IMPROVING STUDENTS' NARRATIVE COMPREHENSION THROUGH A MULTIPLE STRATEGY APPROACH. EFFECTS OF DIALOGIC STRATEGY INSTRUCTION IN SECONDARY SCHOOL

MICHAEL TENGBERG*, CHRISTINA OLIN-SCHELLER* & ANNA LINDHOLM**

*Department of Educational Studies, Karlstad University ** Department of Swedish, University of Gothenburg*

Abstract

This article reports on an intervention study using a multiple strategy approach – called Dialogic Strategy Instruction (DSI) – to improve seventh grade students' narrative comprehension in terms of students' ability to generate inferences and develop interpretations of characters and events. DSI combines strategy instruction with dialogical principles for classroom discussion and response writing. After a 13-week intervention period, students in the intervention condition demonstrated significant improvement in generating inferences and interpretations compared to pre-test results. However, the average effect was small and the improvement was not significantly different relative to controls. An analysis of students with low pre-test scores showed that the intervention group had made significant and strong improvement relative to controls. Results for middle and high achievers were non-significant in this respect. Therefore, although DSI appears valuable for low achievers, results indicate that the instructional principles implemented may not offer instructional support for all children. Possible reasons for the absence of a general effect are discussed and directions for future research are suggested.

Keywords: dialogue, intervention study, narrative comprehension, reading, strategy instruction

1.   INTRODUCTION

There is ample evidence that both deepened understanding and appreciation of narrative texts are facilitated by a number of different comprehension strategy approaches (Snow, 2002; Janssen, Braaksma, & Couzijn, 2009; Souvignier & Mokhlesgerami, 2006). In the past decade, particular attention has been given to the testing of more complex set of comprehension strategies (c.f. Spörer, Brunstein, & Kieschke, 2009; Andreasson & Bråten, 2011). A practical advantage of multiple strategy approaches – given that they are proven effective – is that the scope and variety of multiple strategy activities make them suitable for long-term programs in authentic classrooms. The instructional mode of educational interventions need to deal not only with the social and cultural complexities that school settings make up (Deshler, Hock, Ihle, & Mark, 2011), but also with the broad variation of student motivations and cognitive aptitudes for learning and development. On the flip side, even with efficient multiple strategy programs, it is more or less impossible to determine which strategies contribute to significant improvement in students' reading comprehension (Graesser, 2007). Yet, no multiple strategy approach can be accounted for by summing up the discrete contributions made by any single strategy on its own. Rather, one would expect forms of interaction taking place between strategies that are combined in practice (Pearson, 2009).

In this study, a mix of strategies, selected in particular to promote comprehension of narrative texts, is combined with principles for dialogue and collaborative inquiry, on the one hand, and with principles for responsive writing, on the other hand. Drawing on the works of Vygotsky (1978) and Bakhtin (1981), Almasi, O'Flahavan, and Arya (2001), Nystrand, Gamoran, Kachur, and Prendergast (1997), Soter, Wilkinson, Murphy, Rudge, Reninger, and Edwards (2008), and others, all argue that open-ended, exploratory discussion may enable a public sharing of literary understanding and a co-construction of knowledge. Vygotsky's idea of the development of conceptual understanding through interaction with more capable peers, and Bakhtin's emphasis on the dialogic nature of understanding itself correspond well with the modelling of and interaction on text comprehension that are central to strategy instruction. In DSI, exploratory dialogue includes both discussions and extended response writing that is circulated among peers and teachers in order to make complex thinking visible and thereby subject to collaborative, interpretive dialogue.

The joint instructional focus on these three principles is labeled *Dialogic Strategy Instruction* (DSI) (Tengberg & Olin-Scheller, 2013). Compared with other well-studied approaches to strategy instruction, such as Reciprocal Teaching (RT) (Palinscar & Brown, 1984), DSI includes a larger number of strategies. These strategies are related to the reading of narrative text and aim specifically at supporting students' ability to generate inferences and to develop interpretations of characters and events in the stories they read. Strategies are also introduced one at the

time over a longer period. In this respect, DSI shares a number of traits with Transactional Strategy Instruction (TSI) (Pressley, Beard El-Dinary, Gaskins, Schuder, Bergman, Almasi, & Brown 1992). Unlike TSI, however, DSI does not relate specifically to reader-response theory, but takes dialogic theory as a starting point. In addition, the explicit and systematic integration of shared response-writing in DSI is not found in TSI.

The specific purpose of this research is to examine empirically to what extent DSI implemented over the course of a term (13 weeks) in authentic seventh grade classroom settings (students 13–14 years of age) may contribute to the improvement in students' narrative comprehension by supporting their ability to generate inferences and develop interpretations. In Rosenshine and Meister's (1994) review of studies on RT, a particular concern was that little attention had been paid to intervention effects in relation to student ability level. At the same time, a great deal of research in this field has emanated from the concern with struggling readers, and has indicated that explicit strategy instruction yields strong comprehension effects for students with learning difficulties (Brown, Pressley, Van Meter, & Schuder, 1996; Gersten, Fuchs, Williams, & Baker, 2001). In a recent meta-analysis of Scammacca, Roberts, Vaughn, and Stuebing (in press), the average effect size for standardized measures was found to be smaller (d = .21) than reported in previous studies, which could be explained by more rigorous measures, by differences in participant characteristics and also by general improvements in reading instruction in schools, which would impact on the control conditions in the intervention studies. Still, the authors of the study conclude that the "research base continues to show that teaching reading comprehension strategies to struggling readers in Grades 4 to 12 is beneficial" (Scammacca et al., in press, p. 18). It has also been suggested that strategies play a more important role in the early stages of reading development (Skaftun, 2011) or when decoding or comprehension proves difficult to the reader even at the later stages of development (Afflerbach, Pearson, & Paris, 2008). Against this background, our study also serves the purpose of determining whether DSI has a differential impact on secondary students with varying reading ability. Varying reading ability here refers to the range of normally achieving students found in ordinary classrooms, and does not include students receiving special education.

2.   THEORETICAL FRAMING OF THE INTERVENTION

2.1 Narrative comprehension

Research indicates that the particular strategies for understanding narrative text, with reference both to text structure and to the purpose for reading, are different from strategies for understanding expository text (Caldwell & Leslie, 2003; Graesser, Singer, & Trabasso, 1994; Janssen, Braaksma, & Rijlaarsdam, 2006; Oatley, 1999). In respect to narrative text, children tend to develop a basic knowledge

about structural patterns, or *story grammar*, early in life (Mandler & Johnson, 1977). However, the more complex aspects of narrative comprehension, including interpretation of character and events, generalization to real life, identification of themes and symbols etc., require a continued acculturation and formal training across the years of schooling (Applebee, Langer, Nystrand, & Gamoran, 2003; Janssen et al., 2006). In this study, narrative comprehension instruction is focused specifically on improving students' ability to generate inferences and develop interpretations of characters and events. These skills cut across the different classes of inferences discussed by Graesser et al. (1994) such as *causal antecedents* (bridging between actions or events in different passages of the text), *subordinate goals* (motivating character intensions or actions), and *thematic inferences* (integrating ideas, main points or moral in the text). In our study, these skills are treated as a unitary construct, and measured before and after the intervention through items taken from the Swedish national reading test (Swedish National Agency for Education, 2006, 2007, 2009, 2011). Set within a Swedish context of declining results on the PISA test, which is chiefly explained by a significant drop on the subscale *Integrate and interpret*,[1] special attention is naturally paid to improving interpretive reading (Fredriksson, 2012). Genereux and McKeough (2007) have demonstrated that 12-year-olds display "an emerging but clearly unrefined ability to combine two separate story morals" (p. 862) and that the 10–14 age group represents a period in which a clear shift from literal story meaning towards more interpretive forms of meaning making takes place. Although Genereux and McKeough analyze these conceptual changes in terms of social-psychological development, their results indicate that instructional activities on inference making and interpretations may be well suited to the cognitive development of 13–14-year-olds, which is the case in the present study.

2.2 Strategy instruction in the dialogic approach

This research rests on the assumption that the type of dialogic interaction in both speech and writing that is associated with greater improvement in narrative reading performance (Applebee et al., 2003; Nystrand et al., 1997; Newell, 1994; Wong, Kuperis, Jamieson, Keller, & Cull-Hewitt, 2002) may be combined with strategy instruction in order to form a broad, supportive environment for developing complex narrative comprehension. In this study, the instructional design is primarily based on theory of cognition and metacognition, with an emphasis on the mapping of learning processes from the perspective of comprehension strategies (Graesser, 2007; Griffith & Ruan, 2005). In practice, teachers' modelling, or think-aloud, of their reading process is assumed to enhance students' awareness of their own comprehension processes, and thereby improving their monitoring and regulation of comprehension (Israel & Massey, 2005; Snow, 2002). These models of reading

---

[1] *522 p to 494 p in the period 2000–2009.*

comprehension typically conceptualize metacognitive strategies such as *summarizing* (Armbruster, Anderson, & Ostertag, 1987), *making predictions* (Fielding, Anderson, & Pearson, 1990) and *asking questions* (Janssen et al., 2009), strategies which are used in DSI.

The selection of comprehension strategies included in the present study were 1) *summarizing*; 2) *making predictions*; 3) *evaluating*; 4) *visualizing*; 5) *questioning*; 6) *finding gaps and making inferences*; and 7) *comparing with other texts*. Nos. 1–5 and no. 7 are frequently reported in the research literature as effective strategies to improve both narrative and expository comprehension (Afflerbach & Cho, 2011; Block & Duffy, 2008; National Reading Panel, 2002; Pearson, 2009; Rosenshine & Meister, 1994). Finding gaps and making inferences are somewhat similar to the strategy often referred to as inference generation, which has been shown necessary for skilled reading (Elbro & Buch-Iversen, 2013; Thurlow & van den Broek, 1997). However, teaching this strategy in DSI also means promoting metacognitive awareness (cf. Donndelinger, 2005) about inferencing and interpretation by explicitly demonstrating and discussing the function of structured gaps in fictional text (cf. Iser, 1978).

Although a number of studies have already demonstrated the effectiveness of strategy instruction, a major challenge lies in the classroom implementation of multiple strategy approaches (National Reading Panel, 2002) and the context-sensitive, but necessary, integration of strategy instruction with other classroom activities (Pearson, 2009). One of the more viable turns taken in research on the learning and teaching of text comprehension in the last decade is the emphasis on dialogic approaches (Almasi et al., 2001; Soter et al., 2008; Wilkinson & Son, 2011). Although attention to the role of classroom dialogue goes all the way back to the 1860s (Nystrand, 2006), research on how structured discussions may 'scaffold' (Wood, Bruner, & Ross, 1976) higher-order thinking about texts has provided a number of new insights. For example, allowing conflicting perspectives to meet, open-ended (or authentic) questioning and dialogue may support student reflections on alternative interpretations and encourage them to elaborate on their understandings in ways that foster narrative comprehension (Applebee et al., 2003; Murphy, Wilkinson, Soter, Hennessey, & Alexander, 2009; Nystrand, 2006). In addition, both large-scale, observational studies (c.f. Taylor, Pearson, Clark, & Walpole, 1999) and small-scale, experimental studies (Saunders & Goldenberg, 1999) have linked higher-level discussions to better performance on measures of reading comprehension. An important insight in this respect is that it is not the amount of talk but the kind of talk that promotes comprehension (Murphy et al., 2009).

The sociocultural perspective on learning as a culturally embedded process mediated through language and interaction, as explained by Vygotsky (1986), provides a link between comprehension strategy approaches and dialogic approaches. In multiple strategy instruction implemented over longer periods, students are assumed to benefit from the interaction with peers and teachers by internalizing the ways and models of meaning making that are made explicit there. As the teacher

prompts a joint effort (in groups or in whole-class) to enhance everyone's understanding of the text by applying explicitly taught comprehension strategies, students are offered an ample display of each strategy in action (Pressley et al., 1992). The role of interaction with more capable peers (or teachers) provides a useful framework for investigating how comprehension of texts is scaffolded through collective reasoning and meaning making. By displaying their particular understanding, learners bring forth perspectives that their peers may use as temporary standpoints in their own gradually shifting reading experience. Evidence from research also verify that less capable readers are able to use the discourse of discussion as scaffolds for their own thinking (Reninger, 2007) and that strategy interventions in general are effective both for younger and older struggling readers (Edmonds, Vaughn, Wexler, Reutebuch, Cable, Klingler Tackett, & Wick, 2009; Kamil, 2004; Snow, 2002).

Yet, while scaffolding conceptually emphasizes the role of organizing support for learners, e.g. by providing context, language or mental models, learning and development are also driven by challenge and negotiation. In concrete processes of communication, cooperation and conflict are intrinsically interwoven (Nystrand et al., 1997). Discussions about texts are therefore essentially not about "sharing meanings", but rather about jointly constructing baselines for individual understandings (Tengberg, 2011). Drawing on Bakhtin (1981), speech and thinking is fundamentally dialogic, meaning that the way we reason always involves a response to prior experience, a response to the reasoning provided by others. This view of interaction and understanding has proven particularly useful in empirical research on reading and writing instruction (Marshall, Smagorinsky, & Smith, 1995; Nystrand & Gamoran, 1991). The practical, educational applications of this theory connect with the theoretical implications about the benefit of strategy instruction in that they stress the importance of making models of thinking visible to peers. They unite in the belief that comprehension of narrative text may be improved by joint questioning of the text itself and of explicit-made ways of understanding or interpreting the text.

Dialogue in this study refers not only to peer or whole-class discussions but also to the sharing of written responses to the narrative texts. The inclusion of writing assignments as a means of developing narrative understanding is critical. A number of studies have demonstrated that written responses aimed at acquiring a holistic understanding of narratives may have a significant impact on students' literary responses (Marshall, 1987; Newell, 1994; Wong et al., 2002). In a meta-analysis of nine experiments in which students wrote extended interpretive responses to text, Graham and Hebert (2010) found an average effect size of 0.77 on students' reading comprehension. In a previous study of DSI, students' written interpretive responses were exchanged and discussed between peers and elaborated upon by the teacher in whole-class discussion, thus providing appropriate formative feed-back on the students' narrative comprehension (Tengberg & Olin-Scheller, 2013). Overall intervention effects were statistically significant compared to control groups.

Based on the theoretical principles presented above, the current study aims at determining

1) to what extent DSI may contribute to the improvement in students' narrative comprehension by supporting their ability to generate inferences and develop interpretations, and

2) whether DSI has a differential impact on students' narrative comprehension in relation to their levels of reading achievement at the outset.

3. METHOD

DSI was implemented in ten classes and taught from September through December. Test results for narrative comprehension were compared with results in twelve control classes. In order to cause minimal disturbance to the authenticity of the setting, students were not randomized into groups but nested within their respective classes, which were in turn randomly assigned either to the intervention or to the control condition.

3.1 Participants

Twenty-one teachers (19 female and 2 male), all educated teachers of Swedish, with long experience of secondary school teaching, volunteered to participate. The intervention group consisted of 93 girls and 84 boys from nine 7th grade classes and one 8th grade class.[2] The control group consisted of 106 girls and 84 boys from twelve 7th grade classes. Classes were situated at seven different schools in four separate small-sized cities and one medium-sized city. Classes were chosen based on two criteria: (a) representing similar averages of school merits and passing rates at national tests in reading comprehension;[3] and (b) practical availability and willingness to participate (self-selection).

All students were informed of the study's purpose. Verbal consent was gathered from both teachers and students. Since students were under 15 years of age, written consent was also collected from all students' parents. Participants were informed that they had the right to withdraw their consent at any point, and that the data collected would be treated confidentially and used for research purpose only.

With respect to social and cultural background, the sample is comparatively homogenous. The vast majority of students are native Swedish speakers with Swe-

---

[2] *Although being older than the rest of the participants, the 8th grade class included in the sample had scored below expected reading proficiency level on local tests and their results on the pre-test of the study showed that their reading abilities were comparable to the other 7th grade classes.*
[3] *Total school merit averages over the period 2007–2011 was 212.4 for intervention schools and 208.3 for control schools. Passing rates at national tests in reading comprehension for the same period was 90.5% for intervention schools and 88.8% for control schools.*

dish born parents. Eighteen students in the intervention group and 25 students in the control group report a first language other than Swedish, although none of the participants took Swedish as a second language instead of Swedish. The average class size was 22.0 students in the intervention groups and 22.4 students in the control groups.

3.2 Design and procedure

The study uses a pre-test/post-test design to determine the effect of DSI on 7th graders narrative comprehension. Since the design does not include individual random assignment to conditions, the statistical analyses will control for pre-test differences.

Professional training for intervention teachers, including theoretical baselines and plans for classroom procedures, were provided in three 3-hr seminars before the intervention started and in an additional three 3-hr seminars during the intervention. Plans for classroom procedures included detailed written instructions and were presented and discussed in the teacher-researcher group. After minor revisions, these instructions served as scripts for lesson plans. From early September until Christmas break (13 weeks altogether), DSI was implemented 2 lessons (40–50 minutes each) per week in the ten intervention groups. Control group teachers were contacted only in order to administer pre- and post-tests.

Pre-tests were taken by all students in late August or early September. Post-tests were similarly taken by all students in two weeks in January. Both tests were administered in regular lesson time, and students were given a maximum of 80 minutes to complete it. All tests were administered by either first, second or third author. Observations were conducted in the intervention groups during the second half of the intervention period in order to establish the fidelity of implementation of DSI in each classroom.

3.3 Intervention

The instructional frame designed for the intervention was based on nine short stories/passages from novels, eight out of which were originally written in Swedish and one originally written in English and read in translation (see list in References). The stories were chosen to match the strategies, or more precisely to make the strategies called for during the process of comprehension. A couple of stories, however, were matched with several strategies, and also bridged sequentially between them (see Appendix for full description). For instance, the work on Axelsson's short story "Across the track" (2002) linked the end of working with the strategy 'making predictions' to the introduction of the strategy 'evaluating'. Altogether, the nine narratives formed the outset from which DSI was implemented in the classrooms according to the theoretical principles discussed above.

3.3.1   Introducing comprehension strategies

Comprehension strategies (*summarizing*; *making predictions*; *evaluating*; *visualizing*; *questioning*; *finding gaps and making inferences*; *comparing with other texts*) were introduced one at the time in the mentioned order. Each story was chosen to exemplify the need and benefit of utilizing a specific strategy, although, as the students gradually became familiar with more strategies, they were prompted to continue making use of the strategies they had already learned. The introduction of strategies was therefore intimately related to the structure and nature of the texts. Nonetheless, a core procedure defined by five successive steps shaped the explicit instruction of each strategy. Inspired by Duke and Pearson (2002), who present a five component model for introducing comprehension strategies and gradually releasing responsibility to students for making use of strategies in action, we applied the following five steps in order to integrate reading, writing and discussion: 1) *Introducing the strategy, what to be gained from using it and how to use it*; 2) *Teacher modelling the strategy while reading aloud*; 3) *Teacher and students using the strategy together in whole-class*; 4) *Students using the strategy individually, in pairs and in small groups*; and 5) *Students and teacher using the strategy together in whole-class discussion*.

3.3.2   Interpretive literature discussions

Dialogues in whole-class, pairs or groups were set as pre-reading, during-reading and post-reading activities in order to support strategy learning, and to scaffold continuous inferencing and interpretive thinking. Whole-class discussions combined a focus on strategy introduction (meaning that teachers were modelling comprehension strategies) with an ambition to follow up on what the students in their written responses had noted as difficult, interesting or evocative in the stories. A pedagogical principle for the teachers' orchestration of the discussions was to support students' active participation. Teachers also challenged students' interpretations of texts rather than prescribed appropriate interpretations. By drawing attention to differences in the interpretations of texts, discussions were intended to raise students' awareness of their own active part in the interpretive process. Likewise, by emphasizing the various ways in which readers apply comprehension strategies, and by comparing the inferential process for different readers, instruction aimed at alerting students to their own active inference generation. Furthermore, discussions aimed at being *exploratory* (Mercer, 1996), i.e., to have students engage both critically and constructively with one another's ideas, both building on one another's ideas and constructively challenging them. Pairs and groups were formed with regard to students' reading proficiency in order to let less skilled readers benefit from interaction with more capable peers, in accordance with the Vygotskian framework. Pair-sessions were used to discuss written, interpretive responses, in which specific strategies were applied. These were normally followed

by group discussions, in which two pairs formed one group. The make-up of pairs and groups shifted continuously.

3.3.3   Writing assignments

Writing assignments were designed to elicit higher-level thinking, first of all by providing students with opportunities to formulate literary interpretations in solitude and tranquility, allowing time to gather ideas, questions and personal reactions to bring to discussions with their peers and teachers. For example, students would be asked to reflect on characters' responsibilities and actions in story events. They would also be encouraged to take a personal stance on ethical and moral issues raised in the text. Here is an example of a question on the passage from *A Comedian Growing Up* (Gardell, 1992) in relation to strategy no. 3 (evaluating). In the text, an eleven-year-old boy is abused by his classmates on his way home from school. The text provides a number of clues regarding the personal responsibilities for letting this happen.

> In your opinion, who is/are responsible for the abuse of Thomas? Write an extended response where you describe how you reflect on this issue. Be sure to formulate arguments that support your opinion. Make use of previous knowledge and experiences, but also make sure that your arguments have a clear backing in the text.

In this way, writing assignments supported students' active inferencing and gave way to a more substantial use of their previous experiences in order to comprehend the stories. This type of pre-discussion activity also served to enhance students' ability to participate actively and substantially by challenging them to develop their literary responses. Challenge in this case refers to questions that required students to respond to texts from more than a single perspective, questions targeting critical issues in texts and called for holistic interpretations and elaborated answers.

Writing assignments were usually given as a development of a topic discussed in whole class. Students were prompted to produce personal or analytical, interpretive responses to a story or an aspect of the story, and these responses were then exchanged and discussed in pairs and eventually in groups of four. Students were encouraged to explore the nature of similarities and differences between their responses and to use comprehension strategies to expand on their story understanding. Findings from these discussions were then analyzed together in whole class.

3.4 Control groups

Control teachers were not trained at DSI. Nor were they requested to adopt any other type of reading instruction. Instead, they were encouraged to teach in the way they were used to and to make no significant changes. Yet, the awareness of the fact that their students' progress over the term was to be measured may have

motivated a higher ambition in their reading instruction. However, this obviously goes for the intervention teachers too.

Control teachers were provided with the same compilation of short stories that were used by the intervention groups, and they were asked to teach as many of them as possible in any way they pleased.

During the term, teachers in the control condition also kept journals of their reading instruction. When using any of the short stories, they were to jot down a few lines, describing in what way these texts were read and taught in class. From these notes, we were able to get an overview of the reading instruction with which the intervention was to be compared. To this end we also conducted group interviews with 8 of the 12 control condition teachers. In regard to time spent on reading instruction in the control condition, the average of teachers' estimations was that two out of three lessons a week (each lesson 50 minutes long) were devoted to reading instruction. Therefore, it is fair to say that an equal amount of time was spent on narrative comprehension instruction in intervention groups and control groups.

3.5 Narrative comprehension measures

While dialogically-oriented strategy approaches have proven effective on many research-developed reading tests as well as on standardized reading inventories (Rosenshine & Meister, 1994), we wished to know whether the training effects of a dialogic strategy intervention would also impact on measures used in the national curriculum assessment system. Therefore, in testing narrative comprehension, we compiled two tests (X and Y) using texts and test items from previous Swedish national tests of reading comprehension (Swedish National Agency for Education, 2006, 2007, 2009, 2011). Each test included two short stories (one 500–600 words long and one 1100–1200 words) and 7 items (per test) on inferential and interpretive level. Response formats were mainly constructed-response, ranging from sentence to paragraph in length. While most reading comprehension tests mainly use multiple-choice items, the constructed-response items in the national test were expected to match the format of student responses in the instructional setting (discussion and writing). In order to illustrate the character of these items, we provide an example taken from one of the longer texts included, which was an extract from the novel *The Kite Runner* (Hosseini, 2003), pp. 16–18. This text (and item) was included in the Swedish National reading test in 2009.

Explain   why   Amir   wrote   his   first   story   this   particular   evening.

The scoring rubric to this item requires that students relate the appreciation Amir receives for making up a story to his growing confidence in writing/eagerness to write. In this way the items in the test targeted students' ability to generate infer-

ences and to develop interpretations of characters and events in the stories. Both tests also included items in which students were asked to provide character descriptions or to make inferences about ideas and plot elements that were not explicitly mentioned in the text, and to interpret figurative language, to explain motives for characters' actions, and to explain events related to the story themes. An example of items requesting inferences across several classes according to Graesser et al. (1994) (also related to the extract from The Kite Runner) is the following:

> Explain the ending of Amir's story about the man and the magic cup. The consequence of the man's actions should be included in your response.

To receive full credit, student responses are supposed to link the man's actions both to intentions and consequences and to emotions and story moral.

In this repeated measures study, we used an ABBA design, meaning that half of the students (each class was split) received the X-test for pre-test and Y-test for post-test, while the other half received the tests in the reversed order (Y for pre-test and X for post-test). Items in the two tests were estimated by qualitative analysis to target the same type of comprehension. An independent t-test indicated that there was no significant difference between mean pre-test scores for the X-test (M = 6.84, SD = 3.46) and mean pre-test scores for the Y-test (M = 6.96, SD = 3.64; t (365) = .336, p = .74). The two versions were therefore assessed to be equally difficult. The ABBA design is useful in order to eliminate practice effects, which may occur if the same test is used for both pre- and post-test, as well as to eliminate order effects caused by minor differences between the tests.

Students were instructed to deal with one text at the time, although they were able to return to a text as many times as they liked. They were also instructed to remain at their desk for 60 minutes, after which they could leave the classroom if they had finished (80 minutes were afforded altogether).

Items were scored on a scale 0–1, 0–2 or 0–3 (depending on the complexity of the task) using the national test scoring guidelines including detailed scoring rubrics for each item (Swedish National Agency for Education, 2006, 2007, 2009, 2011). Thus, some items would generate a maximum of one point for correct solution while others could generate up to three points for a correct solution. Maximum score for each test was 15 points. All items were blind-scored by three trained raters (the authors of the article). Cohen's kappa for inter-rater agreement were .86, .81 and .84 on pre-test and .95, .93 and .94 on post-test (all p < .001), which suggests a very high level of agreement between raters.

Statistical analyses to answer the first research question included paired-samples t-testing and one-way between-groups analysis of covariance (ANCOVA). For the second research question, investigating whether students' ability levels would work as a moderator variable, a two-way between-groups analysis of covariance (ANCOVA) was used, for which all cases were grouped based on pre-test results (low achievers = pre-test scores 0–4; middle achievers = pre-test scores 5–9; high achievers = pre-test scores 10–15). This was then complemented by ANCOVAs for groups on each achievement level in order to obtain comparative effect measures for low achievers, middle achievers and high achievers respectively. The reason for choosing a cut-off at 5 points was to be able to split the scale of possible test results in three ways. In a single-group design, an alternative procedure might have been to consider the 25% lowest scores, the 50% in the middle, and the 25% at the top. In this case, however, such a procedure might have yielded between-groups mean differences, in which case the comparison would have included students of different rather than of similar ability levels.

3.6 Observations and interviews

To ensure fidelity of the intervention, teachers and researchers met three times during the intervention period in order to discuss and make sequential adjustments to the implementation of DSI. Furthermore, all intervention classes were observed twice (once in the middle of the period and once in the end). On these occasions, teachers were interviewed individually. Observations included time, type of activity, organization, content, material and products. The observations were not conducted for systematic or in-depth analysis of classroom procedures, but rather to determine whether or not instruction in the intervention classroom aligned with the principles of DSI. Observation protocols (covering the areas mentioned above) also included a commentary on learning objectives of the lesson and (if applicable) the occurrence of unexpected observations. Interviews similarly targeted the implementation of DSI and served as a supplement to the observations. Although control groups were not observed, eight of the control teachers were interviewed based on the journal notes they made about their reading instruction. Obviously, observational and interview data are not directly comparable, yet these data provided an overview of instructional similarities and differences between the two conditions.

4.   RESULTS

4.1 Comparison between intervention and control condition

From the documentation of the two teaching conditions, we were able to make a brief comparison as to the quantity and quality of reading instruction conducted in the period.

Control teachers reported having read at least six of the nine texts. They also reported having employed a number of different types of instruction, including silent reading, group discussions, short-answer questions, writing descriptions of characters and setting, writing their own short stories, analyzing themes and motifs etc. Three of the control teachers also reported having utilized comprehension strategies (summarizing, making predictions and questioning) in 3–5 lessons altogether. Some of these activities may, thus, resemble what was going on in the intervention instruction, which is hardly surprising, but these are single examples collected from all the twelve control teachers. Contrary to the intervention condition, these activities were not brought together in any single classroom, nor were discussions or comprehension strategies, according to the interviews, implemented as systematically as they were in the intervention classrooms. Considering that the entire period ran over 13 weeks, 2 lessons per week on an average, we estimate that there is a clear boundary between intervention and control conditions.

Based on regular meetings with the intervention teachers as well as on classroom observations, we were able to confirm frequent classroom dialogue and systematic instruction of comprehension strategies. As far as we can determine, all texts were read and treated in accordance with the provided lesson plan. Similarly, all strategies were introduced and the same writing assignments and type of discussions were conducted in each classroom. The modeling of comprehension strategies was performed vividly and explicitly and in such a way that students could participate with relevant questions and dared to elaborate on some of their own uses of a particular strategy. Dialogues in pairs, groups and whole class were frequent according to both interviews and observations. Writing assignments were continuously reviewed and discussed in the researcher-teacher-group, and generally found appropriate for the purpose. Obviously, social and cultural differences between the groups meant that the instruction was not identical in terms of discussion topics, intertextual or experiential references etc. However, the observations, complemented by the interviews, verified that DSI was being successfully implemented in the intervention group.

In sum, to a significant extent the intervention and control conditions treated the same type of reading material during the term, although they did so in ways that were systematically different. Therefore, the comparison between the two conditions can be regarded as relevant according to the study rationale.

4.2 Does DSI contribute to a significant improvement in narrative comprehension?

The first research question was to find out to what extent DSI would contribute to the improvement in students' narrative comprehension by supporting their ability to generate inferences and develop interpretations. Results on the narrative comprehension tests were analyzed in order to compare whether the improvement rates of the intervention group were significantly larger than any improvements made by the control group. Descriptive test statistics are reported in Table 1. All

scores were approximately normally distributed with all values for skewness and kurtosis within -1 to 1 and thus appropriate for parametric analysis.

Table 1. Test scores for all subjects on pre-test and post-test, scale 0–15

Measure	Pre-test		Post-test		Within-groups diff. (partial η^2)	Between-groups diff. (partial η^2)
	M	SD	M	SD		
Intervention (N=177)	6.70	3.50	7.24	3.48	0.04**	ns
Control (N=190)	7.08	3.58	7.38	3.60	ns	

* significant at < .05
** significant at < .01

There was a small but non-significant (p = .307, 2-tailed) difference between pre-test means in intervention and control groups. First, we looked for within-groups differences, and found a statistically significant progress for the intervention group from pre-test (M = 6.70, SD = 3.50) to post-test (M = 7.24, SD = 3.48), t (176) = -2.64, p = .009 (2-tailed). Eta squared statistics (.04) indicated a small to moderate effect size according to Cohen's (1988) guidelines. According to Standard Mean Difference (SMD) statistics, the effect was small (Cohen's d = .15). For the control group, there was also a small but non-significant progress from pre-test (M = 7.08, SD = 3.58) to post-test (M = 7.38, SD = 3.60), t (189) = -1.48, p = .14 (2-tailed). The indication of a training effect for the intervention condition, however, does not confirm that the size of improvement is significantly larger than the improvement in the control group.

To evaluate the effectiveness of the intervention, a one-way ANCOVA was conducted using pre-test scores as covariates, condition (intervention and control) as grouping variable and post-test scores as dependent variable. Preliminary checks were made to ensure that assumptions of linearity, homogeneity of variances and homogeneity of regression slopes were not violated. After adjustment by covariates, the analysis showed that the difference of improvement between the two conditions at the post-test was not statistically significant and that the effect size was close to zero [F(1, 364) =.213, p = .65, partial η^2 = .001, Cohens d = .07]. With regard to test score means for all students, the statistical analysis, consequently, indicates that DSI does not improve students' narrative comprehension significantly better than ordinary reading instruction does.

4.3 Does DSI have a differential impact on students' narrative comprehension in relation to their levels of reading achievement?

The second research question was to find out whether DSI would have a differential impact on students with different levels of reading achievement. Reading achievement levels were defined by pre-test results (low achievers = pre-test scores 0–4; middle achievers = pre-test scores 5–9; high achievers = pre-test scores 10–15). A 2 by 2 ANCOVA was conducted to assess the effectiveness of the intervention for students of different achievement levels. Independent variables were, thus, condition (intervention, control) and achievement level. After adjusting for pre-test scores, there was a significant interaction effect [$F(2, 360) = 3.33$, $p = .04$] although with only a small overall effect size (partial $\eta^2 = .02$). These results suggest that students of various levels of reading proficiency responded differently to the intervention.

Table 2. Test scores for all subjects on pre-test and post-test, scale 0–15

Measure	Pre-test		Post-test		Within-groups diff. (partial η^2)	Between-groups diff. (partial η^2)
	M	SD	M	SD		
Low achievers						
Control (N=49)	2.53	1.21	3.65	2.41	0.19**	
Intervention (N=50)	2.52	1.30	4.54	2.49	0.43***	
						.04*
Middle achievers						
Control (N=91)	7.07	1.51	7.86	2.99	0.08**	
Intervention (N=82)	6.71	1.39	7.02	2.77	ns	
						ns
High achievers						
Control (N=50)	11.56	1.63	10.16	2.39	0.22***	
Intervention (N=44)	11.36	1.35	10.66	2.75	ns	ns

* significant at < .05
** significant at < .01
*** significant at < .001

Descriptive statistics for different proficiency levels are reported in Table 2. As shown, the improvement for low achievers was statistically significant with large effect sizes in both the intervention [$F(1, 49) = 37.73$, $p=.000$, partial $\eta^2 = .43$, Cohen's $d = 1.07$] and in the control condition [$F(1, 48) = 11.26$, $p = .002$, partial $\eta^2 = .19$, Cohen's $d = .62$]. Yet, results from an ANCOVA reveal that the improvement is significantly larger in the intervention condition [$F(1, 99) = 3.71$, $p = .03$, partial $\eta^2 = $

.04, Cohen's *d* = .46], suggesting that for low-achieving readers DSI is particularly beneficial for improving narrative comprehension.

For middle achievers, the mean score for the control group was actually higher on post-test than for the intervention group, even after controlling by covariates, although the difference in gains was not statistically significant [$F(1, 170) = 1.92$, $p = .17$, 2-tailed, partial $\eta^2 = .01$, Cohen's *d* = .21]. For high achievers, though, the intervention group scored higher than the control group, even after controlling by covariates, yet neither here was the difference in gains statistically significant [$F(1, 91) = 1.41$, $p = .12$, partial $\eta^2 = .02$, Cohen's *d* = .33]. It is worth noting, however, that the high achievers in both conditions scored lower on post-test than on pretests. This drop was statistically significant for the controls but not for the intervention groups.

5. DISCUSSION

The results provide a fairly straightforward answer to both research questions. Although within-group measurements indicated a small but statistically significant impact from the intervention on students' ability to generate inferences and develop interpretations of stories, the study does not provide evidence that DSI is more efficient in this respect than other instructional models, at least not for students in general. First of all, the improvement in the intervention condition was not statistically significant when compared with the improvement in the control condition. Second, it was revealed that a large portion of the improvement was accounted for by low achievers, while for middle and high achievers the intervention had little or no effect. For low achievers, the improvement on the narrative comprehension measure was significantly larger in the intervention condition than in the control condition. Thus, it seems that DSI has a differential impact on students' narrative comprehension related to their levels of reading achievement at the outset. For less skilled readers, DSI turns out to be more beneficial than the ordinary reading instruction used in the control condition.

This is valuable information since interventions that promote significant improvement in struggling readers' inferencing and interpreting are dearly needed. The fact that the intervention turns out to be more effective for low achievers may not be surprising in itself. The result is consistent with some of the findings in previous research (Brown et al., 1996; Dole, Nokes, & Drits, 2009; Edmonds et al., 2009; Reninger, 2007). With regard to the theoretical foundation on which these interventions were developed in the first place, it should rather be expected that low achievers respond particularly well both to explicit strategy instruction and to structured text discussions. Favored comprehension strategies are mainly drawn from think-aloud interviews with proficient readers aimed at displaying the strategies they implicitly bring to use in skilled reading (Duke & Pearson, 2002; Pressley & Afflerbach, 1995). It is also suggested that strategies are essential when comprehension is challenging, while for more capable readers, the necessary skills for

smooth comprehension are already available, and the need for deliberate and controlled strategy use is much smaller (Afflerbach et al., 2008; Skaftun, 2011). Skaftun (2011) even argues, with reference to recent theoretical development, that competent reading is something more, and something different, than automatized strategy processing. Instead, the move towards competent, or proficient, reading is "a shift away from dependency on rules, maxims and plan-making towards holistic recognition and discrimination." (Skaftun, 2011, p. 140) This transition means that understanding, e.g. by way of inferencing, comes to rely more on previous reading experience than on high-speed processing of already-learned strategies. If this is true, then instruction that turns the attention "back to" comprehension by aid of conscious strategy use may potentially have little to offer the more competent readers. In this study, however, texts were deliberately chosen to provide a challenge for all students. Discussions and writing prompts were expected to contribute a broad scope of instruction to fit the different levels of reading ability. Therefore, although there may be theoretical aspects from which the value of strategy instruction for already-competent readers may be questioned, we would still expect high achievers to draw at least some benefit from the intervention condition instruction.

From a methodological perspective, one should consider whether the differential impact relates to issues of test construction, e.g., if the lack of effect on high achievers is related to ceiling effects in the test. Those who perform poorly on a pre-test naturally have more room for improvement than those who perform well. In this case, pre-test mean score for high achievers was not very close to maximum (see Table 2), suggesting that ceiling effects may after all be a less favorable explanation, since there was room for more improvement even in this group. On the other hand, the same pattern of differential impact is revealed by the control groups' results. This suggests either that high achievers draw less benefit from the instruction in both conditions or that the test does not adequately identify the type of progress made by students who are already skilled readers. Further research on this particular topic would be valuable.

Problems related to the test can take many forms. First of all, the test may not contain enough items to allow for an adequate discrimination between students' performances. Secondly, there may be a mismatch between what was taught and what was measured. Rosenshine and Meister (1994) found larger effect sizes for experimenter-developed tests than for standardized comprehension tests. One of the explanations suggested was that the former type of test normally contains passages similar to those practiced during the intervention, whereas the latter type contains a larger variety of passages. In our study, the texts practiced during the intervention were in many respects similar to those used in the test. Items in the test, on the other hand, differed more from the writing assignments and from the formative feed-back provided in the intervention. These items were taken from previous national reading comprehension tests, and were thus not experimenter-developed, although the selection of items was made by the research group. Therefore, the test used should be categorized neither as experimenter-developed nor as

standardized. The intervention was designed to promote the kind of reading skills targeted in the test, but the actual learning effects might have been different from what was expected. How could this be? Perhaps DSI, because of its focus on both strategy introduction and dialogicity in talk and writing, promoted an interpretive approach to narrative texts that may be characterized as open-ended. Attending to dialogic principles means for one thing that texts (and readers) are perceived as layered with meaning in several dimensions simultaneously. It is a widely accepted notion that literary reading is not about assigning texts with a single meaning or interpretation, but to discern and appreciate the multiplicity of meaning in litera-ture (Machor & Goldstein, 2001). Having encouraged students to seek for multiplic-ity of meaning in narrative texts may to a certain extent have run counter to the perspective brought to use in the national test items, in which all questions have single, predetermined answers. In the test, reasoning in itself provides students with no extra points unless the interpretation they reach is in line with the one de-fined by the scoring rubric (for a more developed analysis of the Swedish national test items, see Tengberg, 2014). Thus, while the response format of test items was expected to match key components of the instructional setting, other factors, re-lated to the perception of what a reasonable interpretation is, may have affected the scores in an opposite direction. The particular challenge of targeting literary reading in correspondence with literary theory is discussed further by Frederking, Henschel, Meier, Roick, Stanat, & Dickhäuser, (2012). If, however, this line of ex-planation to the lack of intervention effect could be validated by the use of a dif-ferent test instrument showing a larger effect, this would, on the other hand, indi-cate that DSI generates only a very limited transfer effect.

Taking the above into consideration, the absence of a general intervention ef-fect is still somewhat puzzling. Previous studies on similar multiple strategy ap-proaches show positive results (Andreassen & Bråten, 2011; Spörer et al., 2009), although these approaches have not targeted narrative comprehension in particu-lar. A possible explanation relates to the assumed ordinariness of the control condi-tion. The teachers were chosen by way of self-selection and volunteered to have their students' progress over a term measured by researchers without receiving any preparation. Teachers who are willing to participate under such conditions are most likely skilled and ambitious teachers of Swedish who take a serious interest in their students' reading development. They might not therefore provide an entirely accurate representation of the average seventh grade reading instruction in Swe-den. In addition, as pointed out above, Scammacca et al. (in press) suggest that the decrease of intervention effects found in recent meta-analyses may at least partly be a consequence of a general instructional improvement in schools. The control condition may in other words also contain a relevant instructional support for stu-dents' reading comprehension development. A limitation of the study is thus that we know too little about the control condition instruction in order to provide a well-informed evaluation on this point. This should also be a point to consider in future research.

Obviously, since observations and interviews indicated that DSI was satisfactorily implemented, the low effects on middle and high achievers may also imply a potential weakness of the model. Perhaps seven different comprehension strategies taught over the course of one semester is a challenging task for both teachers and students. The complexity of the program may require more teacher preparation than was afforded in this case. And a useful management of comprehension strategies to support the understanding of narrative text may require more time for student practice. It has been argued that, in order to train higher-level skills such as narrative comprehension, the duration of the intervention is critical (Gersten et al., 2001; Perfetti, Landi, & Oakhill, 2005; Pressley et al., 1992). The 13-week- long intervention in this study might after all be too short a period to achieve significant effects in this respect.

In summary, this study adds to previous research that low-achieving readers seem to benefit from dialogic strategy instruction – combining interpretive literature discussions with strategy instruction and challenging response writing – as they improved significantly in their ability to generate inferences and develop interpretations. The model may thus provide a good alternative method for working with struggling readers' narrative comprehension in the secondary classroom and should preferably be implemented over an extended period of time, at least for a whole term. In order to enhance the performance of high achievers', dialogic strategy instruction may provide a foundation but need to be combined with other activities that have previously been proven effective for the development of high achievers' narrative comprehension.

6.   CONCLUDING REMARKS

Whether interventions of this type really offer less (or no) support for high-achieving readers' narrative comprehension remains uncertain. More details about improvement rates for high-achieving readers in this type of intervention would be valuable, since particularly strong effects for low achievers may otherwise imply a correspondingly lower effect for middle and high achievers. The indications of the present study need to be corroborated by subsequent studies. In addition to this, future reading research ought to work specifically on instructional designs that emphasize the importance of addressing student diversity not only in terms of reading motivation and reading preferences, but also in terms of reading proficiency.

ACKNOWLEDGEMENTS

REFERENCE LIST

Afflerbach, P. P., & Cho, B. (2011). The classroom assessment of reading. In M. L. Kamil, P. D. Pearson, E. Birr Moje, & P. P. Afflerbach (Eds.). *Handbook of reading research, vol. IV* (pp. 487–514). New York: Routledge.

Afflerbach, P. P., Pearson, P. D., & Paris, S. G. (2008). Clarifying differences between reading skills and reading strategies. *The Reading Teacher, 61*(5), 364–373. doi:10.1598/RT.61.5.1

Almasi, J. F., O'Flahavan, J. F., & Arya, P. (2001). A comparative analysis of student and teacher development in more and less proficient discussions of literature. *Reading Research Quarterly, 36*(2), 96–120. doi:10.1598/RRQ.36.2.1

Andreassen, R., & Bråten, I. (2011). Implementation of explicit reading comprehension instruction in fifth-grade classrooms. *Learning and Instruction, 21*(4), 520–537. doi:10.1016/j.learninstruc.2010.08.003

Applebee, A. N., Langer, J., Nystrand, M., & Gamoran, A. (2003). Discussion-based approaches to developing understanding: Classroom instruction and student performance in middle and high school English. *American Educational Research Journal, 40*(3); 685–730. doi:10.3102/00028312040003685

Armbruster, B. B., Anderson, T. H., & Ostertag, J. (1987). Does text structure/summarization instruction facilitate learning from expository text? *Reading Research Quarterly, 22*(3), 331–346. doi:10.2307/747972

Bakhtin, M. M. (1981). *The dialogic imagination. Four essays by M. M. Bakhtin*. (M. Holquist Ed.). Austin, TX: University of Texas Press.

Block, C. C., & Duffy, G. G. (2008). Research on teaching comprehension. In C. C. Block & S. R. Parris (Eds.), *Comprehension instruction: Research-based best practices, Vol. 2* (pp. 19–37). New York: Guilford Press.

Brown, R., Pressley, M., Van Meter, P., & Schuder, T. (1996). A quasi-experimental validation of transactional strategy instruction with low-achieving second-grade readers. *Journal of Educational Psychology, 88*(1), 18–37. doi:10.1037/0022-0663.88.1.18

Caldwell, J., & Leslie, L. (2003). Does proficiency in middle school reading assure proficiency in high school reading? The possible role of think-alouds. *Journal of Adolescent and Adult Literacy, 47*(4), 324–335.

Cohen, J.W. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Ass.

Deshler, D., Hock, M. F., Ihle, F. M., & Mark C. A. (2011). Designing and conducting literacy intervention research. In M. L. Kamil, P. D. Pearson, E. B. Moje, & P. Afflerbach (Eds.), *Handbook of reading research, Vol. IV* (pp. 66–83). New York, NY: Routledge.

Dole, J. A., Nokes, J. D., & Drits, D. (2009). Cognitive strategy instruction. In S. E. Israel & G. G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 347–372). New York, NY: Taylor & Francis.

Donndelinger, S. J. (2005). Integrating comprehension and metacognitive reading strategies. In S. E. Israel, C. C. Block, K. L. Bauserman, & K. Kinnucan-Welsch (Eds.), *Metacognition in literacy learning: Theory, assessment, instruction and professional development* (pp. 241–260). Mahwah, NJ: Lawrence Erlbaum Ass.

Duke, N. K., & Pearson, P. D. (2002). Effective practices for developing reading comprehension. In A. E. Farstrup & S. J. Samuels (Eds.), *What research has to say about reading instruction* (3rd ed.) (pp. 205–242). Newark, DE: International Reading Association, Inc.

Edmonds, M. S., Vaughn, S., Wexler, J., Reutebuch, C., Cable, A., Klingler Tackett, K., & Wick, J. (2009). A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. *Review of Educational Research, 79*(1), 262–300. doi:10.3102/0034654308325998

Elbro, C., & Buch-Iversen, I. (2013). Activation of background knowledge for inference making: Effects on reading comprehension. *Scientific Studies of Reading, 17*(6), 435–452. doi:10.1080/10888438.2013.774005

Fielding, L. G., Anderson, R. C., & Pearson, P. D. (1990). *How discussion questions influence children's story understanding* (Tech. Rep. No. 490). Urbana, IL: University of Illinois, Center for the Study of Reading.

Frederking, V., Henschel, S., Meier, C., Roick, T., Stanat, P., & Dickhäuser, O. (2012). Beyond functional aspects of reading literacy: Theoretical structure and empirical validity of literary literacy. *L1-Educational Studies in Language and Literature, 12*, 1–24.

Fredriksson, U. (2012). Hur 15-åringars läsning förändrats mellan 2000 och 2009. Resultat från PISA-undersökningarna. [Changes in 15-year-olds' reading 2000 through 2009. Results from the PISA test.] In SOU 2012:10 *Läsarnas marknad, marknadens läsare – en forskningsantologi* [SOU 2012:10 *The readers' market, The market's readers: A Research anthology*] (pp. 95–110). Stockholm, SW: Fritzes.

Genereux, R., & McKeough, A. (2007). Developing narrative interpretation: Structural and content analyses. *British Journal of Educational Psychology, 77*(4), 849–872. doi:10.1348/000709907X179272

Gersten, R., Fuchs, L.S., Williams, J. P., & Baker, S. (2001). Teaching reading comprehension strategies to students with learning disabilities. A review of research. *Review of Educational Research*, *71*(2), 279–320. doi:10.3102/00346543071002279

Graesser, A. C. (2007). An introduction to strategic reading comprehension. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 3–26). Mahwah, NJ: Lawrence Erlbaum Ass.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101*(3), 371–395. doi:10.1037/0033-295X.101.3.371

Graham, S., & Hebert, M. (2010). *Writing to read: Evidence for how writing can improve reading*. Washington, DC: Alliance for Excellent Education.

Griffith, P. L., & Ruan, J. (2005). What is metacognition and what should be its role in literacy instruction? In S. E. Israel, K. L. Bauserman, C. C. Block, & K. Kinnucan-Welsch (Eds.), *Metacognition in literacy learning: Theory, assessment, instruction and professional development* (pp. 3–18). Mahwah, NJ: Lawrence Erlbaum Ass.

Hosseini, K. (2003). *The kite runner*. New York, NY: Riverhead Books.

Iser, W. (1978). *The act of reading. A theory of aesthetic response*. Baltimore, MD: Johns Hopkins Univ. Press.

Israel, S. E., & Massey, D. (2005). Metacognitive think-alouds: Using a gradual release model with middle school students. In S. E. Israel, K.L. Bauserman, C. C. Block, & K. Kinnucan-Welsch (Eds.), *Metacognition in literacy learning: Theory, assessment, instruction and professional development* (pp. 183–198) Mahwah, NJ: Lawrence Erlbaum Ass.

Janssen, T., Braaksma,, M., & Couzijn, M. (2009). Self-questioning in the literature classroom. effects on students' interpretation and appreciation of short stories. *L1 – Educational Studies in Language and Literature, 9*(1), 91–116.

Janssen, T., Braaksma, M., & Rijlaarsdam, G. (2006). Literary reading activities of good and weak students: A think aloud study. *European Journal of Psychology of Education, 11*(1), 35–52. doi:10.1007/BF03173568

Kamil, M. L. (2004). Vocabulary and comprehension instruction. Summary and implications of the National Reading Panel findings. In P. McCardle & V. Chhabra (Eds.), *The voice of evidence in reading research* (pp. 213–234). Baltimore, MD: Paul H. Brookes.

Machor, J. L., & Goldstein, P. (Eds.). (2001). *Reception study: From literary theory to cultural studies*. New York, NY: Routledge.

Mandler J. M., & Johnson N.S. (1977). Remembrance of things parsed. Story structure and recall. *Cognitive Psychology, 9*(1), 111–151. doi:10.1016/0010-0285(77)90006-8

Marshall, J. D. (1987). The effects of writing on students' understanding of literary texts. *Research in the Teaching of English, 21*(1), 30–63.

Marshall, J. D., Smagorinsky, P., & Smith, M. W. (1995). *The language of interpretation: patterns of discourse in discussions of literature*. Urbana, IL: NCTE.

Mercer, N. (1996). The quality of talk in children's collaborative activity in the classroom. *Learning and Instructio*n, *6*(4), 359–377. doi:10.1016/S0959-4752(96)00021-7

Murphy, P. K., Wilkinson, I. A. G., Soter, A. O., Hennessey, M. N., & Alexander, J. F. (2009). Examining the effects of classroom discussion on students' comprehension of text: A meta-analysis. *Journal of Educational Psychology, 101*(3), 740–764. doi:10.1037/a0015576

National Reading Panel (2000). *Teaching children to read. An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* [online]. Available at: http://www.nichd.nih.gov/publications/pubs/nrp/documents/report.pdf

Newell, George E. (1994). The effects of written between-draft responses on students' writing and reasoning about literature. *Written Communication, 11*(3), 311–347. doi:10.1177/0741088394011003002

Nystrand, M. (2006). Research on the role of classroom discourse as it affects reading comprehension. *Research in the Teaching of English, 40*(4), 392–412.

Nystrand, M., & Gamoran, A. (1991). Instructional discourse, student engagement, and literature achievement. *Research in the Teaching of English, 25*(3), 261–290.

Nystrand, M., Gamoran, A., Kachur, R., & Prendergast, C. (1997). *Opening dialogue: Understanding the dynamics of language and learning in the English classroom*. New York: Teachers College Press.

Oatley, K. (1999). Why fiction may be twice as true as facts. Fiction as cognitive and emotional simulation. *Review of General Psychology, 3*(2), 101–117. doi:10.1037/1089-2680.3.2.101

Palinscar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction, 1*(2), 117–175. doi:10.1207/s1532690xci0102_1

Pearson, P. D. (2009). The roots of reading comprehension instruction. In S. E. Israel & G. G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 3–31). New York, NY: Taylor & Francis.

Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The acquisition of reading comprehension skill. In M. J. Snowling, & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 227–247). Malden, MA: Blackwell. doi:10.1002/9780470757642.ch13

Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Pressley, M., Beard El-Dinary, P., Gaskins, I., Schuder, T., Bergman, J. L., Almasi, J., & Brown, R. (1992). Beyond direct explanation: Transactional instruction of reading comprehension strategies. *The Elementary School Journal, 92*(5), 513–555. doi:10.1086/461705

Reninger, K. B. (2007). *Intermediate-level, lower-achieving readers' participation in and high-level thinking during group dicussions about literary texts*. The Ohio State University. https://etd.ohiolink.edu/ap/10?0::NO:10:P10_ACCESSION_NUM:osu1180009715

Rosenshine, B., & Meister, C. (1994). Reciprocal teaching: A review of the research. *Review of Educational Research, 64*(4), 479–530. doi:10.3102/00346543064004479

Saunders, W., & Goldenberg, C. (1999). Effects of instructional conversations and literature logs on limited- and fluent-English-proficient students' story comprehension and thematic understanding. *The Elementary School Journal, 99*(4), 279–301. doi:10.1086/461927

Scammacca, N. K., Roberts, G., Vaughn, S., & Stuebing, K. K. (in press). A meta-analysis of interventions for struggling readers in grades 4–12: 1980–2011. *Journal of Learning Disabilities*. doi:10.1177/0022219413504995

Skaftun, A. (2011). Understanding reading development: A phenomenological perspective. *L1-Educational Studies in Language and Literature, 11*, 127–148.

Snow, C. E. (Ed.) (2002). *Reading for understanding: Toward a R&D program in reading comprehension*. Washington, DC: RAND Corporation.

Soter A. O., Wilkinson, I. A. G., Murphy, P. K., Rudge, L., Reninger, K. B., & Edwards, M. (2008). What the Discourse Tells Us: Talk and Indicators of High-Level Comprehension. *International Journal of Educational Research, 47*(6), 372–391. doi:10.1016/j.ijer.2009.01.001

Souvignier, E., & Mokhlesgerami, J. (2006). Using self-regulation as a framework for implementing strategy instruction to foster reading comprehension. *Learning and Instruction, 16*(1), 57–71. doi:10.1016/j.learninstruc.2005.12.006

Spörer, N., Brunstein, J. C., & Kieschke, U. (2009). Improving students' reading comprehension skills: Effects of strategy instruction and reciprocal teaching. *Learning and Instruction 19*(3), 272–286. doi:10.1016/j.learninstruc.2008.05.003

Swedish National Agency for Education (2006). *National test in Swedish/Swedish as a Second language 2011. Part A. To read and to comprehend*. Stockholm, SW: Swedish National Agency for Education.

Swedish National Agency for Education (2007). *National test in Swedish/Swedish as a Second language 2011. Part A. To read and to comprehend.* Stockholm, SW: Swedish National Agency for Education.

Swedish National Agency for Education (2009). *National test in Swedish/Swedish as a Second language 2011. Part A. To read and to comprehend.* Stockholm, SW: Swedish National Agency for Education.

Swedish National Agency for Education (2011). *National test in Swedish/Swedish as a Second language 2011. Part A. To read and to comprehend.* Stockholm, SW: Swedish National Agency for Education.

Taylor, B. M., Pearson, P. D., Clark, K. F., & Walpole, S. (1999). Effective schools/accomplished teachers. *The Reading Teacher, 53*(2), 156–159.

Tengberg, M. (2011). *Samtalets möjligheter. Om litteratursamtal och litteraturreception i skolan*. Stockholm/Stehag, SW: Symposion.

Tengberg, M. (2014). Att pröva elevers läsförmåga i årskurs nio. En analys av uppgiftkonstruktioner i det nationella provet i svenska, delprov A "Att läsa och förstå". *Utbildning & Demokrati, 23*(2), 109–132.

Tengberg, M. & Olin-Scheller, C. (2013). Improving reading and interpretation in seventh grade. A comparative study of the effects of two different models for reading instruction, *Education Inquiry, 4*(4), 689–713. doi:10.3402/edui.v4i4.23221

Thurlow, R., & van den Broek, P. (1997). Automaticity and inference generation. *Reading and Writing Quarterly, 13*(2), 165–184. doi:10.1080/1057356970130205

Vygotsky, L. S. (1978). *Mind in Society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Vygotsky, L. S. (1986). *Thought and language.* Cambridge, MA: The MIT Press.

Wilkinson, I. A. G. & Son, E. H. (2011). A dialogic turn in research on learning and teaching to comprehend. In M. L. Kamil, P. D. Pearson, E. B. Moje, & P. Afflerbach (Eds.), *Handbook of reading research, Vol. IV* (pp. 359–387). New York, NY: Routledge.

Wong, B. Y. L., Kuperis, S., Jamieson, D., Keller, L. & Cull-Hewitt, R. (2002). Effects of guided journal writing on students' story understanding. *The Journal of Educational Research, 95*(3), 179–191. doi:10.1080/00220670209596588

Wood, D. J., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychiatry and Psychology, 17*(2), 89–100. doi:10.1111/j.1469-7610.1976.tb00381.x

SHORT FICTION

Ajvide Lindqvist, J. (2006). Pappersväggar [Paper walls]. In *Pappersväggar: tio berättelser*. Stockholm, SW: Ordfront.

Andersson, G. (1998). Vågen [The wave]. In *Torrsimmaren: berättelser från ett ungt 50-tal.* Stockholm, SW: En bok för alla.

Axelsson, M. (2002) Över spåret [Across the track]. In *Färdlektyr 2002*. Stockholm, SW: Svenska lokaltrafikföreningen.

Frame, J. (1951/2002). Dossy [Dossy]. In *Lagunen och andra berättelser*. Stockholm, SW: Bonnier.

Gardell, J. (1992). section (pp. 84–91) from *En komikers uppväxt* [*A Comedian Growing Up*]. Stockholm, SW: Norstedts.

Lagerkvist, P. (1924/1997). Far och jag [Father and I]. In *Onda sagor*. Stockholm, SW: Aldus/Bonniers.

Nauman, C. (2009). Mörkret [The darkness]. In *Kulor i hjärtat*. Stockholm, SW: Alfabeta.

Nesser, H. (2000). Om någonting händer [In case something should happen]. In *Färdlektyr 2000.* Stockholm, SW: Svenska lokaltrafikföreningen.

Nesser, H. (2002). section from *Kära Agnes!* [*Dear Agnes!*]. Stockholm, SW: Bonnier.

APPENDIX

Table 3. Matching of strategies and texts

Strategies	*Texts*
1. Summarizing	"The wave", Andersson (1998) "Dossy", Frame (1951/2002)
2. Making predictions	"The darkness", Nauman (2009)
3. Evaluating	"Across the track" Axelsson (2002)
4. Visualizing	*A Comedian Growing Up*, Gardell (1992)
	"The wave", Andersson (1998)
5. Questioning	"Paper walls" Ajvide Lindqvist (2006)
6. Finding gaps and making infer-ences	*Dear Agnes!* Nesser (2002) "In case something should happen", Nesser (2000)
7. Comparing with other texts	"Father and I", Lagerkvist (1924/1997)