This is the published version of a paper published in *Education Inquiry*.

Citation for the original published paper (version of record):

Improving Reading and Interpretation in Seventh Grade: A Comparative Study of the Effects of Two Different Models for Reading Instruction
*Education Inquiry, 4*(4)
https://doi.org/10.3402/edui.v4i4.23221

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

© 2013 Michael Tengberg & Christina Olin-Scheller This is an Open Access article distributed under the terms of the Creative Commons Attribution-Noncommercial 3.0 Unported (CC BY 3.0) License (http://creativecommons.org/licenses/by-nc/3.0/), permitting all non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:kau:diva-26813
Improving Reading and Interpretation in Seventh Grade: A Comparative Study of the Effects of Two Different Models for Reading Instruction

Michael Tengberg & Christina Olin-Scheller

To cite this article: Michael Tengberg & Christina Olin-Scheller (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, 23221, DOI: 10.3402/edui.v4i4.23221

To link to this article: https://doi.org/10.3402/edui.v4i4.23221

© 2013 Michael Tengberg & Christina Olin-Scheller

Published online: 06 Dec 2013.

Article views: 2097

View related articles

Citing articles: 1 View citing articles
Improving Reading and Interpretation in Seventh Grade: A Comparative Study of the Effects of Two Different Models for Reading Instruction

Michael Tengberg* & Christina Olin-Scheller*

Abstract
This article investigates the potential for developing advanced reading and interpretative skills in Swedish secondary school. Over six weeks, four separate study groups in seventh grade participated in an intervention study. The purpose was to gather more knowledge about how different teaching strategies affect students’ development of advanced reading skills. Data were collected from written assignments and an experiential questionnaire. The study attempts to provide empirical corroboration of previously theoretically founded propositions. The findings indicate that the choice of teaching strategy plays a significant role in students’ learning. However, in order to make normative judgments on the advantages of one teaching model over another, further evidence is necessary. Moreover, studying the implementation and effects of instructional change is a challenging field of research which requires an array of sophisticated methods that combine qualitative and quantitative analyses.

Keywords: reading instruction, intervention study, interpretation, comprehension strategies, dialogue

1. Reading instruction at risk?
During the last few years several reports have verified that the reading achievements of Swedish students are declining, in both an international comparison and in comparison with earlier national surveys (Gustafsson and Rosén 2005; Skolverket 2007a; Skolverket 2010a). Tasks related to the reading of fiction and of continuous expository texts are pointed out as an area of special concern. After examining reading achievements in third grade, Gustafsson and Rosén found in particular that the students’ ability to read extended texts had decreased over a 10-year period. In the PISA study of 2009, Sweden was one of four countries whose overall performance in the reading literacy section had decreased significantly. Whether or not these tests can offer a satisfactory account of the complexity of students’ reading literacy is of course an interesting issue in itself, although it will not be addressed in
this article and neither will the question of the accurateness of the reported decline in Swedish performances. Instead, we are interested in what schools can do to foster more advanced forms of reading and interpreting literature. More precisely, what difference does the teacher’s choice of instructional model make for the quality of students’ learning?

Several reports from the Swedish National Agency for Education and from the Swedish Schools Inspectorate have indicated that explicit reading instruction at a higher level is rare in Swedish schools and that there is a pressing need to develop creative and challenging reading environments (Skolverket 2004, 2007b; Skolinspektionen 2010, 2012). These tendencies are supported by the prevalent in-depth research on literature instruction. Yet there is little agreement about the way in which the complex art of strategic and critical reading can be more strongly cultivated (Arfwedson 2006; Schmidl 2008; Tengberg 2011). The scarcity of explicit reading instruction in schools is internationally recognised (Roe 2011). Pressley and Wharton McDonald (2006) underscore that, even though two decades of research has demonstrated that the teaching of comprehension strategies has a crucial role to play as a continuation of elementary literacy instruction, explicit comprehension instruction is still rare in classrooms of grades 4, 5 and 6.

The question thus arises: what is it about reading that students should be taught? Our particular research interest is devoted to the relationship between teaching practices and students’ development of advanced skills for reading and interpretation. The term advanced skills is still to be regarded as a provisional conceptualisation, borrowed from the PIRLS study (c.f. Skolverket 2007b), one which we believe requires further elaboration. Here, it designates aspects of reading such as making inferences, interpreting, reflecting and evaluating. The focus of our investigation concerns reading instruction at secondary level. In this case, we focus on the reading of short stories, which will be partially determinative for the instructional design. However, this limitation should be regarded as a way of exemplifying the study of reading instruction rather than excluding reading instruction concerned with expository texts.

In Swedish research on literature instruction, literary reading is ultimately seen as a practice deeply embedded within the larger domain of socialisation, rather than as a skill to be developed (Malmgren 1997; Molloy 2002; Mehrstam 2009; Lundström et al. 2011). Terms like reader identities (following Holland 1975), interpretive communities (Fish 1980) and repertoires (McCormick 1994) have shaped the understanding of literature and schooling, and in a way discarded the probability that readers’ reception is also a product of conscious and strategic choices. In this article, we adopt a different view which emphasises reading and interpretation not only as contingent but also as particular practices that can be developed and cultivated by way of careful teaching. Without doubt, this is also the consensus idea upon which all students are consecutively assessed at national
assessments. We believe that reflective reading instruction can provide students with better opportunities to meet those standards.

2. Purpose of the study

The purpose of the study is to find out whether a combination of carefully chosen strategies for classroom instruction can make a significant difference for students’ development of advanced reading and interpretive skills. The study is the first in a series of related studies, which means that it also serves to put our theoretical and methodological framework to the test.

Previous research in Sweden is partly devoted to the relationship between methods for instruction and students’ learning. Yet the most frequently cited studies in the field are predominantly ethnographic, providing ‘thick descriptions’ of a limited number of teachers’ teaching (Molloy 2002; Olin-Scheller 2006; Ewald 2007). These studies are excellent when it comes to generating critical analyses and theoretically solid perspectives on school culture, teachers’ choices of texts, particular teaching strategies etc. Among other things, they stress the scarcity of collective teaching processes and disclose the shortcomings of the widespread use of restricted writing tasks. However, the pedagogical alternatives to practical constraints of this kind are rarely empirically tested. Therefore, prevailing research tends to provide limited support for the development of new teaching strategies. Envisioning an alternative pedagogy is perhaps a necessary start, but researchers must also formulate strong hypotheses about teaching and learning, and subject them to empirical testing, in order to provide well founded answers to serious and practically relevant questions. Theories are necessary; it should be obvious that empirical testing without theoretical motivation is superficial but, on the other hand, theoretical propositions without empirical corroboration are equally inadequate.

In this article, by empirical means we investigate some theoretically-based propositions about the teaching of literature and its potential for empowering students’ reading. Our research questions are:

1. Will a carefully designed instructional frame — based on dialogical principles (Nystrand and Gamoran 1991; Marshall et al. 1995) and the introduction of reading comprehension strategies (Duke and Pearson 2002) — have measurable, positive effects on students’ reading comprehension, compared with students who follow an alternative instructional frame that is not in line with those principles?

2. What are the qualitative differences between the students’ reading responses after having followed either of the two instructional frames?

3. In what way do the students themselves describe their learning after having followed either of the two instructional frames?
We also intend to use these analyses to reflect on some more fundamental scientific issues, such as under what conditions may the effects of instructional interventions really be discerned by way of empirical measurements?

3. Dialogue, comprehension strategies, and extended writing assignments

In order to provide a functional foundation for the improvement of literature education, researchers have tried to locate activities or teaching perspectives that could serve as indicators of a more productive learning environment. Three areas are continuously pointed out as significant:

- **authentic and structured dialogue on mutually read texts** (Nystrand and Gamoran 1991; Applebee et al. 2003; Murphy et al. 2009);
- **explicit instruction in comprehension strategies** (Duke and Pearson 2002; Pressley 2006; Duke and Carlisle 2011); and
- **extended, analytic writing assignments** (Marshall 1987; Newell 1994; Wong et al. 2002; Graham and Hebert 2010).

As a complement to more specialised studies in these three areas, we will examine in what way a balanced combination of them affects students’ learning and their reading of literature.

Interest in dialogue-based approaches to the reading of literature has become a significant field of research during the last decade, and has revealed a great deal of potential for providing professionals with both incentives and recommendations for improving their practices. One of the baseline arguments concerning literature discussions is the capability of making the conversation facilitate students’ analytical and reflexive thinking about their reading experiences. Discussion is seen as a tool for thinking together and elaborating on each other’s understandings (Barnes 1976; Mercer 1996). One empirical example is Roberts’ and Langer’s (1991) in-depth analysis of an entire discussion in which they demonstrate, step by step, how the teacher and the students support and enrich each other’s interpretations of the read text. In the Northern countries, the study by Nystrand and Gamoran (1991) had a crucial impact on the development of a dialogic turn in the research on both reading and writing. The approach of Nystrand and Gamoran is based on a pragmatic and professionally oriented understanding of Bakhtin’s dialogism in which the multiplicity of voices (i.e. perspectives) in the classroom produce tension, or even conflict, “as one voice ‘refracts’ another” (Nystrand 1997, 8). This tension lies at the heart of the interactional dynamic through which students’ understandings are reconfigured and structured and by which a transformation of shared knowledge may occur. On a similar basis, a vast body of research has emerged that consists of content-oriented approaches to, for instance, gender practices (Arizpe 2001; Eriksson Barajas 2008;
Tengberg 2009), combinations of macroanalytic and microanalytic analyses in order to measure the effects of different interactional patterns (Almasi et al. 2001; Chinn et al. 2001), and meta-analyses in order to expose discourse features that may serve as indices of high-level learning and high-level comprehension of text (Soter et al. 2008; Murphy et al. 2009). In order to provide a plausible account of how literature discussions help to enrich and encourage students’ comprehension of texts, we first need theoretical models to make strong connections between discourse features and students’ changes of understanding and, second, empirical studies to support the relationship between enduring interventions in the classroom and improvement of students’ reading skills.

Research on the use of comprehension strategies has an equally long tradition, especially in American educational research. One strand of this research is devoted to what skilled readers do when they read. The idea is to target skilled readers’ comprehension processes in order to formulate goals for future reading instruction. Wyatt et al. (1993), for example, used verbal protocol methods to investigate the reading processes of 15 professors from different areas. Among other things, they identified three key patterns that could account for skilled reading. First of all, the readers continuously summarised what they read, making anticipations and selections between more and less important passages. Second, they monitored their own reading, paying attention to the text’s level of difficulty to them, and to whether the ideas presented in the text were new to them or not. Third, the readers repeatedly evaluated what they were reading, making strong judgments as to whether the author was presenting ideas consistent with the readers’ own thinking on the subject. In line with the last point, Janssen et al. (2006) found in a think-aloud study of tenth-grade students’ reading of short stories that stronger readers tended to be more evaluative and emotional in their responses. In an overview of more than 60 think-aloud studies, Pressley and Afflerbach (1995) summarised the findings on skilled reading. An essential component of their conclusion was that the conscious process of skilled reading is not just something that goes on while the reader is physically leaning over the pages of the text. Instead, it begins before reading by overviewing the text, observing its structure and connecting to previous knowledge and reading experiences. This activity goes on during reading and continues afterwards, by rereading or reskimming sections of particular importance, by restating or summarising essential ideas, and by continuing to reflect on the text and its meaning. According to Pressley and Afflerbach, good readers make use of a number of comprehension strategies during their interaction with the text: they make inferences and tentative conclusions, they relate setting and characters, and they make interpretations as they go along that relate the text to previous reading experiences and to world knowledge (Pressley and Afflerbach 1995). Central to these studies on comprehension strategies is that good reading is active reading. It relies on a combination of meta-cognitive strategies used to interrogate, identify
and evaluate the text (Israel et al. 2005). Recent research points out, however, that these strategies cannot be taught as a series of isolated skills but need to be brought together and continuously utilised as a function of previous knowledge (Duke and Pearson 2002; Donndelinger 2005).

The relationship between reading and writing, and the benefit of supporting reading comprehension with personal and analytic writing, is strongly debated among both professionals and researchers. Keeping response journals while reading novels or writing various assignments in response to short stories are relatively common activities in literature classes in secondary school. Yet, according to Wong et al. (2002), only a limited number of studies demonstrate in what way response writing contributes to students’ literary understanding. In an analysis of students’ understanding of short stories as an effect of different types of writing assignments, Marshall (1987) found that extended writing assignments (both personal analytic and formal analytic) significantly enhanced the quality of responses in a post-test compared to when students were given restricted writing assignments (a selection of short answer questions). Following Marshall, Newell (1994) investigated how students reformulated their initial written analyses of short stories as a response to two different types of teacher comments. Investigators, acting as teachers, would either respond to the students’ draft in a dialogic style, which meant they would “pose questions and offer strategies to help the students use what they already had drafted” (Newell 1994, 320), or in a directive style, which meant they evaluated the students’ responses and tried to make them conform to the ‘teacher’s’ interpretation. Responses to post-test questions of story understanding showed that students who were given the dialogue style comments produced “qualitatively more sophisticated answers” across three types of questions (descriptive, interpretive, generalisation). Contributing to these results, Wong et al. (2002) showed that guided-response journals increased students’ understanding of themes and main characters when reading The Great Gatsby, and that the students themselves perceived that the writing activities contributed to their understanding of the novel. An idea emerges from these studies about the type of writing assignments that would produce more relevant effects on students’ reading. In a meta-analysis of 93 experiments on the impact of writing on reading, Graham and Hebert (2010) found that the task which produced the strongest average weighted effect size (.77) on reading comprehension was having students respond to a text by writing personal reactions, analysing and interpreting the text.

These studies serve as a valuable guide for developing classroom practices on reading instruction. However, more than a few of them draw their conclusions from results on a single intervention effect at a particular moment in time. Moreover, changes in students’ reading comprehension are often measured in relation to the very texts on which the intervention was executed (a dependent measure). This means that the experiments demonstrate a capability of providing students with
either more or less support for a deepened understanding of a specific text, but not necessarily that the interventions enhance the students’ general ability to comprehend texts. Therefore, the ambition in our study is to provide a more comprehensive display of the potential for improving reading instruction at secondary level by combining structured dialogues on mutual reading experiences, explicit instruction on a selection of evidently powerful comprehension strategies, as well as extended and challenging writing assignments. Our research focus is to find out whether this combined approach can promote a durable effect on students’ capacity for reading and interpreting literary texts.

4. Methods and material

Participants and site

The instructional intervention was planned and conducted in cooperation with three teachers of Swedish in three seventh-grade classrooms in a small-sized city secondary school (N students = 56). The teachers were selected by way of previously expressed interest in working with systematic reading instruction. They were all educated teachers of Swedish, with long experience in secondary school teaching. It was decided to carry out the study in grade 7 for two reasons. First of all, these students had only recently been grouped together and would therefore have only had very little time to develop specific classroom cultures or specific expectations regarding their teacher’s teaching. Second, according to the teachers these students had limited previous experience of reading short stories and of working explicitly with comprehension strategies.1 This made them suitable as subjects for the intended intervention. However, due to social issues beyond our control, which came to impede the planned teaching, one of the classes had to be left out of the analysis. The lessons in this class were often obstructed, and several of the planned activities could not be carried out. During the end of the school term, this class was even split up as a way of stemming the escalation of the group’s troublesome development. Thus, the data used in the following analysis concern the two remaining classes (N students = 37).

Material

The materials used in this study included three short stories by Swedish authors to be worked on during the intervention.2 Before the intervention, a pre-test was given consisting of a short story (“A Day’s Wait” by Ernest Hemingway) and three concomitant writing tasks. After the intervention a similar test was given as a post-test, consisting once again of a short story (“Far och jag” by Pär Lagerkvist) and three writing tasks (see Appendix I). The purpose of the tests was to assess the students’ reading performances (qualitatively and quantitatively) before and after the intervention. When using short stories, there is obviously no way of ensuring that
two different texts will provide the students with a similar level of challenge in both the pre-test and post-test. The writing tasks in both tests were designed to assess similar types of reading performances (two tasks on a more formal integration and interpretive level, and one task on the level of personal interpretive response). Yet, since one test might still prove to be a little more difficult than the other, a statistical within-group comparison of the results for the pre-test and post-test will be less informative. A more important measure is thus a between-group comparison since students in both conditions (see below) take the same two tests and are thereby provided with the same level of difficulty on both occasions, even though that level may not be entirely invariable going from the pre-test to the post-test.

Further, a questionnaire consisting of 12 rating items and one open-ended question was given about a week after the teaching period. Topic areas were students’ estimations of: 1) previous reading instruction; 2) instructional factors; 3) reading abilities; and 4) learning progress during the intervention (see Appendix II). We also conducted interviews with teachers and groups of students from all four study groups, and performed continuous observations of classroom work. All interviews were conducted after the intervention period and focussed on the participants’ experiences of the classroom work. Observations in all study groups, using a simple observation scheme, provided complementary information about the procedures that were carried out during the lessons.

**Procedure**

The study employs a comparative, quasi-experimental method in which the two classes were divided into four study groups. Two of the groups were subjected to a six-week-long segment of specially designed reading instruction, the *dialogue-based condition* (DBC) (N = 19). The other two groups were subjected to an equally long segment of an alternatively designed teaching, the *monologue-based condition* (MBC) (N = 18). Both of these teaching segments were designed in cooperation between researchers and teachers, and both of the teachers taught one study group in each condition. In order not to cause more practical discomfort than necessary, the students were not randomly assigned to the study groups. Instead, the teachers were instructed to split each of the two classes into two halves, and to ensure that the groups would contain a fairly similar proportion of what they deemed as poor, average and good readers. When comparing the post-tests this procedure alone will obviously not control for pre-existing differences between groups and the pre-test was therefore a necessary component.

The stories were read and worked on at the same pace in all four groups. However, in the DBC group the classroom work pivoted around structured, teacher-led group discussions and extended writing tasks. Further, a selection of comprehension strategies was introduced to each of the texts and, after the third text had been used, the teacher would choose one of the three texts to read again, this time
using a new set of comprehension strategies. Altogether, seven different strategies were taught and exercised during the six weeks: making predictions; asking questions; summarising; evaluating; visualising the text world; comparing with other texts; and finding gaps and making inferences. These strategies were chosen and designed partly in cooperation with the teachers and partly on the basis of previous design-based research. The idea of combining them in a single intervention – which is uncommon in previous studies – was to equip the teachers with a more complete and multidimensional teaching model (necessary for authentic reading instruction).

The classroom work in the MBC group focused primarily on what Marshall (1987) defined as “restricted writing assignments”, i.e. short answer questions on a variety of levels. By continuously prompting students to shift focus, according to Marshall the short answer questions discourage “students’ construction of a coherent and elaborated understanding of the story” (Marshall 1987, 43). The students were assigned to work with a set of short answer questions regarding the stories they had read. Unlike the intention of focusing on comprehension strategies and extended writing tasks, which is more to support a more holistic understanding by encouraging students to make connections and find integral patterns in the text, short answer questions are expected to support an atomistic orientation to a text. Different areas or aspects of text are treated as separate, and students are not prompted to make syntheses of their observations or to look for structural configurations in the text. In theory, this would encourage a more erratic perception of story meaning.

The MBC also contained the writing of book reports, a task that much in the same way is thought to induce an atomistic way of thinking about literature.

**Analyses of the data sets**

In order to examine how the two instructional interventions influenced the students’ reading strategies and their self-estimates of having developed their reading skills, data from the tests and questionnaires were analysed for qualitative and quantitative differences. During the tests, the stories were read aloud by the teachers while the students had copies of the texts in front of them, but no instructional scaffolding was given. The measure therefore focuses on transferred comprehension effects rather than on lesson-text comprehension effects, which is common in intervention studies on reading comprehension instruction (c.f. McKeown et al. 2009). As noted above, the test questions were designed to elicit formal and personal inferences on the story level. Formal story-level inferences mean that the students were to make inferences in order to explain aspects of the story development or of the characterisation that were not explicit in the texts. Personal story-level inferences mean that the students were asked to provide inferences in a more independent interpretive way.
These responses were meant to indicate development of the students’ reading and interpretation. Four categories of results will be presented here. They derive partly from the students’ responses to the pre-tests and post-tests, and partly from the questionnaire data. The test responses were subjected both to qualitative text analysis and quantitative ratings of reading skills. The qualitative text analysis meant that the responses were examined for traces of the students’ use of various reading strategies. Since the focus here is exploratory we did not make use of a predesigned coding scheme; instead, the categories of interest were inductively assembled through a combination of synchronic comparisons between the DBC and MBC group responses and diachronic comparisons between the pre- and post-test-level responses. The categories we collected included references to text, distinction of characterisation, and visual and symbolic inferences.

The statistical approach consisted of ratings of the student performances in the pre-test and post-test using a three-grade scale defined in the following way: 1 = The student demonstrates no or little ability to respond to the three tasks. 2 = The student demonstrates a good ability to respond to the three tasks. 3 = The student demonstrates an excellent ability to respond to the three tasks. Before the ratings, detailed criteria for each story, task and level of rating were discussed and agreed upon. Two raters then independently rated the entire dataset and attained a strong inter-rater agreement (Cohen’s Kappa .936; p < .001).

If the DBC provides students with a substantially better opportunity for developing advanced reading and interpretive skills, one would expect the DBC students to score significantly better in the post-test than the MBC students. To explore the differences between the two groups’ test results an analysis of covariance (ANCOVA) was used. Since ANCOVA controls for any pre-test differences, it is particularly useful when you have small sample sizes – which is the case here – and it also increases the sensitivity of the test for main effects. However, since the sample is small, the statistical analysis should still be viewed as complementary to the more elaborate descriptions of the content of the students’ progress provided by the qualitative analysis.

The third result category deals with the students’ own experiences of the instructional frame and to what extent they estimated that the six weeks of teaching had contributed to their reading skills. We explore three sets of questions in particular. First of all, we were interested in whether the students in the DBC would experience a stronger sensation of having developed their reading and interpretation skills than the students in the MBC. Second, it was expected that the students in the DBC would experience a more distinct contrast between the designed teaching and the teaching they were used to from earlier schooling. In the interviews, both students and teachers had reported that the reading instruction in grades 4, 5 and 6 had not accustomed the students to comprehension strategies, structured
dialogues about texts or to extended response assignments. On the other hand, they were accustomed to writing book reports and responding to short answer questions. Therefore, we also measured the students’ sense of novelty or rupture when subjected to the teaching we had designed.

The fourth result category is concerned with the students’ own accounts of their learning about reading and interpretation. Based on an open-ended item in the questionnaire we compared the particular content of what the students from the DBC and the MBC groups express that they have learned. Taking as a point of departure that being able to express one’s understanding in itself suggests a more complex quality of understanding, we set out to examine whether the students in the DBC display a more developed capacity for articulating what they have learned about reading than the students in the MBC.

**Ethics**

All participants were informed of the study’s purpose. Verbal consent was gathered from both the teachers and students. Since the students were under 15 years of age, written consent was also collected from all of the students’ parents. The 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 purposes only.

5. **Results**

**Qualitative analysis of the test responses**

The response data were divided into four groups: DBC pre-test, DBC post-test, MBC pre-test and MBC post-test. When comparing the DBC pre-tests with the MBC pre-tests, the homogeneity of the reading strategies displayed by the students in the two groups was found to be significant. Generally, all students were *plot-oriented* (Tengberg 2011), meaning that their responses mainly focused on the continuation of the plot, which is natural when considering that the questions had been partly designed to map the orientations of the students’ active participatory responses. For instance, in the last question most students focused on whether the boy in the story (“A Day’s Wait”) would either live or die. In a few of the DBC responses, however, there was a slight tendency towards a perception of anxiety in the boy’s disposition. Their suggestions could be interpreted as stating that not only does the boy need to recover from a medical illness, but he also has to recuperate mentally from the prospect of dying from his fever. These responses may thus be seen as a display of an extended use of reading strategies (attending to both plot *and* to characterisation and the character’s feelings). Bearing in mind that these examples are in a clear minority, they still indicate a possible, minor, initial difference between the readers who were to follow either the DBC or the MBC.
However, the post-test responses proved to contain more evident differences. Although the combination of a new story ("Far och jag") and a new set of questions may have spurred a conversion of the reading strategies utilised, this conversion was not identical in the two groups. Our analysis showed that three particular features distinguished the DBC responses from the MBC responses. Although these features cannot be generalised at group level, they appear repeatedly in the DBC post-test responses and not at all in the pre-test responses or the MBC post-test responses. These three features are:

1. **Text references.** Students used examples from, or direct references to, text passages in order to motivate their responses.
2. **Distinction of characterisation.** Students displayed more distinction, or finer nuances, in their description of story characters.
3. **Inner visualisations and symbolic inferences.** Students used visualisations (one of the comprehension strategies) and symbolic inferences to describe their meaning-making processes in relation to specific text passages.

From diachronic comparisons of the pre-test and post-test responses it was also found that students in both groups gave clear examples of historical inferences in order to make meaning of the end of the post-test story. From the interviews with the teachers, we were able to link this particular feature to the history class running parallel to our intervention. However, the three features distinguishing the DBC responses indicate a contribution from the intervention that we designed. More or less closely, these strategies relate to the teaching that occurred in the DBC classrooms. However, since these aspects of data have not been coded and tested by, for instance, independent raters, they should be treated as a provisional aspect of the intervention’s outcome and should as such be used as a hypothesis in upcoming studies.

**Statistical analysis of the test responses**

The statistical analysis of the test responses provides an additional validation of the differences suggested by the qualitative conceptualisations of students’ progress in the DBC group.

<table>
<thead>
<tr>
<th></th>
<th>Pre-test</th>
<th>Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DBC</strong></td>
<td>M = 2.105 SD = 0.658</td>
<td>M = 2.211 SD = 0.713</td>
</tr>
<tr>
<td>N = 19</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MBC</strong></td>
<td>M = 2.111 SD = 0.676</td>
<td>M = 1.778 SD = 0.732</td>
</tr>
<tr>
<td>N = 18</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
As shown in Table 1, the pre-test differences were marginal (DBC: M = 2.105, MBC: M = 2.111), while the post-test differences were noticeable (DBC: M = 2.211, MBC: M = 1.778). An analysis of covariance (ANCOVA) was conducted to evaluate the effectiveness of the intervention. The pre-test scores were used as the covariate, the condition (DBC vs. MBC) as an independent variable and the post-test scores as a dependent variable. Preliminary checks were made to ensure that assumptions of linearity, homogeneity of variances and homogeneity of regression slopes were not violated. After an adjustment for the pre-test scores, the analysis showed that the difference between the two groups in the post-test scores was not statistically significant \([F(1, 34) = 3.95, p = .055]\). The partial eta squared for effect size was .104 (a moderate to large effect).

The \(p\)-value is close to but still slightly above .05, which means that generalisations beyond the group level would be speculative. Yet, since the sample size is small and the effect size is moderate to large, it is still feasible not to discard the hypothesis altogether about the instructional benefits of the DBC compared with the MBC.

The mean scores for the DBC group increased by 5% \((2.105 \rightarrow 2.211)\), whereas the means for the MBC group decreased by 16% \((2.111 \rightarrow 1.778)\). This may seem to indicate that the MBC students’ reading skills decreased over the six weeks. The statistical analysis cannot reject this possibility. However, since the qualitative analysis demonstrated that the DBC group had acquired a number of new strategies in short story reading, a more reasonable explanation is that the post-test was more difficult than the pre-test. Within-group comparisons are therefore less informative than a between-group comparison. The two tests assess similar types of reading performances, even if one turns out slightly more difficult than the other. We thus argue that the combination of qualitative and quantitative analyses provides modest (although not generalisable) support for the instructional value of the dialogue-based teaching design. Nevertheless, if the pre-test and post-test are not comparable this is a methodological defect which needs to be addressed when preparing future studies.

**Students’ experiences of the instructional frame**

Three items in the questionnaire specifically targeted the students’ experience of learning about reading during the intervention. Their task was to rate, on a five-point Likert scale, the extent to which they agreed with the following statements: Item 5 “I believe I have learned a lot during this project”; item 10 “This project has

<table>
<thead>
<tr>
<th>Group membership</th>
<th>Mean</th>
<th>N</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBC</td>
<td>11.25</td>
<td>20</td>
<td>1.77</td>
</tr>
<tr>
<td>MBC</td>
<td>11.44</td>
<td>16</td>
<td>1.67</td>
</tr>
<tr>
<td>Total</td>
<td>11.33</td>
<td>36</td>
<td>1.71</td>
</tr>
</tbody>
</table>
made me better at understanding short stories”; and item 12 “During this project I have learned things that make me a better reader”. A rating of 1 on the scale meant “strongly disagree” and a rating of 5 meant “strongly agree”. Item 5 is obviously open to interpretation: learned what exactly? We surmise, however, that the students, in situ, perceive its orientation adjacently to the way they perceive the orientation of items 10 and 12. In order to produce a general representation of the prevalent attitudes to learning about reading, the three items were combined to form a 3–15 scale index. This index was then measured as a factor of the instructional frame.

Since the scale runs from 3 through 15 the median value is 9. Thus, mean comparisons indicate that both groups have a generally positive experience of having developed their reading skills. The seemingly stronger positive effect in the MBC is, however, not statistically significant ($p = .75$).

A different set of items in the questionnaire specifically targeted students’ experience of familiarity with, and attitude to, the particular substance and form of teaching (items 1, 4, 7 and 8, see App. II). Our first hypothesis here was that all students had little experience of reading short stories or of working on texts in half-class groups. However, if the dialogue-based teaching was as rare as we suspected, and if the more traditional monologic teaching had dominated the students’ previous schooling, then students’ in the DBC group would experience a much more distinct breach when subjected to the intervention teaching than students in the MBC group. This was our second hypothesis, which was tested by item 4. While the first hypothesis was supported by the questionnaire data (see Table 3), the second one was not (see Table 4). The empirical data even seemed to hint at the opposite, i.e. that students in the MBC group experienced a more distinct breach from previous schooling than students in the DBC group, although the mean difference was not statistically significant ($p = .74$).

Thus, according to the students an instructional frame based on dialogue, comprehension strategies and extended writing tasks is evidently not experienced as more deviant from the kind of teaching they are familiar with from earlier schooling than an instructional frame based on silent, individual work on short answer questions and book reports. This conclusion was unexpected. Yet the answers to

Table 3. Means of familiarity with and attitude to the substance and form of teaching (all students) (scale 1–5)

<table>
<thead>
<tr>
<th></th>
<th>Item 1 “I am used to reading short stories from prior experience”</th>
<th>Item 4 “This way of working with texts in the classroom is different from what I am used to…”</th>
<th>Item 7 “In Swedish we often work in half-class groups…”</th>
<th>Item 8 “I think it is good to be divided into half-class groups…”</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>Valid 36</td>
<td>36</td>
<td>36</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>Mean 2.56</td>
<td>3.89</td>
<td>2.36</td>
<td>4.58</td>
</tr>
</tbody>
</table>
items 7 and 8 may suggest at least a partial explanation. Both the MBC and the DBC were set up in half-class groups (generally comprising 10–12 students). Being unfamiliar with reading instruction in this intense form, and at the same time appreciating receiving both more personal teacher time and a more peaceful study environment, the students in the MBC group also experienced a distinct alteration from earlier reading instruction. Moreover, this alteration was clearly seen as beneficial to their reading development, as suggested in Table 2 above, as well as by their accounts of having learned about reading in the open-ended item.

**Students’ description of their learning outcomes**

The last item in the questionnaire served to demonstrate how the students themselves articulated what they had learned about reading during the past six weeks. The question read as follows: “If you, in your own words, were to describe what you have learned during this project, how would you then describe it?”. Responses were segmented and assigned to categories of reported learning outcomes using a simple coding frame of six different categories.

Table 5. Students’ own accounts of their learning about reading (n = 36)

<table>
<thead>
<tr>
<th>Categories of learning outcomes</th>
<th>DBC (n) Total = 20</th>
<th>DBC (prop.)</th>
<th>MBC (n) Total = 16</th>
<th>MBC (prop.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Mentions the general development of reading skills, e.g. “to understand short stories” or “to understand texts”</td>
<td>8</td>
<td>40%</td>
<td>6</td>
<td>37.5%</td>
</tr>
<tr>
<td>2 Mentions specified new knowledge related to the reading process, e.g. “to make predictions”, “to summarise”</td>
<td>8</td>
<td>40%</td>
<td>1</td>
<td>6%</td>
</tr>
<tr>
<td>3 Mentions general new knowledge related to the reading process, e.g. “to read between the lines”, “reading comprehension”</td>
<td>7</td>
<td>35%</td>
<td>5</td>
<td>31%</td>
</tr>
<tr>
<td>4 Mentions specified new knowledge related to text, e.g. “that short stories can be difficult”</td>
<td>3</td>
<td>15%</td>
<td>2</td>
<td>13%</td>
</tr>
<tr>
<td>5 Mentions new knowledge related to the writing process, e.g. “to write short stories”</td>
<td>1</td>
<td>5%</td>
<td>1</td>
<td>6%</td>
</tr>
<tr>
<td>6 Mentions other new knowledge, e.g. “new words”</td>
<td>2</td>
<td>10%</td>
<td>4</td>
<td>25%</td>
</tr>
</tbody>
</table>
As shown in Table 5, students from the DBC and MBC groups demonstrate a fairly similar pattern when it comes to describing their newly acquired knowledge about reading, with the exception of one striking difference. The DBC students demonstrate a decidedly stronger inclination to specify this knowledge in subject-matter-relevant terminology (category no. 2). Here is an example from one of the DBC students which displays the qualitative difference between responses of different categories.

One can understand texts easier now, I have learned to make predictions and to draw conclusions and to question the text. It was fun (SDBC4).

The opening clause, being a general remark on the student’s experience of reading development, represents category no. 1, whereas the following statement, up until the first punctuation, provides a finer distinction of specific comprehension strategies with which the student has become acquainted. These are also easily tied to the teaching performed in the DBC group. Now, terminology in itself may certainly be memorised and used without connection to a deeper level of understanding and thus written remarks do not necessarily prove that an intellectual appropriation, corresponding to our pedagogical objectives, has occurred. Yet we argue that mere access to a relevant terminology will undoubtedly provide students with a more adequate basis for developing a deep understanding of subject matter. From a theoretical perspective, it should also be noted that the gradual acquisition of new distinctions with which we organise our experiences is closely connected to a more comprehensive self-awareness (Marton 2006), and in this case to a more comprehensive understanding of an individual’s reading process.

A detailed examination also revealed that the DBC students’ written responses to the open-ended question were 25% longer than the accounts provided by the MBC students (M = 12.4 words/student compared to M = 9.9 words/student). Obviously, longer answers in themselves do not mean that students have learned more and the sample is, especially in this respect, very small. Yet the fact that the DBC responses are both longer and more subject-matter-specific suggests that the DBC contains promising pedagogical potential.

6. Discussion
In line with our two-fold research purpose, which was oriented both to the application of research methods and to the particular intervention effects, the discussion will begin with the methodological problems of studying the effects of instructional change. In our case, these problems fall into two categories: one related to specific data collection procedures and one related to the intervention study design in general. We then discuss some remaining promises of the intervention method that may justify further investigations.
Weaknesses of the data collection procedures

Although the qualitative text analysis revealed interesting differences on the group level, for example that the DBC group had acquired a set of new reading strategies, the increase in test scores was modest. This may indicate that the two text-assignment compounds was not a suitable instrument for exposing learning effects from the intervention. Our methods for testing the students’ reading skills should therefore be developed in a way that facilitates greater consistency between different types of analyses. One way of pursuing such standardisation would be to adapt both texts and types of assignments to those used in the reading literacy section of the grade 9 national tests in Swedish. Yet, while such an adjustment to national policy would at least be perceived as strengthening our instructional cause, it remains a matter of discussion whether those tests would also provide students with a solid opportunity to display the kind of active and creative readership that our intervention hopes to foster. For that reason, some additional testing procedures will be necessary.

Further, the gradual socialisation into new and more qualified ways of reading and interpreting literature is without doubt a long-drawn-out process. The six weeks dedicated here to reading and discussing three short stories are then perhaps not enough time spent on instruction for the slowly changing patterns of response to be detected by the standards of our test instruments. Even with more sophisticated methods for testing the students’ reading skills, an extended intervention period would most likely be preferable.

A third aspect related to the range of the measurement is the degree to which one is testing what the students can actually perform on their own. In the testing procedures in our study, the teachers read the stories aloud, thereby providing many students with better opportunities to focus on interpreting the text rather than on decoding it. When comparing the effectiveness of different instructional approaches in grade 5 classrooms, McKeown et al. (2009) noted that “it would be desirable to measure further transfer, for example, what students could do if given a text to read completely on their own” (McKeown et al. 2009, 245). The emphasis placed by McKeown et al. on “to read completely on their own” is interesting in our case since our aim is to prepare the students to comprehend texts without the aid of teachers’ oral reading. However, the notion that teachers’ oral reading would unlock the text for students’ interpretations by freeing them from decoding is of course but a simplification of the actual case. First of all, the teachers’ oral reading of stories will regulate and conform the pace of reading and interpretation, thereby providing little opportunity for students to adapt the pace of reading to their individual linguistic, emotional and intellectual responses to the text. Second, teachers’ oral reading is by no means a neutral transmission of textual content. On the contrary, it significantly influences the students’ comprehension and interpretation (c.f. Greene Brabham
and Lynch-Brown 2002). For these reasons, a test situation like the one used here, with the particular purpose of exposing the effects of during-reading instruction, would perhaps be more suitable without the aid of teachers’ oral reading.

Finally, in several previous intervention studies on reading development the quality of students’ learning, as a result of the particular intervention, is assessed by their ability to perform operations very similar to ones that were taught during the intervention. Likewise, the results of an intervention are often measured using the same texts that were given during the intervention itself. The analysis by McKeown et al., for example, partly relies on dependent measures (e.g. recognition and recall of text content), whereas our analysis focuses on an independent test situation. Without this independent measure, one cannot conclude whether the intervention has affected the students’ general reading skills, or only their understanding of the text in use. These should obviously be regarded as methodological shortcomings only as far as one is particularly interested in the more general effects, which is not necessarily always the case.

**Weaknesses of the intervention study design**

Carlisle et al. (2011) points out that research on the effectiveness of teachers’ instruction needs to focus not simply on aggregates or averages of already theoretically favoured instructional models in order to explain the variance in students’ progress in reading. Instead, researchers should focus on teachers’ actual activities in the classroom and measure how emphasis on different dimensions of instruction accounts for variance in students’ reading comprehension achievements. The problem that may be encountered in the first case is that even if significant gains in students’ reading comprehension may be associated with a relatively higher proportion of a favoured instructional element (e.g. authentic questions or high-level questions), it is not necessarily safe to determine that this particular element really is the variable that explains the variance in students’ achievements (c.f. Taylor et al. 2003). A similar critique of research on strategy approaches in general is summarised by McKeown et al. (2009, 222). This critique certainly touches on our study design in which an instructional frame consisting of three particular features is believed to account for students’ progress in reading and interpretation. Yet the alternative approach taken by Carlisle et al. has another problem, as revealed by their discussion. Even if some instructional strategies are found to have a stronger relationship with student gains in reading and some others a weaker relationship with student gains, the theoretical explanation of why some strategies are more effective than others is missing, or at least insufficient. The fact that earlier studies have reached a similar result is of course important, but does not assist in explaining the relationship between teachers’ teaching and students’ learning. It will, for instance, be difficult for teachers to construe the adequacy of a particular strategy.
when brought into the local and complex social setting of their own classroom if there is no firm theoretical basis to explain the cause-and-effect relationships.

**Remaining promises**

Bearing these limitations in mind, this study still offers some promising cues for future research and for instruction. McKeown et al. compares *strategy-based instruction* (focused on monitoring the individual’s reading process) with *content-based instruction* (focused on the textual content and meaning-making by questioning the text) and argue for the benefit of the latter over the former. Their critique of strategy-based instruction is that it encourages a focus on the strategies themselves rather than a focus on the text. In terms of terminology, our study would primarily be associated with strategy-based approaches, but the instructional point we wish to make is to emphasise a *balanced* teaching method (c.f. Pressley 2006; Weaver 1998) which uses strategies in order to promote students’ deep understanding of, and creativity with, texts rather than to learn about the strategies as properties in their own right. We thereby combine the imperative values of both approaches to form a meaningful whole. The students’ reported learning suggests that this is the case, although the teaching we designed proved not to be the only factor influencing their reading development.

Although further research is necessary, the instructional frame at issue that comprises a mixture of text-centred discussions, embedded instruction on comprehension strategies, and extended and challenging writing tasks, appears to provide a potentially rich environment for the improvement of reading and interpretation. Students subjected to this instructional frame display a more evident tendency to employ relevant interpretive strategies, and are better equipped to describe the qualitative change in their reading proficiency. Hence, the general implication is that we are on the right track with the development of teaching strategies, but that we need better instruments to detect their effects on students’ reading and interpretation.
Notes

1 This was verified by the students themselves in the questionnaire responses (items 1 and 4).
2 “Om någonting händer” by Håkan Nesser, “Över spåret” by Majgull Axelsson, and “Första snön” by Cecilia Davidsson.
References


Graham, Steve & Hebert, Michael (2010). Writing to read: evidence for how writing can improve reading. Washington, DC: Alliance for Excellent Education.


Appendix I

Pre-test – “A Day’s Wait”

1. The story is entitled “A Day’s Wait”. What kind of wait do you think the title is referring to? Explain what you think.
2. The boy in the story misinterprets something the doctor says. How would you like to describe this misinterpretation?
3. The text does not reveal what happens to the sick boy. What do you think? Write a couple of lines about what could possibly happen.

Post-test – “Far och jag” (“Father and I”)

1. The story is entitled “Far och jag”. What is your view on the relationship between the two main characters in the story? Describe what you think.
2. They spend a day together. How would you like to describe the I:s (that is the child’s) feelings during this day?
3. At the end of the story a black train with darkened cars comes rushing by on the railway. Explain your view on this train and why it is in the story.

Appendix II

Questionnaire H1101

The teaching you have had in Swedish during the last few weeks, in which you have read and worked with different short stories, has also been part of a research project. The research project is about reading and reading instruction. As we finish this we would like to ask you to respond to a questionnaire. The questionnaire contains a number of statements. You respond by marking whether you agree or not. Try to respond spontaneously.

Respond to the statements below by marking the point of the scale that best describes your opinion.

1. I am used to reading short stories from prior experience.

   □ Strongly disagree □ Disagree □ Neither □ Agree □ Strongly agree

2. I personally believe that I am a good reader.

   □ Strongly disagree □ Disagree □ Neither □ Agree □ Strongly agree

3. It has been exciting to work with short stories the way we have been doing now.

   □ Strongly disagree □ Disagree □ Neither □ Agree □ Strongly agree

4. This way of working with texts in the classroom is different from what I am used to from previous teaching of Swedish (e.g. in grades 5 and 6).

   □ Strongly disagree □ Disagree □ Neither □ Agree □ Strongly agree

5. I believe I have learned a lot during this project.

   □ Strongly disagree □ Disagree □ Neither □ Agree □ Strongly agree
6. The short stories that we read were interesting and worth reading.

[ ] Strongly disagree  [ ] Disagree  [ ] Neither  [ ] Agree  [ ] Strongly agree

7. In Swedish we often work in half-class groups similar to the way we have been doing during this project.

[ ] Strongly disagree  [ ] Disagree  [ ] Neither  [ ] Agree  [ ] Strongly agree

8. I think it is good to be divided into half-class groups when you read and work with short stories in the way we have been doing during this project.

[ ] Strongly disagree  [ ] Disagree  [ ] Neither  [ ] Agree  [ ] Strongly agree

9. I think the short stories that we read in this project were difficult to understand.

[ ] Strongly disagree  [ ] Disagree  [ ] Neither  [ ] Agree  [ ] Strongly agree

10. This project has made me better at understanding short stories.

[ ] Strongly disagree  [ ] Disagree  [ ] Neither  [ ] Agree  [ ] Strongly agree

11. Most of the time I find it easy to read and understand texts.

[ ] Strongly disagree  [ ] Disagree  [ ] Neither  [ ] Agree  [ ] Strongly agree

12. During this project I have learned things that make me a better reader.

[ ] Strongly disagree  [ ] Disagree  [ ] Neither  [ ] Agree  [ ] Strongly agree

13. If you, in your own words, were to describe what you have learned during this project, how would you then describe it?

___________________________________________________________________
___________________________________________________________________
___________________________________________________________________
___________________________________________________________________

Finally, we would also like you to answer a few questions about yourself.

14. In which group were you?  Blue  [ ] Green  [ ]

15. Gender?  Boy  [ ] Girl  [ ]

16. In which class are you?  7_____

Thank you for your response! Now you are finished.