

VU Research Portal

Moving Beyond Words

Bos, L.T.

2016

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Bos, L. T. (2016). *Moving Beyond Words: Supporting Text Processing Using a Situation Model approach*.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

Chapter 3: Effects of an Inconsistency-Detection Training Aimed at Improving Comprehension Monitoring in Primary School Children

Wassenburg, S. I., Bos, L. T., De Koning, B. B., & van der Schoot, M.

Published in: *Discourse Processes*, 52, 463–488. doi: 10.1080/0163853X.2015.1025203

Note: The intervention study described in this chapter is also part of the dissertation of S. I. Wassenburg



Abstract

Grounded within the situation model framework, this study investigated the effectiveness of a reading comprehension strategy training aimed at improving children's comprehension monitoring strategies. Sixty-four third and fourth graders received a 4-week training targeted at situation model updating, evaluative and self-regulatory strategies, and metacognitive awareness. A group ($n = 51$) following the school's regular reading comprehension curriculum served as a control group. The inconsistency-detection training was expected to show a pretest-to-posttest enhancement in comprehension monitoring strategies (measured with the inconsistency-detection paradigm), general reading comprehension, and reading motivation. Results showed that, compared to the control group, fourth grader's inconsistency-detection performance significantly improved after the inconsistency-detection training. Third graders did not show a significant gain. General reading comprehension and reading motivation scores were promising for children receiving the inconsistency-detection training. In sum, our inconsistency-detection training was an effective means to enhance children's use of monitoring strategies required for constructing and updating a coherent situation model, and to transfer these strategies to novel texts.

A substantial number of primary school children have a poor level of reading comprehension. For example, according to the National Center for Education Statistics (2011), 33% of fourth-grade students and 24% of eighth-grade students in the U.S. are performing below basic standards. Periodic surveys of reading comprehension in the Netherlands (known as PPO) show comparable results (van Berkel, Krom, Heesters, van der Schoot, & Hemker, 2007; van der Schoot, 2008). Furthermore, although the most recent international Progress in International Reading Literacy Study (PIRLS) assessment indicates that the average reading achievement of fourth-grade students in the Netherlands is relatively good compared to other countries, Dutch students perform significantly lower than in 2001 (Mullis, Martin, Foy & Drucker, 2012).

One way to improve reading comprehension is teaching children how to use cognitive reading strategies (National Reading Panel, 2000; Pressley, 1998). In the past decades, a considerable number of reading strategies have been proposed in order to improve text comprehension (for an overview, see De Koning & van der Schoot, 2013; Trabasso & Bouchard, 2002; van der Schoot, Vasbinder, Horsley, & van Lieshout, 2008). The Dutch national institute for curriculum development (SLO) formulates these in terms of sub-goals for reading comprehension for the Dutch educational curriculum. Children in Grade 4 are, for example, required to infer meaningful relations between sentences and paragraphs, recognize inconsistencies, and plan, navigate, monitor, and control their own reading behavior (Aarnoutse & Verhoeven, 2003). However, a gap remains between empirical findings and actual instruction (Aarnoutse & Weterings, 1995; Andreassen & Bråten, 2011; Liang & Dole, 2006; van Keer & Verhaeghe, 2005). That is, curricular textbooks for reading comprehension are often inadequate and the effectiveness of the included reading strategies is not always supported by empirical evidence (Droop, van Elsäcker, Voeten, & Verhoeven, 2012; Houtveen & van de Grift, 2012; Stoeldraijer & Forrer, 2012). Moreover, relatively little teaching time is spent on reading comprehension in Dutch primary schools with even less time devoted to explicit instruction (Houtveen & van de Grift, 2007; Periodic Survey of Educational Level, 2013; van Elsäcker, 2002). In particular, the high-level comprehension processes by which readers evaluate their understanding of a text have typically been overlooked (e.g., Houtveen & van de Grift, 2007). The goal of the present study is, therefore, to improve deep-level text comprehension in primary school children by teaching them evaluative reading

strategies. For these purposes, we adopted the situation model theory (van Dijk & Kintsch, 1983) to serve as our framework in which we set up the training program.

Situation Model Theory

According to the situation model theory, deep-level text comprehension involves the construction of a situation model (Kintsch, 1988; van Dijk & Kintsch, 1983). A situation model is a mental representation of the situation described in the text, rather than of the text itself (e.g., Zwaan & Radvansky, 1998). In order to construct a situation model, the reader is required to integrate information across the text, as well as to integrate text information with his or her background knowledge. Another requirement for readers who construct a situation model is that they have to constantly update the evolving model with new incoming information (Zwaan & Madden, 2004). The purpose of both integration and updating processes is to establish and maintain coherence in the situation model along a number of key narrative dimensions such as time, space, causation, and the protagonists' characteristics, goals, and emotions (Graesser, Singer, & Trabasso, 1994; Zwaan, Langston, & Graesser, 1995; Zwaan, Magliano, & Graesser, 1995). Readers who effectively engage in these processes increase their chance to acquire deep-level understanding of the text; those who do not or do so sub-optimally may benefit from reading strategy instruction to engender these processes (Andreassen & Bråten, 2011; Dole, Brown, & Trathen, 1996; Guthrie et al., 1998, 2004; van der Schoot, Horsley, & van Lieshout, 2010).

A situation model strategy which has been much discussed in the literature on reading comprehension intervention research is inference making (Cain, Oakhill, Barnes, & Bryant, 2001; Elbro & Buch-Iversen, 2013; Mcgee & Johnson, 2003; Yuill & Oakhill, 1988). The reason for this is straightforward. In order to resolve breaks in continuity which typically appear in narrative texts (e.g. gaps in time and space, or causal discontinuities), readers have to learn how to infer the missing information on the basis of text-based or real-world knowledge (McKoon & Ratcliff, 1992). Whereas fostering the integration process during situation model construction seems to be well represented in the literature, surprisingly less attention has been devoted to another important situation model strategy which is mainly targeted at the updating process, i.e., comprehension monitoring (van der Schoot, Reijntjes, & van Lieshout, 2012). Comprehension monitoring refers to the processes by which readers evaluate

their understanding of a text. Skilled readers who evaluate their comprehension constantly ask themselves if what they are reading makes sense. Or, to put it in terms of situation model updating: skilled readers constantly ask themselves whether new textual information is consistent with the information already present in their current situation model. If it does not, they apply repair strategies to resolve the inconsistency or incoherence and restore comprehension in their situation model. The ability to be aware of the ongoing cognitive process and its results as well as adequately adapting cognitive processes accordingly is at the core of comprehension monitoring in particular and reading comprehension more generally, and is often viewed as a key factor that distinguishes skilled readers from less skilled readers (McNamara, Ozuru, Best, & O'Reilly, 2007; Pressley, Borkowski, & Schneider, 1987). The present study aimed at teaching children these types of monitoring strategies required for situation model updating. More specifically, we tested an inconsistency-detection training which was developed to target comprehension monitoring strategies in order to maintain and restore coherence during reading.

The Inconsistency-Detection Training

The inconsistency-detection training consists of four stages, each targeting one of the four goals of the training. In the first stage, learning goals focus on the reader's awareness of whether or not comprehension is occurring (e.g., August, Flavell, & Clift, 1984; Baker & Brown, 1984; Paris, Lipson, & Wixson, 1983). Obviously, it is useless to teach comprehension monitoring strategies to children if they do not know what monitoring is and why it is important (Graesser, 2007; Yuill, 2007). Therefore, children first acquire knowledge about the metacognitive aspects of reading comprehension in general and comprehension monitoring strategies in particular. It is important to note that, in our study, comprehension monitoring strategies include relevant repair strategies (e.g., generating inferences, explanation, elaboration) in order to restore comprehension when necessary. The importance of maintaining coherence is indicated by introducing them to narrative passages containing inconsistencies (Baker & Zimlin, 1989; Zipke, Ehri, & Cairns, 2009). Although previous studies have demonstrated the effectiveness of specific strategy instruction with regard to reading strategy use and reading comprehension (e.g., Brown, Pressley, Van Meter, & Schuder, 1996; Dole et al., 1996; Guthrie, Wigfield, & VonSecker, 2000; Jitendra & Kay Hoppes, 2000; Yuill & Oakhill, 1988), the present study is, to the best of our knowledge, the first reading intervention providing strategy instructions based

on encouraging children to detect (and resolve) inconsistencies in text. In the second stage of the training, children learn how to execute comprehension monitoring strategies required for situation model updating. Hereby, the focus is on strategies needed to identify and resolve the presented inconsistencies, such as inference making and text-level information integration. In the third stage, children are made cognizant of the multidimensional character of a rich and coherent situation model. Recognizing different types of narrative information in a text is necessary for indexing events along the relevant situational dimensions (Zwaan et al., 1995). Therefore, children are asked to apply the strategies taught in the previous stage to texts containing different types of contradictory information appearing along different situation model dimensions. In the final stage of the training, the primary focus is on the importance of using background knowledge in monitoring comprehension for coherence and misunderstanding. Usually, texts do not contain complete and cohesive descriptions of narrative situations. Therefore, readers are required to make inferences based on background knowledge to resolve breaks in continuity and coherence (e.g., to fill in gaps in time and space or missing causal relations between text events). The training program, including the training and testing materials, is purposefully developed for and delivered to children in the third and fourth grade of primary school. In the Netherlands, in second grade, children receive formal instruction in reading comprehension for the first time which primarily concentrates on basic reading comprehension strategies such as identifying main ideas, summarizing, generating questions, and clarifying (Palinscar & Brown, 1984). This forms the basis for learning the higher-order strategies, such as comprehension monitoring, that are crucial for deep-level understanding of texts that need to be acquired later on, from the third grade onwards. We took this into account while designing our training program. In order to explore possible grade-related developmental trends in the data, grade was used as a separate factor throughout the analyses.

Evaluation of the Training's Effectiveness

As stated earlier, the inconsistency-detection training was designed to enhance children's use of monitoring strategies supportive to situation model updating and hence deeper comprehension of text. The inconsistency-detection paradigm is often used to assess a reader's ability to maintain coherence by connecting currently processed information with preceding context (e.g., Albrecht & O'Brien, 1993; Long

& Chong, 2001). Readers are required to activate prior text information and update their developing situation model. To evaluate the effectiveness of the training, we therefore used an inconsistency-detection test to assess the comprehension monitoring strategies at pre- and posttest. In this test, participants read texts in which some piece of target information (e.g., being satisfied with one's exam grade) is either consistent or inconsistent with regard to the "context" information (e.g., passing or failing the exam respectively) presented earlier. The difference in reading times between inconsistent and consistent passages provides information about their relative processing difficulty. The main assumptions of the inconsistency-detection paradigm are that: (a) inconsistencies interfere with the updating process as these are difficult to integrate with the information already present in the evolving situation model (van der Schoot et al., 2012), and (b) inconsistency detection and resolution processes are associated with longer reading times (i.e., the so-called inconsistency effect; Hyönä, Lorch, & Rinck, 2003). Based on these assumptions, we hypothesized that a more effective use of the reading strategies acquired in the inconsistency-detection training would be reflected in a pretest to posttest enhancement of the inconsistency effect. When readers engage in updating and monitoring processes, it is more likely they detect the inconsistency. In addition, we expected the inconsistency-detection training to lead to increased general reading comprehension levels. This expectation is based on the notion that comprehension monitoring is an integral part of teaching reading comprehension and, as such, it was reasonable to hypothesize that training benefits would transfer to texts other than the ones developed specifically for our experimental purposes (i.e., the inconsistency-detection test). To increase the likelihood that improved monitoring strategies, due to the training, would transfer to general reading comprehension, children practiced with texts differing in length, difficulty, and types of incoherencies. In addition, to further promote transfer of acquired monitoring strategies, we used naturalistic or only loosely constrained texts that corresponded to the children's own world of experience. Finally, we explored the possibility that our training in comprehension monitoring strategies would influence reading motivation. In particular, children may develop a more positive attitude towards reading comprehension after the training than they had before because the training was specifically designed to raise awareness of self-evaluation during reading. From research on training in metacognition, we know that children who actively deploy such self-regulatory strategies are more engaged in their learning and academic

performance and thus are more motivated (Bandura, 1997; Guthrie et al., 2004; Guthrie et al., 2007; Schunk & Zimmerman, 1997).

Method

Participants

Participants were 115 third and fourth graders from four regular primary schools in a large urban area in the Netherlands. The schools where the intervention took place were in collaboration with the university, but participation in the intervention study was voluntary. In accordance with a procedure preferred by the schools and endorsed by the ethical committee of the faculty, parents were provided a letter about the aim and methods of the study and could allow or deny the participation of their child by returning a preprinted objection note.

Children with dyslexia ($n = 12$) and an IQ less than 85 ($n = 6$), as indicated by school records, were excluded. In addition, children with diagnosed autism ($n = 6$) were excluded from the study. This resulted in the removal of 24 children from the initial sample. Of the remaining children, 64 children participated in the inconsistency-detection training group and 51 children formed the control training group, which followed the school's regular reading comprehension curriculum.

Random assignment of children was not possible due to practical and organizational reasons imposed by the schools. For example, schools preferred not to make within-class divisions between groups of experimental and control children. However, after carefully assigning classes to conditions, it turned out that gender ratio, age, and IQ (as determined by raw scores on Raven's Standard Progressive Matrices - Short Form) were comparable for the experimental group (31 boys, 33 girls; $M_{\text{age}}: 9.65$, $SD = 0.73$; $M_{\text{Raven}} = 21.38$, $SD = 3.93$) and control group (22 boys, 29 girls; $M_{\text{age}}: 9.62$, $SD = 0.71$; $M_{\text{Raven}} = 20.87$, $SD = 2.80$). All children had grade-level decoding skills.

Design

The study employed a pretest-posttest control group design in which the independent variables were time (pretest vs. posttest), training group (monitoring vs. control), and Grade (3 vs. 4), and the dependent variables were inconsistency detection (i.e., reading times on consistent vs. inconsistent sentences), general

reading comprehension level, and reading motivation. Pre- and posttests were administered by trained research assistants in the three weeks before and after the monitoring training, and consisted of different versions of the same tests. The order of the tests was counterbalanced across participants.

Inconsistency-Detection Training

Comprehension monitoring was taught in a 4-week training program containing eight 30-minute lessons (two lessons per week). Specifically, half of the lessons were instruction lessons conducted in groups (lessons 1, 3, 5, and 7), the other half were computer-based lessons conducted individually (lessons 2, 4, 6, and 8). Instruction lessons and computer-based lessons were taught alternately, so that each instruction lesson was followed by a computer-based lesson, together covering one of the four stages of the training every week. That is, the four main aspects, or goals, of the training program were taught in four corresponding stages such that in each stage an instruction lesson and a computer-based lesson comprised one of the four goals. All lessons consisted of a balanced approach of direct instruction, modeling, guided practice, and individual practice (Baumann & Schmitt, 1986; Brown et al., 1996; Houtveen & van de Grift, 2007; Palincsar & Brown, 1984). Particularly, depending on the type of lesson, relatively more time was spent on direct instruction, modeling, or guided practice (in the instruction lessons), or on guided or individual practice (in the computer-based lessons). The trainers who conducted the lessons followed standardized instructions and had received elaborate training before the start of the inconsistency-detection training.

From the first stage of the training onwards, children were provided with practical guidelines which coincided with the goals set out for the training and served as a means to help the children perform, and learn from, the different exercises used to induce the required comprehension monitoring strategies. The guidelines were cumulative in nature in that new guidelines built on already established guidelines and practiced strategies. Inevitably, the different monitoring strategies that were trained were related to each other and difficult to isolate, so the presented guidelines overlapped in their content and use. Also, throughout the training, the group-based discussions about how to carry out the monitoring processes revolved around these guidelines. The guidelines had a prominent, recurrent role in the instruction lessons

and they were practiced with progressive difficulty during individual practice in the computer-based lessons.

To promote the children's engagement and motivation for the training, a number of measures were taken. First, the training program was presented to the children as a 'professor training', meant to teach them, among other things, how to 'mark' passages of text and indicate what is wrong with them in terms of internal coherence (within-text inconsistencies) or external coherence (inconsistencies between text information and prior knowledge). Second, an abundant variety of relevant, level-appropriate, and appealing text examples was used to further explain the guidelines and instructions, and guide the children in reaching the training goals. Third, we used scaffolding techniques including (a) gradual fading of teacher support and (b) transfer of responsibility from the teacher to the students once the latter were beginning to become more competent (Guthrie, Taboada, & Coddington, 2007; Houtveen & van de Grift, 2007). Finally, each lesson ended with a reflective discussion about what and how the children had learned and why that is important.

Instruction lessons. Instruction lessons were provided to children in groups of 5 to 6 children in a separate classroom. Instead of relying primarily on a top-bottom approach, the instruction lessons were collaborative and interactive, with students in the role of engaged learners. Throughout the training, children practiced with passages of narrative texts containing some incoherence or inconsistency. These passages were split up into two parts. The first part was always consistent, whereas the second part contained an inconsistency with regard to the first part. After each part, a few questions were asked to check whether the children were engaging in the instructed strategies.

The first instruction lesson (Lesson 1) was more or less introductory. Children were given general information about the training to make them familiar with its method and goals. The training goal for the first lesson specifically was to make children aware of their own comprehension and to teach them what they can do to monitor relevant information in the text. Furthermore, they were made aware of the possibility to use repair strategies when necessary (e.g., rereading, inference making). Questions that were asked during reading were of a metacognitive nature (e.g., 'what am I reading about?', 'does this still make sense?', or 'can I resolve this?'),

in order to raise consciousness of the comprehension process (King, 2007; Scardamalia & Bereiter, 1992). When the children indicated that a passage did not make sense anymore, they were asked to explain why.

In the second instruction lesson (Lesson 3), the focus was on specific monitoring strategies in order to foster situation model updating: recognizing what is important information and what is not, keeping the important information available during reading, validating new incoming information against it, and integrating all information into one situation model (Albrecht & O'Brien, 1993). Children were instructed to elaborate on what they read and integrate new information from the text with earlier encountered information. To illustrate this, children read passages in which an action of the protagonist (e.g., eating a hamburger at McDonald's) was inconsistent with a description of his/her character or goals given earlier (e.g., a vegetarian, or someone who wanted to go wining and dining in a five-star restaurant) (Albrecht & O'Brien, 1993). After the first part of the passage (character description), participants were asked to describe what they knew about the protagonist. After the second part (containing the inconsistent information), children were asked to underline in the text what did not make sense to them and to explain why. Finally, they were asked to come up with plausible information that would possibly resolve the inconsistency (e.g., the protagonist may have ordered a veggie burger).

In the third instruction lesson (Lesson 5), children practiced again with these comprehension monitoring strategies. However, in addition to Lesson 3, children were made cognizant of the different types of information in a narrative text. In particular, it was explained to them that the situational dimensions of protagonist (including his/her intentions and emotions), time, space, and causation should be considered the "cornerstones" laying the "foundation" of a situation model. Children were taught how to process these dimensions so as to construct and gradually update a rich and coherent situation model. They read passages of text of which the main topics could be classified into the narrative dimensions of time, space, causation, and emotion, comparable to the texts used at pre- and posttest in the inconsistency-detection paradigm (see Table 1 for an example situation within each dimension). After describing what was inconsistent in a passage and why, they had to indicate the "situational nature" of the contradiction.

In the fourth instruction lesson (Lesson 7), the focus was shifted from internal inconsistencies (i.e., inconsistencies within the text) to external inconsistencies (i.e., inconsistencies between text information and background knowledge) to teach children to involve their own background knowledge during reading (Elbro & Buchlversen, 2013; Pressley et al., 1987). They were instructed to elaborate on what they read and integrate this with prior knowledge. To practice this, children read passages containing information inconsistent with regard to general world knowledge (e.g., 'blue with green makes red', 'eating soup with a fork', 'it is cold in Africa'). They indicated what was wrong and explained why.

Table 1
Samples of consistent vs. inconsistent situations for the four narrative dimensions

Context sentence	Target sentence
Emotion:	
Eric had just heard that he got <i>an A</i> on his last exam.	Eric was <i>very happy</i> and could not wait to tell everyone.
Eric had just heard that he got <i>a F</i> on his last exam.	
Time:	
John's train arrived 20 minutes <i>later</i> than Sarah's train.	Sarah was <i>already waiting</i> when John arrived at the station.
John's train arrived 20 minutes <i>earlier</i> than Sarah's train.	
Space:	
Construction workers were busy in the <i>basement</i> of the house.	The landlord <i>walked up</i> all the stairs to bring them a drink.
Construction workers were busy in the <i>attic</i> of the house.	
Causation:	
The boy scout threw some more <i>wood</i> on the campfire.	The campfire <i>flared up</i> and kept everybody nice and warm.
The boy scout threw some more <i>water</i> on the campfire.	

Note. The combinations of the top context sentence with the target sentence are the consistent versions. The inconsistent versions are obtained by combining the bottom context sentences with the target sentence. Target sentences were the same in both conditions. Relevant information is printed in italics for illustration only. In the near condition, context and target sentences were adjacent. In the distant condition, the context and target sentence were separated by a neutral filler paragraph. We used some example stimuli of Ferstl and von Cramon (2007).

Computer-based lessons. In the computer-based lessons, which took place in a separate computer classroom, children read passages of text of which some contained inconsistencies. Question-answering techniques were used to trigger the monitoring processes that were taught in the instruction lessons. The exercises in the computer-based lessons were constructed in such a way that they followed the cumulative nature of the guidelines. The purpose of these exercises was to practice the learned monitoring strategies in order to automatize them. The passages increased in length, from two-sentence passages to longer multi-sentence passages. With the increasing length of the passages, we were able to manipulate the distance between the two pieces of conflicting text information. Research shows that it is more difficult to detect “distant” than “near” inconsistencies, because of the increased working memory demands (Oakhill, Hartt, & Samols, 2005; Orrantia, Múñez, & Tarín, 2013). Practicing both these types of inconsistencies presumably fosters transfer of strategies to longer discourse.

In the first computer-based lesson (Lesson 2), the focus was, as in the first instruction lesson, on raising awareness about comprehension and text coherence. During reading, children indicated the sentence from where a passage did not make sense anymore by clicking on it. After that, they were required to click on the specific words that caused the inconsistency and to elaborate on exactly why it did not make sense by writing an explanation in a separate textbox under the text. In the second computer-based lesson (Lesson 4), children were specifically asked to explain why a passage did not make sense. Subsequently, they answered questions regarding the situation-relevant context information which served as the source of the inconsistency (in this case, the description of the character of the protagonist) and against which later information in the text (in this case, the actions of the protagonist) could be evaluated. This way, they were able to check their own answer before the correct answer was provided. Finally, they were asked whether they could come up with some plausible explanation for the inconsistency.

The goal of Lesson 4 was for children to practice executing the different monitoring strategies, like integrating information across text into one situation model, updating this model with new incoming information while maintaining coherence, and using repair strategies if necessary. In the third computer-based lesson (Lesson 6), children additionally answered questions about the nature of the contradictory information

(i.e., was the inconsistency related to the temporal, spatial, causal or emotional aspects of the described situation?). In the last computer-based lesson (Lesson 8), in addition to internal inconsistencies, children also practiced with external inconsistencies. In all computer-based lessons, after giving the answer, children were provided with the correct answer as well as an explanation of why this answer was correct and other answers were not.

Although the assignments during the computer-based lessons had to be made individually, the beginning and end of these lessons were held in groups. In the same groups as in the instruction lessons, each computer-based lesson started with a recapitulation of what was taught in the preceding instruction lesson and ended with a discussion in which the children reflected on the what, how and why of the strategic reading activities which were trained.

Control Training

Children in the control training group followed the school's regular curriculum taught by their own teachers. This meant that they attended reading comprehension courses twice a week, just like the children in the inconsistency-detection training group. The reading lessons involved whole class reading as well as a small group reading instruction at their appropriate reading level combined with individual practice. Reading strategies that were taught included predicting, clarifying, and summarizing.

Pre- and Posttests

Pre- and posttests were administered at school. Children individually completed the inconsistency-detection test in a silent room. The reading comprehension test and the reading motivation question were completed in the classroom using a whole-class test taking approach.

Inconsistency-detection test. In the pre- and posttests, comprehension monitoring strategies were measured using an inconsistency-detection test adapted from previous research (Albrecht & O'Brien, 1993; van der Schoot et al., 2012; Wassenburg, Beker, van den Broek, & van der Schoot, 2015). The test consisted of 32 passages of which half contained inconsistencies between different pieces of information, whereas the other half were consistent. Children were instructed to read silently at a

normal pace. They were under the assumption they would read 'normal' passages; no information was provided with regard to inconsistencies. After each text, a question about non-critical information (i.e., information that is held constant for all experimental versions of a passage and is not relevant to the inconsistency) from the passage appeared on the screen. Participants answered the question by saying yes or no to the experimenter. The purpose of these questions was to ensure that participants would read the passages for comprehension. In total, the inconsistency-detection test took approximately 25 to 40 minutes, depending on the participants' reading speed.

In the inconsistency-detection test, participants read a passage in which an action of the protagonist or certain aspect of the situation (i.e., target sentence) was either consistent or inconsistent with information presented earlier in the text (context sentence). The context sentence provided critical information on one of four situational dimensions (i.e., time, space, causation, and emotion). The context sentence and the target sentence were adjacent (near condition) or separated by a substantial amount of intervening text (distant condition). The purpose of this 4-sentence filler paragraph was to ensure that context information was no longer active in the working memory (Albrecht & O'Brien, 1993; Long & Chong, 2001). Consistency was manipulated by changing one or two words in the context sentence. Within each situational dimension, the number of syllables, words and characters were approximately the same for context sentences across all conditions. Target sentences were exactly the same in all conditions for a certain passage to ensure comparability between the conditions. Passages always started with an introductory sentence and ended with a wrap up sentence. All of them were designed to match the students' reading levels and had the same structure.

In total, there were 16 within-subject conditions formed by crossing three factors: consistency (consistent vs. inconsistent), location (near vs. distant), and situational dimension (time, space, causation, and emotion). Each participant was presented with 32 experimental passages, 2 in each condition. In other words, for every dimension, 8 different passages were constructed, with four versions of every passage (i.e., near/consistent, near/inconsistent, distant/consistent and distant/inconsistent). Table 1 shows an example passage for consistent and

inconsistent situations for each narrative dimension (table taken from Wassenburg et al., 2015).

The stimuli were arranged into four lists, each containing the 32 stories. In order to ensure a full combination of conditions and materials, the four different versions of each story within each dimension were counterbalanced across the lists by means of a Latin square design. Each list was presented to approximately the same number of children. Thus, across lists and across participants, each passage within each dimension occurred equally as often in the near/consistent, near/inconsistent, distant/consistent and distant/inconsistent versions. Passages within each list were presented in a pseudo-randomized order. Reading times on the target sentences were collected using the self-paced moving window method (Just, Carpenter, & Woolley, 1982; for more information, see van der Schoot et al., 2012; Wassenburg et al., 2015).

Reading comprehension. The Grade 3 and Grade 4 versions of the standardized CITO Reading Comprehension Test were used to measure children's reading comprehension skills (Institute for Educational Measurement, 2010). This test is part of the standard Dutch pupil monitoring system and is designed to determine general reading comprehension level in primary school children. It contains two modules, each consisting of a text and 25 multiple-choice questions. The questions were designed to tap both the text-base and situation model representation which can be constructed from the text (e.g., Kintsch, 1988) and pertained to either the word, sentence or text level.

Normed proficiency scores were obtained by rescaling students' raw test scores on the 50 items. The rescaling procedure enabled us to compare the results of the pre- and posttest versions of the CITO Reading Comprehension Test. Furthermore, the proficiency scores allow for comparisons between children from different grades (i.e., Grade 3 and 4). The internal consistency coefficient of the tests was high with Cronbach's alphas no lower than .85 (Feenstra, Kamphuis, Kleintjes, & Krom, 2010).

Reading motivation. We asked the children how much they liked reading comprehension before and after the training. Children had to answer on a 4-point Likert scale represented by cartoon figures (1 = I do not like it at all; 4 = I like it a lot).

After this, children were provided the opportunity to orally explain why they did or did not like reading comprehension. The oral answers were not documented. Their purpose was to give the children the chance to provide their opinion and express their wishes for improvement.

Results

Inconsistency-Detection Test

Analysis of the comprehension questions showed that responses were above chance ($M = 85\%$ correct, range 60-100%), suggesting that, as instructed, participants read the passages naturally. That is, it is likely that they did not just guess or read the passages too thoroughly. Reading times < 80 ms per syllable as well as reading times ± 3 standard deviations from the overall condition mean were excluded from the analyses. This resulted in removal of less than 3% of the reading times.

An overall $2 \times 2 \times 2 \times 2$ mixed analysis of variance (ANOVA) was conducted with Consistency (consistent vs. inconsistent) and Time (pretest vs. posttest) as within-subject factors, and Training Group (inconsistency-detection vs. control) and Grade (Grade 3 vs. Grade 4) as between-subject factors¹. In Figure 1, reading times for target sentences per syllable are represented as a function of Consistency, Time, and Training Group for Grade 3 and 4. The results show that overall reading times were slower for inconsistent than for consistent target sentences, $F(1, 106) = 19.56, p < .001, \eta_p^2 = .16$, indicating that, generally, children were able to detect the inconsistencies. Children read faster at posttest than pretest, $F(1, 106) = 23.81, p < .001, \eta_p^2 = .18$, and children in Grade 4 read faster than children in Grade 3, $F(1, 106) = 9.80, p = .002, \eta_p^2 = .09$. As can be seen in Figure 1, the effects of Consistency, Time, and Training Group on reading time varied as a function of Grade, as was evident in the significant 4-way Consistency \times Time \times Training Group \times Grade interaction, $F(1, 106) = 4.55, p = .035, \eta_p^2 = .04$. Therefore, further analyses on reading times were performed for Grade 3 and 4 separately.

¹ The within-subject factors Location (near vs. distant) and Dimension (time vs. space vs. causation vs. emotion) did not show any effects relevant to the intervention (i.e., there were no significant interaction effects with Time and Training Group). Reporting all effects and discussing the non-significant 5 and 6-way interactions goes beyond the theoretical scope of this paper. Therefore, we decided to exclude both factors from the analyses.

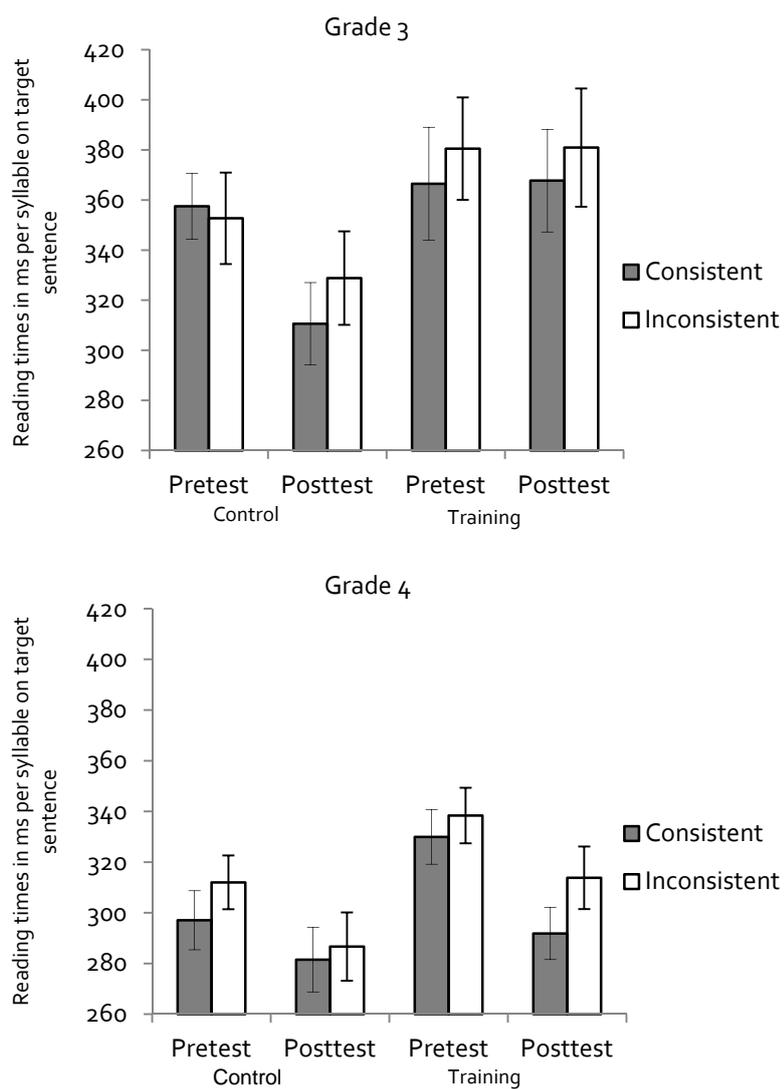


Figure 1. Reading times per syllable on target sentences (in ms) as a function of consistency, time, and training group for Grade 3 and Grade 4. Error bars depict standard errors of means.

Whereas the results showed no significant main effects or interactions for Grade 3, $F(1, 27) < 4.15$, *ns*, children's reading times for Grade 4 showed clear effects of Consistency, $F(1, 79) = 25.87$, $p < .001$, $\eta_p^2 = .25$, and Time, $F(1, 79) = 42.25$, $p < .001$, $\eta_p^2 = .35$, indicating that children in Grade 4 read inconsistent sentences slower than consistent sentences and read faster at posttest than at pretest. More interestingly, however, were the pre- and post-training group differences in reading time patterns for Grade 4. At the pretest, the inconsistency-detection and control training groups performed similarly with regard to showing an inconsistency effect, Consistency \times Training Group: $F(1, 81) = .99$, *ns*. At posttest, however, performance differed for the two groups. Specifically, children in the inconsistency-detection training group showed a significantly larger inconsistency effect (i.e., slower reading times on inconsistent than consistent target sentences) than children in the control training group, Consistency \times Training Group: $F(1, 81) = 5.28$, $p = .024$, $\eta_p^2 = .06$. This pattern of results was evident in the significant Consistency \times Training Group \times Time interaction, $F(1, 79) = 4.47$, $p = .038$, $\eta_p^2 = .05$. In sum, these results indicate that Grade 4 children who had received the inconsistency-detection training slowed down their reading more upon encountering an inconsistency than children who had followed the school's regular curriculum.

Reading Comprehension and Reading Motivation

The inconsistency-detection training did not only result in more effective usage of the trained reading strategies as measured by the inconsistency-detection test. From Figure 2, it can be seen that the inconsistency-detection training also led to increased general reading comprehension levels. An overall $2 \times 2 \times 2$ mixed analysis of variance (ANOVA) was performed on the reading comprehension proficiency scores with Time (pretest vs. posttest) as the within-subject factor, and Training Group (inconsistency-detection vs. control) and Grade (Grade 3 vs. Grade 4) as between-subject factors. The results indicated that the interaction between Training Group and Time was statistically significant, $F(1, 106) = 4.93$, $p = .028$, $\eta_p^2 = .04$, and that this interaction did not vary as a function of Grade, $F(1, 106) = .34$, $p = .562$.

Despite the fact, however, that the Training Group \times Time interaction was not statistically affected by Grade, Figure 2 clearly showed different patterns of performance between the two groups. The Training Group \times Time interaction for fourth graders appeared to be based on the similarity of the training and control

groups at post-test while the training group was less proficient on the pretest. For the third graders, the Training Group \times Time interaction seemed to have arisen, at least in part, from a drop in performance at posttest in the control group. Therefore, we decided to also analyze the data separately for each group. The results showed that the Training Group \times Time interaction was not significant for Grade 4, $F(1, 80) = 2.41$, $p = .125$, and marginally significant for Grade 3, $F(1, 26) = 3.88$, $p = .060$, $\eta_p^2 = .13$.

In addition to the gains in general reading comprehension, we also examined whether children would develop a more positive attitude towards reading comprehension due to the inconsistency-detection training. In Table 2, the mean reading motivation scores are presented as a function of Training Group and Time for both grades. Results from the ANOVA (with the above factors) on the motivation data revealed that children tended to become more positive about reading comprehension from pretest to posttest, $F(1, 108) = 3.02$, $p = .085$, $\eta_p^2 = .03$. Although visual inspection of Table 2 shows that the gain in motivation was larger for children in the inconsistency-detection training group than for children in the control training group, the ANOVA $F(1, 108) = 2.23$, $p = .138$, $\eta_p^2 = .02$.

However, when the ANOVA was recalculated excluding the between-subjects factor of Grade ², the interaction reached a marginal level of significance, $F(1, 110) = 2.88$, $p = .093$, $\eta_p^2 = .03$, with post hoc tests revealing a significant increase in reading motivation for children who had received the inconsistency-detection training, $t(62) = -3.62$, $p < .001$, $d = .46$, but not for children in the control training group, $t(48) = -.52$, *ns*.

Table 2. Mean reading motivation score as a function of Grade and Training Group (standard deviations are shown in parenthesis).

		Grade 3		Grade 4	
		Control	Training	Control	Training
How much do you like reading comprehension?	Pretest	1.92 (1.16)	1.94 (1.03)	2.46 (.90)	2.02 (.91)
	Posttest	1.83 (.94)	2.18 (1.01)	2.59 (.76)	2.48 (.96)

² Grade did not confound the effects of Training Group and Time, as evidenced by the absence of significant interactions between these critical factors and Grade. When using a more conservative value of $p > .25$, it is generally permitted to pool across the levels of a factor in order to increase power and create a more parsimonious model (e.g., Hines, 1996).

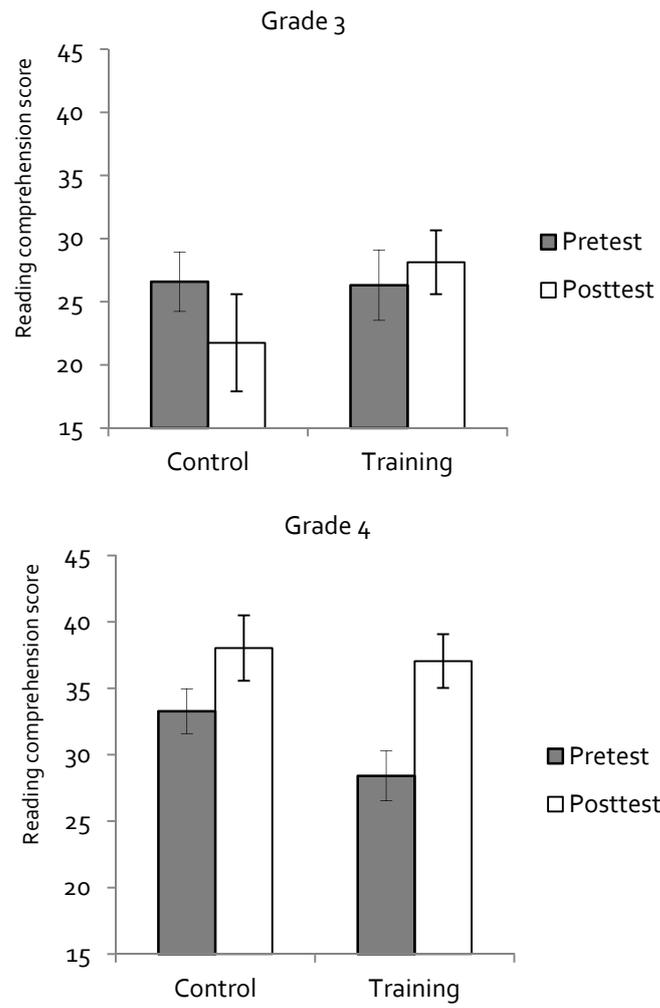


Figure 2. Reading comprehension proficiency scores as a function of Training Group and Time for Grade 3 and 4. Error bars depict standard errors of means.

Discussion

This study examined the effectiveness of a reading strategy training aimed at improving children's evaluative and self-regulatory strategies. Situated within the situation model framework, this intervention moves beyond previous training studies in that it mainly focuses on creating coherence from texts by integrating information and updating the evolving situation model during reading. It was hypothesized that teaching third- and fourth-grade children these comprehension monitoring strategies would improve their situation model updating ability and subsequently enhance deep-level text comprehension. Overall, the present results provide support that the inconsistency-detection training was effective at fostering specific comprehension monitoring processes. Moreover, the training seems to have a positive influence on general reading comprehension performance.

The first major finding of this study is that fourth-grade children showed improved comprehension monitoring strategies (e.g., integration and updating processes) after the inconsistency-detection training but not after the control training, as evidenced by a pretest to posttest enhancement of the inconsistency effect. That is, only after the inconsistency-detection training did fourth graders slow down their reading on encountering inconsistent information, as compared to consistent information. Applying the inconsistency-detection paradigm as a main outcome variable at pre- and posttest is an important added value of this study as it enabled us to provide a process-oriented perspective on the effectiveness of our intervention. Online measures of reading comprehension are thought to have advantages over traditional offline measures such as question answering and self-report of reading strategy use. For example, they are less contaminated by processes not relevant to internal text representation (Gorin, 2005). Also, unlike verbal reports, self-paced reading methodology does not run the risk of biasing the readers' attention to the experimenter's hypotheses (Kirk & Ashcraft, 2001). Combining these aspects with the fact that children were not aware of the nature of the test (i.e., they were not told that some texts contained inconsistencies) suggests that the observed effects are most likely due to the content and nature of the inconsistency-detection training rather than specific task demands or other instructions.

Interestingly, children in Grade 3 did not show evidence of a more effective use of monitoring strategies. A possible explanation is that the tasks or training materials were too difficult for third-grade children, although this is not very likely because our training program was purposefully developed together with teachers to be level-appropriate for both third- and fourth-grade children. The gradual progression from simple and short texts to more complex and longer texts within and between lessons ensured that there were enough level-appropriate passages for each child to practice with. A more plausible explanation is that children in Grade 3 are simply less matured and have had less practice with deep-level reading strategies than children in Grade 4. Children at the beginning of Grade 3 have only received one year of formal instruction in basic reading comprehension strategies such as identifying main ideas, summarizing, generating questions, and clarifying. Presumably, only when they move beyond these more basic comprehension strategies in Grade 4, children are “ready and well-equipped” to acquire more deep-level reading strategies such as situation model processing and comprehension monitoring (Pressley, 2002; Stoeldraijer & Vernooy, 2007). Shifting to more advanced reading strategy use as early as possible is crucial for accurate acquisition of these deep-level reading strategies and to prevent later reading comprehension difficulties (Snow, 2002). It is important to realize, however, that it may take children more than one year to master these metacognitive strategies (Garner & Taylor, 1982). Therefore, training benefits are expected to not only accrue from Grade 3 to Grade 4, but also from Grade 4 up to Grade 5 (or even Grade 6), in spite of the probable existence of substantial individual differences in the rate of strategy acquisition (Baker, 1984; Markman, 1981). Future research should examine this issue more closely.

Related to this developmental explanation is the possibility that third-grade children in our sample were mature enough to learn and improve the trained strategies, but did not execute them spontaneously, as fourth graders did, during the inconsistency-detection test (Flavell, 1970). This would be consistent with research showing that children as young as third grade already have the prerequisite skills to detect inconsistencies in text passages and can improve these skills by inducing constructive reading processes (Rubman & Waters, 2000). In a similar vein, research has pointed out that we may not simply assume that primary school children spontaneously evaluate their understanding of text (Graesser, 2007; Markman, 1979; van der Schoot et al., 2012). Many of them have been found to have insufficient awareness and

knowledge of metacognitive reading strategies (e.g., Baker, 1985; Myers & Paris, 1978; Markman, 1979; Periodic Survey of Educational Level, 2013). However, applied comprehension instruction research conducted in classrooms has typically focused on the teaching of cognitive instead of metacognitive reading strategies (Cromley, 2005). The present study expands this research base by putting a strong focus on metacognitive awareness in the training and shows that even in this case the spontaneous use of monitoring strategies may not show up.

The second major finding of this study is that the effects of the inconsistency-detection training also transferred to general reading comprehension. Children who had followed the inconsistency-detection training showed a pretest to posttest improvement on our general reading comprehension measure. A significant improvement was not seen in children who had followed the school's regular reading comprehension curriculum. However, this result should be put in perspective, given that clearly different things were going on in the third and fourth grade. In Grade 3, the interacting effect of group and time on reading comprehension was influenced by the poor performance of the control group at posttest. Possibly, this drop in performance from pre- to posttest can be explained by the posttest version of the reading comprehension test (for Grade 3) being more difficult than the pretest version. Assuming that this is true, it certainly puts the relatively small gain in reading comprehension performance in the inconsistency-detection training group in a different, more positive light. In Grade 4, the interacting effect of group and time on reading comprehension was influenced by the poor performance of the inconsistency-detection training group at pretest. Due to the schools' preference that classes stay intact, we were not able to match individuals in the different groups on the basis of pretest levels of reading comprehension. Although this can be considered as a limitation of the study design, it also can serve as an explanation for the specific pattern of performance in fourth grade children. That is, children in the experimental group started off from a lower baseline of reading comprehension performance, but due to the inconsistency-detection training, they were able to catch up with their peers who had initially outperformed them at pretest.

Importantly, a transfer of strategies would fit the idea that higher-level complex strategies like comprehension monitoring that support the construction and updating of a coherent situation model during reading, underlie deep-level text

comprehension and that training these strategies fosters comprehension of texts in general (van der Schoot et al., 2010). The fact that for the inconsistency-detection training several measures were taken that would potentially facilitate the transfer of the acquired knowledge to more general reading comprehension is an important contribution. For example, the texts used for training contained several types of incoherence such as that the distance between the two pieces of conflicting text information was either “near” or “distant” (e.g., van der Schoot et al., 2012), and that inconsistencies appeared along various situational dimensions (e.g., Wassenburg et al., 2015). Taken together, the findings suggest that, despite a seemingly unaltered ability to detect inconsistencies in text, children in Grade 3 apparently did benefit from the explicit instruction in monitoring strategies and practice with the carefully designed texts, in comparison to children who had followed the regular curriculum. This is consistent with the view that children in Grade 3 are able to learn and improve the trained strategies, but only use them when it is evident to the children that the situation asks for them (Rubman & Waters, 2000). On the standardized test for general reading comprehension, for example, children answered questions about what they were reading and were encouraged to make inferences, which corresponded to elements of our training used to induce comprehension monitoring. In the inconsistency-detection test, on the other hand, children “naturally” read passages without being guided by questions about the text that supported them to actively make text-to-text or text-to-world connections. So, it seems that the inconsistency-detection test taps into comprehension monitoring processing more implicitly.

A third aspect we investigated was the extent to which reading motivation increased due to the inconsistency-detection training. Obviously, increasing children’s reading motivation is desirable because of its positive influence on reading effort, engagement, and reading growth (Morgan & Fuchs, 2007). It should be emphasized, however, that this study only intended to provide a first exploratory look at whether instruction in a situation model-focused reading strategy would lead to higher motivation. For example, it is unclear from the present study whether our one-item four-point Likert scale is valid and reliable enough to serve as a sole measure of reading motivation. Also, the scale prevented us from gathering more fine-grained information on the various aspects of reading motivation (Wigfield & Guthrie, 1997). Hence, using a more elaborate questionnaire is recommended for future research in

order to more thoroughly investigate the motivational effects of reading strategy trainings such as the one described in this study. Though the positive effects on motivation were not statistically reliable, the trends suggest that, regardless of grade level, children who had received the inconsistency-detection training enjoyed reading more after the training, whereas children who had received the control training did not. This finding is in line with evidence showing that metacognitive awareness and self-regulatory strategies, both of which are central components of our inconsistency-detection training, are related to both reading performance and learning engagement (Bandura, 1997; Guthrie et al., 2004; Schunk & Zimmerman, 1997).

In conclusion, the present study shows that the situation model framework provides a useful context for setting up a reading comprehension strategy training aimed at comprehension monitoring. By comparing this inconsistency-detection training to a control training (i.e., regular reading curriculum) that was as similar as possible in its form (e.g., lessons were held twice a week and involved small-group reading instruction combined with individual practice) and only differed in its content, we were able to rule out alternative explanations related to adverse side-effects like instruction time and/or alternative explanations based on natural development (Houtveen & van de Grift, 2007). Therefore, we conclude that teaching children metacognitive awareness through the explicit instruction of comprehension monitoring strategies during only a four-week period effectively improves strategies that are required for situation model updating, and can be transferred to novel texts.

