aspects are important for achieving the educational standards that are relevant for learning from multiple texts in education (Common Core State Standards, 2010; OECD, 2015).

Learning from multiple texts starts with processing a single text. During reading, each piece of information that is being processed activates associated information in memory, including previous parts of the same text and background knowledge (Albrecht & O'Brien, 1993; McKoon & Ratcliff, 1992; van den Broek, Risden, et al., 1996). When reading multiple texts about the same topic, information from an earlier text can also be activated when reading a later text (Beker, Jolles, Lorch, & van den Broek, 2016; Britt & Rouet, 2012; C. A. Perfetti et al., 1999), leading to co-activation of information from the two texts (Goldman & Varma, 1995; Kendeou & O'Brien, 2014; Kintsch, 1988; McRae & Jones, 2013; van den Broek, Risden, et al., 1996). As a result, a connection can be established between the co-activated information elements from the two texts. This connection can be associative or more meaningful, for example, causal. According to the Landscape Model, two factors determine whether information and relations are integrated in memory: The amount and the frequency of (co-)activation of information during reading (van den Broek, Risden, et al., 1996). A third factor that may affect incorporation of information in memory is whether (similar) information is processed in different contexts, which may lead to an enriched knowledge representation (Beker et al., in press). Given this hypothesized correspondence between integration processes and integration of information in memory, one would expect that a failure to integrate information in memory can be traced back to problems with integration processes during reading. This why it is important to study both the integration process and the resulting knowledge representation. In the next section we provide a brief review about previous efforts to investigate integration processes and integrated knowledge representations in children.

Integration of Text Information

There are three lines of research that focus on integration using texts. For each line of research the most important findings with regard the process of integration across texts and the resulting knowledge representation will be summarized.

Integration Within Single Texts

One line of research focuses on integration of information within single texts. A number of studies measured integration after reading by asking questions that require integration. These studies showed that children are able to integrate information in a single text (Cain & Oakhill, 1999; Oakhill, 1982, 1984) and make connections between text information and background knowledge, at least when explicitly prompted (Barnes, Dennis, & Haefele-Kalvaitis, 1996; Cain, Oakhill, Barnes, & Bryant, 2001). Other studies have demonstrated that children are able to integrate information spontaneously *during reading* (Casteel, 1993; Coté, Goldman, & Saul, 1998; Danner & Mathews, 1980; Lynch & van den Broek, 2007; McMaster et al., 2012). However, it is evident that children sometimes struggle with tasks that require integration of information, such as detecting inconsistencies (Markman, 1977, 1979; Vosniadou, Pearson, & Rogers, 1988). The question is whether these findings apply to integration processes between *multiple* texts as well. Some inherent properties of multiple texts could make integration between multiple texts more challenging. For example, the information is usually separated over a larger distance (Beker et al., 2016), one text can be inconsistent with another text (Stadtler & Bromme, 2014), and overlap between the texts may not be recognized (Kurby et al., 2005).

Integration Across Multiple Texts

A second line of research focuses directly on integration of information from multiple texts. This field has only recently started to emerge, so there are only a few studies that measured integration during reading and its relation with the resulting knowledge representation (e.g. Cerdán & Vidal-Abarca, 2008; D. K. Hartman, 1995; Kurby et al., 2005; Strømsø, Bråten, & Samuelstuen, 2003). Overall, the results from these studies indicate that advanced readers integrate information across texts during reading and incorporate connections across texts in memory. There are only a few studies involving children. In one study children aged 11 to 13 were asked to think-aloud while reading two conflicting historical texts (Wolfe & Goldman, 2005). These children were also asked to answer questions that involved several aspects of learning after reading was finished (i.e. integration, complex reasoning, detecting similarities and differences). The results showed that children integrated multiple texts during reading and that this process was positively related to learning (Wolfe & Goldman, 2005). However there also are indications that developing readers struggle with integration of information across texts (Pearson & Hamm, 2005; Sheehan et al., 2006). For example, in a large reading comprehension assessment, sixth graders particularly struggled with tasks that required integration of multiple texts (Sabatini et al., 2014). Furthermore, one study showed that high school students tend to prefer one source and ignore others in building a representation from multiple texts (Wineburg, 1991). Whether younger children integrate multiple texts spontaneously during reading is largely unexplored, even though children are supposed to master these skills already in primary school, for example in writing or presentation tasks (Common Core State Standards, 2010).

Integration Across Multiple Auditorily Presented Texts

A third line of research concerns integration across texts by very young children using auditorily presented texts (Bauer et al., 2012; Bauer & San Souci, 2010; Bauer et al., 2015). Bauer and colleagues had children aged 4 to 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 4-6 year old children were able to integrate information across auditorily presented texts. It is not clear whether these findings generalize to reading situations. Furthermore, the way children respond to prompts *after* processing does not always reflect what happens spontaneously *during* processing (Rapp, van den Broek, McMaster, Kendeou, & Espin, 2007), so prior findings regarding integration across auditorily presented texts after processing may not generalize to spontaneous integration during reading.

The Current Study

Based on these three lines of research it can be concluded that connecting information across texts is an important skill that needs to be mastered by children in primary school, but that there are indications that children may struggle with this skill. We introduce a controlled way to investigate the spontaneous integration processes in young readers that can shed light on factors that facilitate or hinder integration across multiple texts. The aim of the current study was to determine: 1) whether children spontaneously integrate information across texts during reading and 2) whether they incorporate intertextual connections (i.e. connections linking different texts) in memory. In order to answer these questions the multiple-text integration paradigm was used (Beker et al., 2016). In this paradigm the second text of a text pair contains an internal inconsistency. There are two conditions. In the first condition this inconsistency can be resolved by activating an explanation from the first text. In the second condition the inconsistency cannot be resolved by activating information from the first text. The only difference between the conditions therefore, 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 in 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 was ascribed to 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 (Goldman & Varma, 1995; Kendeou & O'Brien, 2014; Kintsch, 1988; McRae & Jones, 2013; van den Broek, Risden, et al., 1996). The present study examined whether children also process the inconsistent target sentence faster in the condition with explanations than in the condition without explanations, indicating that they spontaneously activate information from a previous text while reading a subsequent text. This is a prerequisite for integration of information across texts. As this study is the first to investigate implicit, spontaneous integration processes across different texts, we purposely kept the distance between consecutive texts small. By providing optimal conditions for integrating information across texts we establish a baseline that allows for comparisons to situations in which integrating information across different texts becomes more challenging.

In order to examine whether children incorporate intertextual connections in the knowledge representation we asked children to recall the texts. Recall can be useful to gain insight into text representations built from multiple texts (Britt & Sommer, 2004). Children were asked to report everything they remembered from the text without interference of 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 how integrated information from multiple texts was in the knowledge representation: The number of switches *between* the texts was taken as a measure of integration. It was expected that if children demonstrate activation of prior text information during reading a subsequent text (as indicated by a difference in reading times between the condition with and without explanation), this should 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, Risden, 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.

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; Garner, 1981; Hacker, 1997; Helder, Van Leijenhorst, & van den Broek, 2016; Long, Oppy, & Seely, 1994, 1997; McMaster et al., 2012; Oakhill, 1982; Oakhill & Yuill, 1986; Oakhill, Yuill, & Parkin, 1988; van der Schoot et al., 2012). Measures of reading comprehension ability and working memory were included in the current study to determine whether they also interact with integration across multiple texts during reading. If reading comprehension ability and working memory positively affect integration across texts, then the effects of conditions (with and without explanations) would more strongly affect children with good reading comprehension skills and good working memory skills than children with poor reading comprehension skills and poor working memory skills on both the reading time measures and the measure of intertextual connections in memory. In addition, reading comprehension ability and working memory would also have a main effect on encoding of intertextual connections in memory. Although not of primary interest, there may also be main effects of reading comprehension ability and working memory on reading times. Faster reading may indicate more automatic decoding processes, leaving more working memory resources for reading comprehension processes (C. A. Perfetti, 1985), and possibly integration processes.

In the current study we included children in elementary school from grade 4 because they master the basic reading skills and because at this age children are expected to integrate information across texts, as prescribed by the national educational standards (Expertisecentrum Nederlands [Expertise Centre Netherlands], 2010). Many skills related to reading comprehension and memory formation continue to develop from childhood into adulthood (Kendeou, van den Broek, White, & Lynch, 2009; Oakhill & Cain, 2012; van den Broek, 1997). Therefore, we also included children from grade 6 to determine whether there are grade-related differences between grade 4 and 6 in the ability to integrate information across different texts. Based on previous studies we expected main effects of grade on reading times (Fuchs & Fuchs, 1993) and integration (Bauer & San Souci, 2010). In addition, we expected that the effects of the conditions with and without explanations more strongly affected in grade 6 than children in grade 4 on the reading times measures, which would reflect grade-related differences in integration during reading.

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 Dutch primary schools. Informed consent was obtained from the parents. Only children with good or corrected eyesight, a lack of developmental and reading disorders, were included in the experiment. Participation was rewarded with a small gift.

Materials and Design

Text materials. We created a child-friendly version of the multiple-text integration paradigm (Beker et al., 2016). Children read expository[1] text pairs, in which the second text contained an internal inconsistency, and the first text either contained or did not contain an explanation, that either could or could not help resolve the inconsistency (the Inconsistent-with-explanation and Inconsistent-without-explanation condition respectively). The texts used in prior research were adapted to fit the reading level of children in Grade 4 and 6. To check whether the difficulty level of the adapted texts was appropriate for children in Grade 4 and 6, a reading index was used that provides an indication of the difficulty of the texts based on a variety of text characteristics, namely, the (Dutch) Cito reading index for primary education, or P-CLIB (Evers, 2008; Staphorsius, Verhelst, & Kleintjes, 1996). The average reading index score of the adapted texts indicated that the texts were appropriate for children in Grade 4 and 6. In a pilot study a

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

separate group of children in grade 4 and 6 judged the texts for consistency (yes/no judgment). Only texts that were appropriate for the experimental manipulation were used in the current study.

The topics of the expository texts were realistic but fictitious, to limit the influence of background knowledge. There were 20 different topics, including animals, persons, objects, countries, and events, which were based on real-world knowledge (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 *omitted* an explanation for the inconsistency. In the condition with explanation, the first text described an explanation that could resolve the inconsistency. In 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 4.1.

Table 4.1 *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 lives mainly on the moors, but sometimes also in the woods. The rulver has a pretty brown fur. This fur can be used to make clothing. Hunters can get a lot of money for this fur. *In the winter the rulver's fur turns white.* *Its' brown fur fell off in the fall.* *After this, new white hairs start to grow.* *White camouflage is better against the snow.*	The rulver is an animal with a short tail. The rulver lives mainly on the moors, but sometimes also in the woods. The rulver has a pretty brown fur. This fur can be used to make clothing. Hunters can get a lot of money for this fur. *That is why they try to catch rulvers.* *They catch fewer rulvers than they used to.* *Because there are not many rulvers left.* *The hunters are not happy about this.*
Text 2	The rulver's fur can be used to make coats. To get this fur, the rulver is being hunted in the summer. The rulver's fur has a special brown color. You don't see this brown color on other animals. In the winter the hunt for the rulver stops. <u>Because then you cannot see the rulver in the white snow.</u> The hunt can resume in June.	The rulver's fur can be used to make coats. To get this fur, the rulver is being hunted in the summer. The rulver's fur has a special brown color. You don't see this brown color on other animals. In the winter the hunt for the rulver stops. <u>Because then you cannot see the rulver in the white snow.</u> The hunt can resume in June.

Note. 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.

The texts within a pair were designed to be independent and could be comprehended individually because of their syntactic structure (with the exception of the part with the inconsistency in the second text). Every text began with an introductory sentence, ended with 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 question was twofold: A) To test whether children were paying attention to the task and B) to indicate that the child finished reading the text. The question always concerned literal information 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) 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; but adapted by Swanson et al., 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, and each level consisted of two sets. The items at the easiest level consisted of two sentences and the items at the most difficult level consisted of six sentences. There were 10 sets in total. The task was stopped either when children were not able to answer the comprehension question correctly or when they were not able to recall 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).

Reading Comprehension Ability. The Cito test for reading comprehension is a national standardized norm-referenced test (Cito, 2013a, 2013b). 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 Grade 4 and 6 were obtained from the teachers of the children. The most recent test results were used. On average, the test was administered two months before the experiment. The ‘level scores’ were included in the analyses, which consist of five levels, ranging from I (i.e. the highest level) to V (the lowest level), each level representing 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. 90-95% of the schools in the Netherlands use the Cito test for reading comprehension, so the norms are representative (Egberink, Janssen, & Vermeulen, 2015). The Cito assessment for reading comprehension in Grade 4 and 6 has good reliability and validity (Egberink et al., 2015).

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 these texts for comprehension and to answer several questions about these texts. Some questions were asked immediately after reading the texts and some questions after a delay (i.e. the recall and application questions were provided 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 capitals 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 2500ms 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 and 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 500ms of the previous press. Also, if children took longer than 15.000ms 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 question. The recall questions were always presented in the same order as the topics were presented to the children in the texts. Children were asked 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 free recall question, an 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 lasted about 70 minutes on average, with a short break after reading twelve text pairs and answering the corresponding questions. 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.

Scoring

Recall. Children’s auditory responses were transcribed, parsed into idea units, and coded. One idea unit generally comprised a semantically meaningful clause (consisting of a subject and main verb), which was coded based on the source 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 (“he was…[silence]”) etc.) and metacognitive responses such as “I don’t remember” were excluded from the analysis. Next, the number of source switches between the first and the second text was counted, ignoring information that could not be traced down to one unique text (code type 3) and that was not from either text (code type 4). 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 texts 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, assuming that these reflect processes that are not of interest in the current study (Ratcliff, 1993). Less than 1% of the data were removed using this criterion. The descriptives are presented in Table 4.2.

Table 4.2 *Descriptives for Each Condition in Each Grade (Reading Times in Milliseconds)*

	Condition	*M*	*SE*
Grade 4	Inconsistent-with-explanation	3654.34	73.90
	Inconsistent-without-explanation	3791.60	74.01
Grade 6	Inconsistent-with-explanation	3360.23	60.51
	Inconsistent-without-explanation	3607.30	66.87

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 R-statistics software and the ‘lmerTest’ and ‘effects’ packages.

Item-level reading speeds were clusters at Level 1 and subjects and items were clusters at Level 2, 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 around the grand mean. Degrees of freedom were estimated with Satterthwaite's approximation method (Kuznetsova et al., 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. 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 4.3.

Table 4.3 *Descriptives for Each Condition in Each Grade (Recall Data)*

		Integration[a]		*Total Recall*[b]	
	Condition	*M*	*SE*	*M*	*SE*
Grade 4	Inconsistent-with-explanation	1.01	.06	4.13	.14
	Inconsistent-without-explanation	.92	.05	4.21	.15
Grade 6	Inconsistent-with-explanation	1.23	.05	4.86	.15
	Inconsistent-without-explanation	.96	.05	4.73	.15

[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*

Recall

There was a moderate amount of missing data: 23% of the responses did not involve content-specific information. This was possibly due to the limited number of cues in the recall question. More specifically, the question contained only one non-specific recall cue (e.g. 'the animal') in combination with the unfamiliar topic (e.g. 'the rulver'). This information may not have been sufficient to recall which of the four preceding unfamiliar topics had to be retrieved. Notably, in 53% of the trials on which children did not report content-specific information during free recall, they *did* recall text information spontaneously in response to the subsequent application questions, possibly because these questions contained additional cues. This suggests that the recall task itself provides a relatively low estimate of what the children have represented of the texts. 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 descriptives are presented in Table 4.4. The integration scores were analyzed using hierarchical linear models using the same procedures and following the same steps as in the previous analyses (Table 4.3). 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 compared to 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 not significant after correcting for total recall ($p = .330$). There were no other main or interaction effects.

Table 4.4 *Model Comparisons*

		Model fit			
Model	Tested Against	Reading Times	Integration	Total Recall	Application[a]
Baseline	Condition	$\chi^2(1) = 15.73$*	$\chi^2(1) = 16.98$*	$\chi^2(1) = .01$	-
Condition	Condition + Condition*RCA + Condition*WM + Condition*Grade	$\chi^2(6) = 8.11$	$\chi^2(6) = 22.63$*	$\chi^2(6) = 16.95$*	$\chi^2(3) = 43.43$*

Note: 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.
* $p < .01$

Total recall was analyzed using hierarchical linear models using the same procedures as in the previous analyses (Table 4.3). 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 may be 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 4.3). 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 from sixth grade performed better on the application questions ($M_{proportion\_correct} = .55$, $SE = .02$) than children from 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 (Common Core State Standards, 2010). This requires processing individual texts, as well as integrating and encoding information from multiple texts. If learning is successful, the knowledge representation constructed from multiple texts can be used to solve novel 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 fourth and sixth grade children integrate information across texts during reading and 2) whether they incorporate intertextual connections in memory. In investigating these questions, differences in reading comprehension ability, working memory, and grade were taken into account.

Integration Across Texts During Reading

The multiple-text integration paradigm was used to determine whether information from previous texts was spontaneously activated during reading of subsequent texts (Beker et al., 2016). 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 creation of connections across texts (Goldman & Varma, 1995; Kendeou & O'Brien, 2014; Kintsch, 1988; McRae & Jones, 2013; van den Broek, Risden, et al., 1996). This is the first study to show that children as young as 9 attempt to relate information across texts by spontaneously activating information from previous texts during reading 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). It extends previous findings by showing that integration across texts occurs *spontaneously* during reading using an unobtrusive measure (Bauer et al., 2012; Bauer & San Souci, 2010; Bauer et al., 2015; Wolfe & Goldman, 2005).

Although the current results seem to conflict with previous studies that showed that children particularly struggle 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 explain the seemingly contradictory conclusions. First, whereas previous studies used explicit questions, we used an implicit measure to inspect integration of information across texts. Second, in the current study we created optimal conditions for integration information across texts (i.e. by using experimenter-designed texts) whereas in previous studies the conditions may have been more challenging (i.e. by using ecologically valid texts). Thus, successfully integrating information across texts may depend on the situation. Future studies should focus on manipulating different aspects of the situation to determine under what circumstances integrating information across texts becomes more challenging. By gradually increasing the difficulty of the materials we could determine when and why children sometimes fail to integrate information across texts.

An unresolved question in the current study is whether co-activating information actually led to integrating the information in a meaningful way. It is possible that overlap in key terms between the first and the second text led to activation of information from the first text but that this did not lead to 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') and instead only an associative connection. Future research could employ think-aloud methods in combination with the multiple-text paradigm to determine whether co-activated information was related and if so, whether the relation was 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. There was more integrated recall when connecting the two texts could restore comprehension, i.e., in the conditions that provided explanations compared to the conditions that lacked explanations. Processing times of the target sentence suggest that that integration during recall was the result of co-activation of information during reading. This is in line with how several models describe the integration process and consistent with several empirical findings (Goldman & Varma, 1995; Kendeou & O'Brien, 2014; Kintsch, 1988; McRae & Jones, 2013; van den Broek & Kendeou, 2008; van den Broek, Risden, et al., 1996). Importantly, the effect was not a byproduct of higher recall in general, because there were no differences between the conditions on total recall. In prior research, adults did not show a condition difference in the integration of information in their knowledge representation (Beker et al., 2016). However, these null-results may have been caused by the use of a different, possibly less sensitive, coding procedure, which makes it difficult to compare the results with those from the present study. Another way to shed light on the apparent discrepancy between adults and children is by conducting a new study that includes different measures of knowledge representations (such as primed recognition measures) 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

In the current study, two sources of individual differences were taken into account, reading comprehension ability and working memory. Reading comprehension ability did not affect the process of activating information from prior texts during reading, nor did it affect the construction of knowledge representations. This may reflect the test used to measure reading comprehension ability, which focused on the ability to comprehend individual texts (Cito, 2013a, 2013b). The processes involved in constructing a representation from a single text may be different from those involved in constructing a representation from multiple texts (Stadtler, Scharrer, Brummernhenrich, & Bromme, 2013).

Working memory did not affect processing speed, nor did it affect the integration of information in children's knowledge representations. The absence of an effect of working memory is in contrast with studies showing that this skill is important for reading single texts (Borella et al., 2010; Cain et al., 2004; Seigneuric & Ehrlich, 2005). It could be the case that the distance between the texts was too small, enabling both low- and high-span readers to keep information from the first text activated. Alternatively, it is possible that working memory is not important for the integration of information across multiple texts, and that instead differences in long-term memory affect the process of intertextual integration (Le Bigot & Rouet, 2007). For example, it may be that information from a previous text was no longer active for all children, so that only those children that efficiently encoded and retrieved information from a previous text from long-term memory integrated information across different texts.

Developmental Patterns in Integration Across Texts

In contrast to what was expected based on previous work (Bauer & San Souci, 2010), there were no differences across grades in the ability to integrate information across texts; children in fourth and sixth grade showed similar processing times and knowledge representations. This may be due to the simplicity of the task. The current study was intentionally designed to minimize the challenges posed by the separate texts, to encourage learning from multiple texts (Beker et al., 2016). Therefore, differences across grades may have been 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, which could increase our knowledge about the boundary conditions that determine success or failure in multiple text situations.

Individual Differences in Transfer

Reading comprehension ability and grade affected the ability to apply information from a text to a new situation (i.e. transfer). Good comprehenders performed better on this task than poor comprehenders. There are several explanations for this effect. Good comprehenders may have constructed better knowledge representations of the texts than poor comprehenders (Oakhill, 1982), or their knowledge representation was more available, which helped them to answer the application questions. Furthermore, children in Grade 6 performed better than 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

One issue concerns the interpretation of the direction of the reading time difference between the conditions. The difference could either reflect a speed-up in the condition with explanation or a reduced slow-down. Although we did not include a baseline measure to distinguish between these accounts, we can speculate about the direction of the effect based on 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 first detected, and this 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 would 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 processing the target sentence, for example due to featural overlap (Myers & O'Brien, 1998; van den Broek, Risden, et al., 1996). In the condition with explanation this would 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 would lead to longer processing times, because the target sentence does not fit the knowledge representation. Recent insights in the field of predictive inferences are in favor of 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 is 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 only differ in when co-activation begins: Before or during reading the target sentence. Future research should be done to gain more insight into the fundamental processes that underlie integration across multiple texts.

Limitations

To increase the distinctive boundary between the two texts three or four cues were provided: An intervening task (a comprehension question), an explicit message ("next text"), implicit text structure cues (e.g. introducing each topic in the second text as if it were new), and, for half of 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*."). We did not include a single text control group so it may be that children did not always perceive the texts in one pair as distinct. Nevertheless, this study provides a foundation for investigating intertextual integration processes in a controlled way in more ecologically valid situations in future research. The paradigm can easily be extended to study spontaneous integration processes during reading in situations in which integration is more challenging for children. We view the current study as an initial step in investigating integration of information across texts in children. A reasonable second step would be to increase the textual or physical distance between the texts to determine which factors decrease intertextual integration in more difficult situations. In addition to factors that affect integration within single texts (e.g. featural overlap, reading strategies, etc.), factors that are particularly relevant in the context of multiple texts could also be taken into account (e.g. reliability of the sources, different writing styles, etc.).

Concluding Remarks

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 of other texts (J. A. Hartman & Hartman, 1994). However, the results in the current study suggest that under certain circumstances children *do* process texts in relation to other texts. Children demonstrated integrative processing across texts during reading and integrated information from different texts in memory. These 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.