IMPROVING READING COMPREHENSION BY ENHANCING METACOGNITIVE COMPETENCES: AN EVALUATION OF THE RECIPROCAL TEACHING METHOD

ANKE DEMMRICH

Dissertation zur Erlangung des Grades Dr. phil.

eingereicht bei der Humanwissenschaftlichen Fakultät

der Universität Potsdam

Potsdam

März 2005

Gutachter: Prof. Dr. Jürgen Baumert
Prof. Dr. Joachim C. Brunstein

Datum der Disputation: 10. Juni 2005

ACKNOWLEDGEMENTS

I would like to express my gratitude to the people that helped and supported me during the last years. It was quite a long way to go and without them, it wouldn't have been possible to conduct this study, analyze all the data and finally to complete this dissertation.

First, I would like to thank the participants of the study: the schools and the children who participated in the training. The children gave their best and worked very hard in their spare-time. It was not only fun to work with them, but also a very special, invaluable experience for me.

The dissertation was conducted at the Max Planck Institute for Human Development in Berlin and supervised by Dr. Cordula Artelt and Prof. Dr. Jürgen Baumert. I am grateful for the technical and financial support of the project by the Max Planck Institute, but more importantly, for the help of many people there. Vicariously I would like to mention two staff members that were especially important for organizational matters: Susannah Goss and Andrea Derichs. I received additional support from Prof. Dr. Joachim C. Brunstein at the department of pedagogical psychology of the University of Potsdam. I would especially like to thank Dr. Cordula Artelt and Prof. Dr. Joachim C. Brunstein for the assistance, critical remarks and encouragement I received from them.

A lot of people were involved in data collection and coding. I wish to thank the trainers Alexis Engelhardt, Joachim Ellermann, Ulrike Hanse and Katleen Ueberfeld. Special thanks also go to Helvi Koch who trained three groups, Bianca Hudjetz and Christian Grundey who were not only involved as trainers, but also in rating the reading comprehension assessments. Andrea Lisker and Cathrin Lehmann were involved in the study as trainers, helped with organizational matters and rated video material and the post-training questionnaires. Barbara Diffendaffer helped with organization and rated the post-training questionnaires. Alexandra Biegay and Wenke Haase developed rating schemata for the videos and together with Nathalie Kruppa, Juliane Quilitzsch and Alexander Gerlach they rated the video material. At this point I want to mention Wenke Haases engagement and invaluable work on the video material. Her work culminated in a diploma thesis for which the video material was complemented with data about the trainers (their personality, behaviour and judgements of the training).

I do not want to close without appreciating the support I received from my family and friends during the past years. They encouraged me, and sometimes also reminded me to take a break and collect new power. I would also like to express the incredible debt of gratitude owed to my parents who were always supportive of me.

TABLE OF CONTENT

APPENDIX

LIST OF TABLES

LIST OF FIGURES

1. THEORY

1.1. The importance of reading ability

The ability to read written material is very important in the civilized world. The letters that compose words which are part of sentences and paragraphs in a book or newspaper, leaflet or advertisement, timetable at a bus station or information material in a hotel serve a variety of tasks. Writing is used to communicate information, express feelings, tell stories, attract attention or pass on knowledge to other people. Written language is a very effective means of communication; it does not rely on the presence of author and addressee at the same time, can be produced in many copies easily and can reach a lot of different people.

In a society characterized by globalization and technological change, where knowledge is becoming increasingly important, reading ability is a key skill for active participation. Access to education, employment opportunities and participation in social settings depend on it. It is necessary for many occupations and a prerequisite for further and life-long learning (Elley, 1994). Other than for educational or informational purposes, people also read for another purpose: pleasure. Reading poetry or prose makes it possible to discover imaginary worlds; reading opens up a fictive perspective and enables people to identify with other characters, to try out different things with them, and so inspires the mind.

The International Adult Literacy Survey (OECD, 2000) found that the probability of having a high-skilled white-collar position is positively associated with high literacy skills and that people with low literacy skills have a higher risk of being unemployed. Literacy level is also a predictor for how well people do in the labor market in addition to and, more important, independently from their educational attainment. There also is a relationship between reading frequency and competent use of other media: people who read a lot are competent users of other media as well, whereas people who read only little or not at all also have deficits using other media (Stiftung Lesen, 2001).

Yet the findings of recent studies indicate a mismatch between the importance of the ability to read and the reading habits and skills of German students. In the PISA 2000 study (Baumert et al., 2001), a large-scale international study, the reading skills of 15-year old students were assessed in 32 countries, most of which are members of the OECD. German students scored below the mean of the other OECD countries

participating in the study. Almost 23% of German students showed comprehension skills at only very rudimentary levels. That is, these students were only able to locate information directly stated in the text if there was very little or no competing information, to detect the main idea of a text if it was repeatedly stated or overtly formulated and to generate simple connections between ideas stated in the text and common knowledge. This lack of knowledge and skills may impact on these students' chances of finding an apprenticeship and, later on, a job.

It is difficult to identify the main causes for the low performance of the German students; but the PISA study did provide a number of pointers. For instance, a path model to predict reading comprehension showed that interest in reading was an important predictor of text comprehension in addition to general cognitive ability, decoding fluency and knowledge about reading strategies (Artelt, Stanat, Schneider, & Schiefele, 2001). Interest in reading also seems to be a precondition for children spending time with reading activities independently from what they read in school: reading for pleasure (e.g., literary texts like detective stories, fairy tales, etc.) and acquiring knowledge about topics they are interested in (e.g., reading a sports magazine or searching for information about dinosaurs in the internet). At the same time, a large proportion of students stated in a questionnaire that they never read for pleasure (42%; with a higher percentage for boys than for girls).

This study also provided more evidence for the importance of reading literacy: when achievement in the PISA mathematics test was predicted, students' reading skills proved to be the most important predictor (path coefficient of .55, see Klieme, Neubrand, & Lüdtke, 2001). The other factors included in the model were general cognitive ability, gender, mathematical self-concept and socio-economic status of the parents. This finding underlines the importance of reading skills; they are a prerequisite for achievement in other school subjects and not just important in language arts classes.

In summary, reading comprehension is a skill necessary to succeed in educational and vocational settings and to participate in society. Yet students' reading abilities are often alarmingly low. Improving students' reading abilities therefore is a very important task. But how can the ability to read and comprehend textual material be improved? What should be the focus of training programs, and which factors determine the effectiveness of these programs? In the next section, these questions will be addressed by considering a number of factors that influence reading ability.

1.2. How can reading comprehension be promoted?

Numerous factors that contribute to the ability to decode and comprehend texts have been identified; they include general cognitive ability, prior knowledge, decoding fluency, knowledge about reading strategies, interest in reading, the goal of the reading activity and features of the text. Some of these factors are characteristics of the reader himself, others involve external factors. To structure the many factors that have an influence on how well a text is comprehended, a tetrahedron model representation originally developed by Jenkins (1979) for learning experiments and adapted by Artelt (2004) for the domain of reading may be useful. The model (see figure 1) contains four categories of determinants of reading comprehension: two that are related to the text itself – characteristics of the text and task demands – and two that deal with the reader – characteristics and activities of the reader.

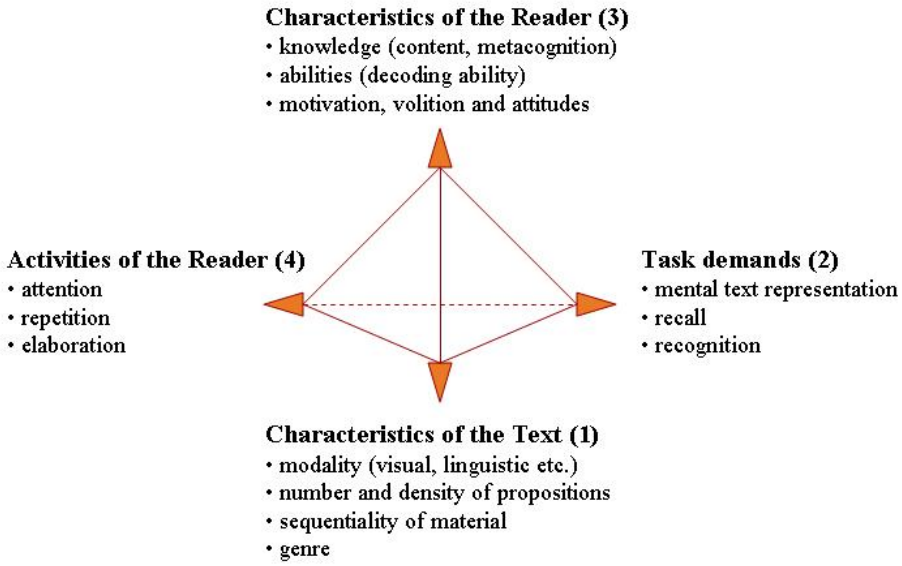


Figure 1. Determinants of Reading Comprehension (Artelt, 2004).

(1) Text characteristics that determine the difficulty of a text and influence how well texts are understood are, for instance, how the text is presented, whether it is illustrated with pictures or schemata, how many propositions it contains and in what order the ideas are presented, the genre of the text, etc. (2) It is also important to consider the goal of the reading activity – what is demanded by the task. Imagine that students are required to read a chapter of a history book for a test the following day. The demands of the test may be free recall of the content ("Write a summary of

Chapter 7"), cued recall ("What were the main causes of World War II?") or recognition in the form of multiple-choice questions. (3) Which characteristics of readers impact on the ability to comprehend texts? Very important is what kind and how much knowledge (content knowledge about the topic of the text and metacognitive knowledge) a reader possesses, how well he can decode words and, of course, his general cognitive abilities. Aside from these, motivation to read, volition and attitudes also play a role. (4) All these characteristics also have an influence on the activities the reader engages in when he reads: how attentive he is and what kind of strategic activities like repeating text content or elaborative efforts he undertakes in order to understand and try to remember the text.

The components of the model interact within and between the four areas, of course. Depending on the task demands, students engage in different activities when working with a text, and their representation and memory of the text will differ greatly. Both free and cued recall require the reader to actively construct a mental representation of the text that allows him to retrieve information from memory at the time of test-taking. This may not be necessary in the case of recognition; here it is possible that the presentation of the correct answer as one of the response alternatives automatically activates the relevant knowledge. Students may therefore adjust their reading activities accordingly: in the case of recognition it may be sufficient to deploy surface-level processing strategies, whereas free and cued recall will require them to engage in deep-level processes like elaborative strategies. Interest in the topic and prior knowledge also play a role: it has been shown that both factors induce deeper-level processing.

A lot of research has been devoted to the influence of textual features on reading comprehension. For example, reading time is dependent on the syntactical complexity of sentences and on the number of propositions the text contains (Richter & Christmann, 2002). Texts in which the topic remains the same or changes in the topic are marked by syntactic cues are easier to process (Schnotz, 1994). For the construction of propositions, co-referential relations have proven to be important (van Dijk and Kintsch, 1983); the formation of macro-propositions is fostered by advanced organizers, topic sentences, summaries, headings, comparisons, etc. (see Schnotz, 1994).

On the part of the reader, numerous abilities, skills and attitudes are of relevance to the complex skill of reading. The theoretical part of this thesis will focus on some of the most important factors. Based on cognitive psychological research, Richter and Christmann (2002) conclude that decoding processes (speed and efficiency of

identification of words and activation of their meaning), working memory capacity and prior knowledge are the most important factors. Other researchers (Baumert et al., 2001; Schiefele, 1996) have also underlined the importance of general intelligence, motivational variables, metacognitive knowledge and strategy knowledge. Various models will be introduced and discussed to explain the impact of the many different factors involved in reading and the complex interactions between them. I will start with a model of text comprehension that helps to explain the component processes that occur in reading comprehension and illustrates how text characteristics interact with the knowledge, expectations and beliefs of the reader: the situation model of text comprehension by van Dijk and Kintsch (1983). The main components of the memory system and its content will then be described, concentrating on a very special part of memory - metamemory.

The reader actively engages in activities that draw upon his knowledge and skills when trying to read, comprehend and remember texts. The most efficient reading activities are strategic activities, that is, application of techniques that allow the reading goal to be attained with minimal effort and maximal success. The functions and applicability of strategies will be discussed and different classification systems will be introduced before a model that addresses a variety of components relevant to text comprehension, the *Good Strategy User* (Schneider & Pressley, 1997), is presented. This model is a specific instantiation of the *Good Information Processor Model* (Pressley, Borkowski, & Schneider, 1989) and serves to introduce some very important concepts (strategies and strategic behavior, metacognition) and to explain how the various cognitive, motivational and situational components interact. Then, a few motivational issues will be addressed.

The learning context is also of interest: the school setting, instructional methods and the interactive processes that occur during instruction. A short overview of developmental changes that occur in the structural and procedural components involved on the part of the reader serves to complete the picture and to provide information useful for the last section, in which I will try to answer the question of what form an ideal training program should take to help students engage in activities leading to better text comprehension.

1.2.1. A process model of text comprehension

What is meant by reading ability? There is widespread agreement nowadays that text comprehension is a cognitive activity whereby information is actively re-constructed by the reader. It is viewed as the result of the interaction between the reader and the text, depending on the context in which the text is read, the purpose or goal of reading, the knowledge the reader possesses and textual features. The information in the text is associated with prior knowledge and incorporated into the existing knowledge network of the reader. Thereby information can be added to the reader's network of knowledge, but there can be changes in the structure of the network as well.

Over the past thirty years, a great deal of research in cognitive psychology has been carried out to describe and explain the complex processes involved in reading and interindividual differences in reading ability. The process components that people with good reading skills master better than people who read poorly formed the starting point for this research. Despite differences in emphasis, all modern theories assume hierarchical models of reading comprehension. Letter and word recognition are the basal analytical processes on the lowest level, semantic and syntactic analyses of sequences of words form the next level of the hierarchy, and the integration of larger parts of text and the construction of a coherent structure as well as the formation of a global representation and interpretation of the text represent the highest level of the hierarchy.

There are two main views as to how these levels are related. Proponents of the first kind of model, (e.g., Fodor, 1983), argue that the processes involved in reading comprehension are autonomous and function independently from one another. Processes higher in the hierarchy start only after processing at lower levels has been completed. Others advocate interactionist models, in which text comprehension is viewed as an interaction between the text (information) and the reader (and his knowledge). The main assumption of interactionist models is that there is no strict order of processes, but that processes on different levels overlap or occur in parallel. It is also assumed that there are various kinds of interactions between processes on the different levels and that bottom-up (text-driven processes) and top-down processes (knowledge-driven processes) are executed at the same time. A very influential and prominent example is the construction-integration model introduced by Kintsch (1982, 1998) and van Dijk and Kintsch (1983). To illustrate the processes that occur on

different levels of comprehension and to explain the complex interplay between processes on the different levels and the interaction of the reader's knowledge and expectations with the text, this model will be described in detail.

Five different process components are distinguished in the model: (1) construction of a propositional text representation and (2) formation of local coherence as lower-level components, and (3) formation of global coherence or macrostructure, (4) recognition of superstructures and (5) recognition of rhetorical strategies as process components on the higher hierarchy level. According to van Dijk and Kintsch (1983), the component processes occur mainly in parallel and there is feedback between lower-level process components and components higher in the hierarchy. The result of the complex interplay is an analogous, content-specific and vivid representation that consists of textual information and prior knowledge in an integrated form and is independent from linguistic structures. This representation, also called a situational model, is a form of a mental model (Johnson-Laird, 1983). Let me now consider the five process components that lead to the formation of a situational model in detail:

1. The preconditions for *constructing a propositional text representation* are word recognition and the comprehension of the meaning of sequences of words. Word recognition is primarily a visual process. According to (McClelland & Rumelhart, 1981), the identification of words already known by the reader is an alternating process of activation and inhibition of letters and words. Their meaning is then retrieved from long-term memory. Unknown words are identified phonologically or on the basis of morphological units (also called decoding). Sequences of words are connected with one another on the basis of their semantic relationships and are integrated into propositions (argument-relation structures) (Kintsch, 1982). The construction of semantic relations has priority during reading and is only supplemented by the syntactic analysis of the sentence(s), by identification of words with syntactical functions, if this proves to be necessary (Richter & Christmann, 2002). The semantic units of the propositions are appropriately semantically and syntactically annotated expressions of the text or newly established units if no slot in memory already exists. These propositions are constructed by an interplay between the textual information, the way in which this information is expressed (syntax) and the organizational principles retrieved from the reader's knowledge base.

2. *Local coherence* is formed by discovering and constructing semantic relations between propositions and between sentences. To infer the relations between the various information stated in the text, the reader can draw either upon his knowledge base or upon cues provided in the text. Cues in the text can be co-references (word repetitions, pronomina, anaphora, cataphora or contiguity relations), topic-comment relations or conceptual relations. A lot of research has been conducted on text-related cues as to how propositions are related, in particular (see Garrod & Sanford, 1994; Schnotz, 1994 for an overview of co-reference and topic-comment strategy).
3. The result of forming a *global coherent text representation* is the *macrostructure* of a text. Sequences of propositions are connected and condensed and result in a global structure or a mental model consisting of a hierarchy of propositions. Some of the macro-rules used to condense information are selection, deletion and generalization. New information is added to the representation by the construction of inferences. The reader draws upon his prior knowledge to construct inferences: items that are strongly linked to the text are retrieved from long-term memory and become part of the text representation. There are also diverse textual cues that can foster the formation of a macrostructure and that tell the reader which portions of the text are likely to be important: among them are topic sentences, summaries, titles, comparisons, abstractions and examples (Schnotz, 1994).
4. *Superstructures* in the sense of schemata explain the global organization of texts with specific conventional structures, like narrative texts, research reports, advertisements or laws. It is assumed that superstructures are saved in long-term memory and guide the formation of macrostructures (top-down strategy, see Richter & Christmann, 2002). Research on superstructures for narrative texts has shown that reading time is longer when narrative texts, a text type to which much research has been devoted (see Mandler, 1984 for a review), differ from their conventional structure (Kintsch & Kozminsky, 1977).
5. Another important tool for the adequate interpretation of the meaning of a text is an understanding of rhetorical, stylistic and argumentative strategies. These are often applied to accentuate certain information in the text and are relevant for drawing higher-level inferences and understanding, for instance, metaphor, irony and humor.

In a later extension of this model, Kintsch (1998) also made some assumptions about memory. All of the cognitive work occurs in working memory, which has direct access to both current input and short-term memory containing the currently dominant macropropositions of the text in the form of chunks of atomic units. When a new representation is formed in working memory, the old one remains available in episodic text memory. In this way, a sequence of interrelated text propositions is constructed in episodic text memory during reading.

Many researchers have explored the component processes in which good and poor readers differ. There are three groups of influences that are especially relevant: processes of word recognition, working memory capacity and prior knowledge. Most research has dealt with processes on the level of words. Lexical access, that is assigning meaning to words, and vocabulary knowledge are very highly correlated with general reading abilities (Graves, 1989; Jackson & McClelland, 1979). For processes higher in the hierarchy, like the construction of a situational model of the text, the reader's prior knowledge is of great importance. Content knowledge about the topic of the text is a very influential predictor of text comprehension: in the frequently cited studies by Voss et al. (1985), prior knowledge about baseball predicted comprehension even after reading skills had been controlled for. Similarly, Körkel (1987) showed that 3rd grade students with expertise in the domain of soccer outperformed even 7th graders with less knowledge on that topic on measures of memory, construction of appropriate inferences and detection of contradictions after reading text passages about soccer. There is also evidence that good and poor readers differ in their use of the context to identify words: good readers rely more on lexical access, which is an automated, fast and context-independent process, to identify words. All possible meanings of the words are activated and good readers are able to inhibit irrelevant meanings very quickly (Gernsbacher & Faust, 1991). In contrast, it seems that poor readers are able to compensate at least in part for their deficits in lexical access by relying on using the context of the sentence to discover the meaning of words (Perfetti, 1989).

Another factor on the part of the reader is the capacity of working memory, a characteristic of the reader that influences processes on higher (e.g., construction of a situational model) as well as lower (e.g., syntactic analysis) levels of the hierarchy. For instance, strong evidence that limited working memory capacity plays a crucial role in text comprehension has been provided by research with "garden path" sentences

(sentences containing syntactic ambiguities that are resolved at the end of the sentence), where large differences in reading time and comprehension are always observed between good and poor readers. Good readers seem to be able to process multiple interpretations simultaneously in working memory until the ambiguity is resolved (King & Just, 1991). There are also differences with respect to basic perceptual processes: poor readers have longer fixation times (Everatt & Underwood, 1994) and also show more regressions, that is, they jump back to a word already seen (Olson, Kliegl & Davidson, 1983), but it is very unlikely that these processes are causal factors that help to explain low performance.

In summary, according to Kintsch (1998), reading comprehension is the result of a text-driven construction process and a knowledge-driven integration process. Deliberate control is only necessary when the information given in the text interferes with the knowledge base of the reader, or when it is difficult to form a coherent representation of the text. Depending on the goal and ability of the reader, there are different ways to process the text and different forms of representation and levels of understanding. When students have to learn a poem to recite it, the representation will focus more on the surface of the text than when they read a text about history and have to comprehend the connections between certain events – the latter representation will be more elaborated. The situational representation is the most elaborated form of representation. This is accomplished by creating an interconnection between the textual information and prior knowledge and can also include additional information generated by inferences.

1.2.2. Memory and knowledge

People construct new knowledge and understandings based on what they already know. This knowledge provides the framework within which new information is incorporated. Teachers' and students' activities, their goals and subjective theories about learning and knowledge, as well as the learning material itself also impact on students' learning process. However, learning (and also learning from texts) should result in changes in the learner's knowledge or skills. Knowledge is organized, processed and stored in memory. Cognitive psychology has contributed to the understanding of learning by providing models of how memory is organized and how information is processed. Information processing theory, which has been the dominant theory of learning and memory for the past twenty years, can provide the framework

and some basic concepts and processes to explain how information is absorbed, selected, stored or forgotten, and how teachers and students themselves can take advantage of this process to help students retain critical information and skills (Slavin, 1994).

According to global models of information processing, stimuli are first processed by *sensory registers*. Those are modality-specific, they have a large capacity but are only of very short duration (less than 5 seconds). Some of this information is attended to and processed further by *short-term memory*. Some of the information processed in the short-term store is transferred to *long-term memory*, and other information is forgotten, decays or is destroyed by interference. Short-term memory, or *working memory*, is a system of limited capacity and short duration (less than 15 seconds). If the information is processed in some way, it can become part of long-term memory – a system with unlimited capacity and very long duration. To access the content of long-term memory the knowledge must be retrieved. It thus becomes part of short-term memory again and is available for further use.

The content of the knowledge base stored in *long-term memory* exists in declarative or procedural form. *Declarative knowledge* is factual knowledge about the world; it can be semantic or episodic. *Procedural knowledge* consists of algorithms that capture how something works. The current conditions in reality are checked with the production rules to decide which rule will be executed. In the case that multiple productions match the current conditions, conflict-solving rules are applied to choose one production. There is no exact equivalence between procedural knowledge and strategies, but a very close relationship (see Schneider & Pressley, 1997). What information is stored in memory, and how this occurs, depends on the attention, perception and processing that occurs at the time of learning (Craik & Lockhart, 1972). It is assumed that there are various levels of processing, ranging from shallow or physical analysis of the stimulus to deep or semantic analysis. Craik and Lockhart (1972) assumed that the level of processing has a large effect on the memorability of the information; with deeper levels of analysis producing more elaborate, longer lasting, and stronger memory traces than shallow levels of analysis (see also Eysenck & Keane, 2000).

The concept of *short-term* or *working memory* is so important because some memory mechanisms and capacities (based on biological substrates) are more fundamental than others; they underlie and can explain higher-order memory

components and their operation, like strategies (Schneider & Pressley, 1997). Many researchers have been concerned with *short-term* or *working memory*; some have conceptualized it as a unitary system (Atkinson & Shiffrin, 1968), while others assume that it consists of several subsystems. For instance, Baddeley (1986, 1992) proposed that working memory consists of a central executive and at least two slave systems, the phonological loop and the visuo-spatial sketchpad. There are different views as to whether working memory is but one part of short-term memory or if the two terms actually subsume the same construct. Working memory focuses more on an active, processing capacity that is used to transform information; whereas short-term memory is conceptualized as a rather passive system or as a "container" model and is believed to involve storage and reproduction of information. Case (1972), for instance, assumed that operating space and storage space together form the total processing space, which is limited. However, all researchers agree upon one central feature of this memory system: its limited capacity. In contrast to the unimaginably large capacity of long-term memory, the working space of memory (short-term memory) is very limited. Short-term, or working memory, capacity poses the constraints within which all strategies, metacognition and knowledge must operate; it determines how much can be consciously contemplated at one moment. Short-term memory capacity is also attentional capacity. One measure of working memory capacity is the memory span; which is conceptualized as the maximum number of elements one is able to hold in memory. This is usually assessed by presenting a list of items and asking the subject to recall them. Miller, (1956) found that the memory span of adults is usually 7 ± 2 chunks representing meaningful units like words, numbers or letters. Daneman and Carpenter (1980) assessed working memory capacity with the reading span task. This is a dual task requiring simultaneous processing and storage. The reading span task and similar measures of working memory capacity have successfully predicted performance in many cognitive tasks, such as reading and listening comprehension, vocabulary learning, following directions, note-taking and complex learning (Engle, 1996).

Working memory, with its limited capacity, is so important because it is assumed that all conscious processing occurs there: the system is used for the selection and temporary activation of information from the long-term memory (retrieval), for the manipulation of elements activated by means of transformation, and for coordination of the slave systems (if the theoretical model assumes these to exist; e.g., Baddeley 1986, 1992). It is also responsible for regulative and control activities and allocation of time.

Performance on complex tasks largely depends on working memory capacity, as does reading ability and reasoning about moral dilemmas. In a study by Cariglia-Bull and Pressley (1990), in which the dependent variable was children's ability to execute a capacity-demanding imagery representational strategy for sentence learning, it was possible to predict whether the children would benefit from the imagery instruction on the basis of their working memory capacity, measured with classic working memory span tasks (e.g., word span).

A theory of text comprehension that incorporates the working memory construct and stresses the importance of limited capacity is the capacity theory of Just and Carpenter (1992). According to them, working memory is used for both storage and processing during text comprehension. Because storage demands need to be kept manageable during language processing, each word is thoroughly processed when it is first encountered instead of being stored for further processing. Central assumptions of the theory are that working memory capacity is strictly limited and that there are individual differences in working memory capacity, with substantial effects on language comprehension. A lot of empirical evidence congruent with the predictions of capacity theory has been collected. For instance, individual differences help to determine whether meaning affects initial syntactic parsing; only subjects with greater capacity (indicated by higher reading spans) took advantage of inanimate nouns in sentences with unexpected syntactic structures, resulting in shorter reading times. Also, subjects with higher reading spans seem to be able to retain both or all meanings of sentences with syntactic ambiguities in memory as they process the sentences; resulting in slower processing of ambiguous versus non-ambiguous sentences; especially at the passages where the ambiguities are resolved. In contrast, subjects with less working memory capacity do not take advantage of inanimate noun cues, nor do they seem to detect syntactic ambiguities in sentences.

It is not only working memory capacity that is important for comprehension, however: as mentioned above, people construct new knowledge and understandings based on what they already know. This knowledge provides the framework within which new information is incorporated. Basically, two kinds of knowledge can be distinguished. One kind of knowledge important for reading ability is content knowledge (referred to as prior knowledge). The other kind of knowledge that plays a role for text comprehension is metacognitive knowledge. This knowledge is positively related to comprehension activities and effective learning with texts in general. Because

of the great importance of metacognitive knowledge for learning from texts (Schiefele, 1996), the next section will focus on the many aspects that contribute to this type of knowledge. Many studies have shown that *prior (content) knowledge*, which is stored in long-term memory, plays a large role for reading comprehension. Children learn more when studying new information related to their prior knowledge than when studying content from an unfamiliar domain. This has been demonstrated, for instance, by Pearson, Hansen and Gordon (1979) in a study with 2nd graders. Here, experts in the domain of snakes were better at answering both text-explicit and text-implicit comprehension questions about short texts dealing with snakes than were novices in the subject. Similar results have been found by Körkel (1987) with 3rd and 7th graders reading text passages about soccer. Content knowledge impacts not only on the understanding of text, but also on the recall of categorizable lists or serial positions of figures on a chess board (Chi, 1978). Students' as well as adults' understanding of text is influenced not only by their prior knowledge of content, but also by their knowledge of the schematic structure of text (story grammar). When stories conform almost perfectly to ideal story grammar forms, developmental differences in recall are less pronounced than when stories deviate from the ideal structure (Schneider & Pressley, 1997). Furthermore, both children and adults tend to adjust stories at recall to make them consistent with story grammar (Mandler, 1978).

1.2.3. Metamemory

Metamemory is a special part, or kind of memory. The term "metamemory" was first used in 1971 by Flavell (1971) to refer to knowledge about one's own memory. It was a very global conception and covered all possible aspects of information storage and retrieval, including knowledge about memory functioning, limitations, difficulties and strategies.

In 1977, Flavell and Wellman presented a taxonomy of memory phenomena, including metamemory, as a framework for classifying studies focusing on developmental differences. They proposed two main categories for metamemory: sensitivity and variables. Sensitivity refers to knowledge of when memory activity is necessary. It is the awareness that a particular task in a specific setting requires the use of memory strategies. For variables, three subcategories were distinguished: person and task characteristics relevant to memory, and potentially applicable strategies. One example of a person variable is the mnemonic self-concept, the ideas a person has about

his own strengths and weaknesses as regards memory. Task characteristics are knowledge about factors that make tasks easier or harder, like familiarity of material, number of items to be remembered or available study time. Knowledge about encoding and retrieval strategies is also important. These three types of variables are not independent from one another; there is considerable overlap and, more importantly, interaction between the categories. There can, for example, be person by task or person by task by strategy interactions. The interactionist perspective is very important; the focus of research should be on the integration and synthesis of information about memory tasks and processes. "Metamemory should be viewed as the individual's theory of mind, which is ... likely to be a highly integrated set of notions, propositions and concepts" (Wellman, 1983).

Metamemory is not isolated from knowledge about other aspects of the mind. Somewhat later Flavell (1987) generalized the metamemory taxonomy to metacognition in general. As in the previous taxonomy, metacognition about persons, tasks and strategies was proposed. Another term was also introduced, that of metacognitive experiences. Those are occasions during cognitive processes when new insights about cognitions arise. Also, there is constant interaction between metacognitive knowledge, metacognitive experiences and cognitive behavior and goals. In summary, "metacognition refers to one's knowledge concerning one's own cognitive processes or products or anything related to them, e.g. the learning-relevant properties of information or data" (Flavell, 1976).

Other researchers have also developed ideas and specified theories about metacognition and described related phenomena. One important contribution was made by Brown (1987; Brown & DeLoache, 1978), who reconceptualized metamemory, focusing on "here and now memory monitoring" within the framework of the competent information processor model. The competent information processor is a person with an efficient executive that regulates cognitive behavior. Brown (1987) distinguished between stable and expressible knowledge about cognition, and the regulation of cognition (executive), which was described as unstable, not necessarily expressible and dependent on the concrete situation and task. Also, personality traits like achievement motivation or self-concept are part of metamemory in the respect that they include knowledge about the self as a learner. Analysis, planning, monitoring and evaluation are executive metacognitive processes. They serve to identify and characterize learning demands, help select appropriate learning strategies, supervise information processing,

monitor and evaluate the usage of strategies and learning outcomes, and decide whether to continue with the current strategies or to replace them with more efficient ones. Awareness of the limits of the system and strategies is also a function of the executive. Memory monitoring plays a large role in executive actions. The effects of metacognitive activities are of great importance for cognitive regulation.

Kluwe (1980, 1984) provided a more complete and differentiated description of metacognition. The frame for this was provided by general theories of information processing. In line with these general theories, he distinguished between knowledge of data ("knowing that") and knowledge of processes ("knowing how"). This corresponds to declarative and procedural knowledge, as outlined above. Kluwe's model of metacognition includes knowledge about minds in general and about individual differences in cognition. The focus is on procedural components of metacognition; these are the active agents. Kluwe identified control processes that evaluate ongoing cognition, or monitor current performance. Other control processes are responsible for regulation, allocation of attention, selection of other processes to be applied, or determining the intensity with which a strategy is applied.

Wellman (1983) distinguished between four types of metacognition. The first, metamemory, refers to factual long-term knowledge about cognitive tasks, processes and strategies. The second is called knowledge monitoring, and entails knowledge about the state of one's own current memory, its contents and limits. The third type, strategy regulation, concerns the control and regulation of cognitive processes or strategies. Finally, there are conscious affective feelings or states related to cognitive activity.

In 1988, Borkowski, Milstead and Hale (1988) published their componential theory of metamemory, in which metamemory was described in terms of a number of interactive, mutually related components. These components, listed in order of their developmental emergence, are specific strategy knowledge, relational strategy knowledge, general strategy knowledge and metamemory acquisition procedures (MAPs). This model, shown in figure 2, will be described in more detail, because it provides a useful framework for describing the measures of metamemory used in the experimental study and explains how they might be interrelated.

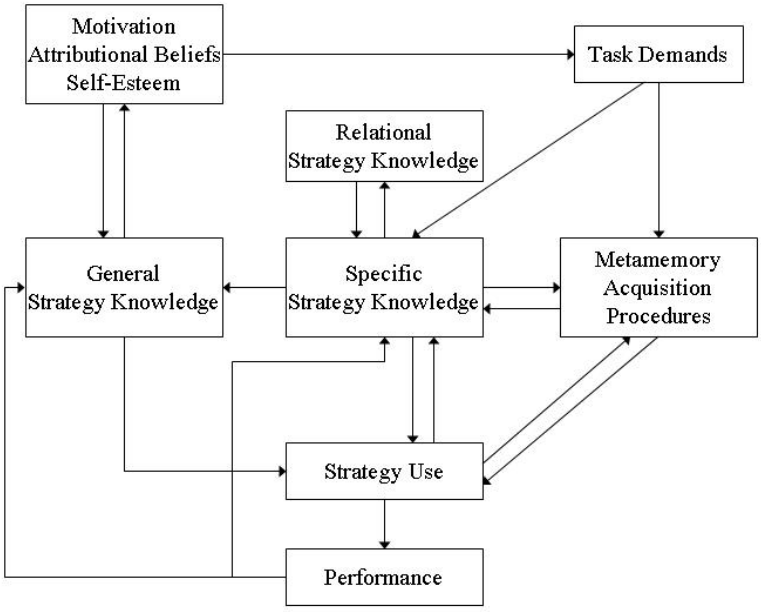


Figure 2. A model of mature metamemory (Borkowski et al., 1988).

Specific strategy knowledge is knowledge about particular strategies that involves an understanding of the following attributes of that strategy: its goals and objectives, the tasks for which it is appropriate, its range of applicability, the learning gains expected from consistent use of the strategy, the amount of effort that needs to be invested and whether the strategy is enjoyable or burdensome to use. Specific strategy knowledge accumulates slowly; the attributes of a strategy are acquired after guided prolonged instruction in its use or after more independent extensions of the strategy. After several strategies have been acquired, general attributes (common features) of all strategies become apparent. This knowledge about key attributes should make the acquisition of new strategies more rapid and durable. A dynamic, causal bidirectionality is assumed between actual strategy use and specific strategy knowledge: specific strategy knowledge guides the deployment of individual strategies and, in turn, the continued use of a strategy results in the expansion and refinement of knowledge about that strategy.

Relational strategy knowledge helps learners to understand the comparative merits associated with a number of specific strategies; it enables a classification system to be formed for contrasting the strengths and weaknesses of strategies. It provides useful comparative information for strategy selection and revision of these decisions (see also metamemory acquisition procedures).

General strategy knowledge reflects learners' understanding of the value of a planned and strategic approach, and of the fact that effort is required to apply strategies.

This knowledge is enhanced by feedback about a strategy's effectiveness. As already mentioned, there is an interaction between general and specific strategy knowledge; general attributes of strategies are acquired after several strategies have been learned, and this general strategy knowledge facilitates the acquisition of new strategies. The unique property of general strategy knowledge is its motivational character: it results in expectations of efficacy which motivate learners to confront challenging learning tasks.

The relationship between general strategy knowledge and motivational beliefs is bidirectional: learners who have frequently engaged in strategy-based learning are likely to understand the general utility of behaving strategically and believe in themselves as effective and efficient learners. At the same time, learners who attribute their success to effort rather than to ability or task characteristics (which are not controllable) are usually found to be both more strategic and higher in meta-memorial knowledge (see Kurtz & Borkowski, 1984). However, general strategy knowledge only promotes performance if learners believe that they have the necessary ability and skills. This underlines the importance of including attributional retraining in cognitive and metacognitive training procedures.

The mechanisms necessary to decide how and when to use a strategy are metamemory acquisition procedures (MAPs). These serve a dual function. First, they enhance the development of lower-level specific strategy knowledge by detecting insufficient strategy information and filling in gaps in instructions. Second, they supply regulating processes useful in implementing and modifying specific strategies by repeated checking and monitoring of a strategy's effectiveness. Chi (1987) used the term "meta-strategies" for these higher-order processes that are useful in making decisions about when and how to use a strategy or to switch to another one when the strategy deployed is ineffective. MAPs also interact with other components: repeated checking and monitoring of a strategy's effectiveness leads to knowledge about its benefits (specific strategy knowledge) and the relative difficulty involved in its implementation (relational strategy knowledge). Overt signs of metamemory acquisition procedures in operation are independent problem solving, task perseverance in the face of errors and trial and error behaviors.

One of the strengths of this model of metacognition is that it helps to explain the problem of generalization of strategies, or their transfer to new contexts or learning situations. It provides an explanation for the very interesting observation (see Kurtz, Reid, Borkowski, & Cavanaugh, 1982 and Kurtz & Borkowski, 1984) that knowledge

about a variety of task-irrelevant strategies predicts the acquisition and transfer of new, task-relevant strategies. The explanation of Borkowski et al. (1988) is as follows: the old knowledge provides the context in which a newly acquired strategy receives its interpretation and meaning. If there is no rich knowledge base, it may be the case that a strategy becomes welded to the training task. This occurs because the information about some of the strategy's attributes is incomplete or missing entirely. This results in problems of strategy generalization. Several studies that have included the training of higher-order executive processes (MAPs) as part of large-scale instructional packages (for instance Reciprocal Teaching, as developed by Palincsar & Brown, 1984) suggest that metamemory acquisition procedures guide the implementation and monitoring of lower-level strategies during generalization. Other researchers have also examined the problem of transfer; I will come back to this in the following chapter.

Metacognition has been described as a "fuzzy concept" (Wellman, 1983). Metacognition monitors and regulates cognition, but how is metacognition itself monitored and regulated? There is clearly a parallel to the *Homunculus* problem faced by Goethe's Faust: is there a human-like creature directing the human being? The many researchers who have contributed to this discussion do not agree in their definitions of what is cognitive and what is meta-cognitive, and even where they do agree, it is not easy to classify a strategy. However, metacognition is a constituting characteristic of most concepts of self-regulated learning. At the same time, regulation of learning is a key feature of the concept of metacognition. Learning is self-regulated if the learner himself is enabled to make the major decisions regarding if, what, when, how and why he learns (Weinert, 1984).

Metacognitive knowledge and metacognitive skills play an important role for reading comprehension and memory performance. For instance, Paris and colleagues (Paris, Cross, & Lipson, 1984; Paris & Jacobs, 1984) reported significant correlations between children's performance on different text comprehension tasks and their awareness of various cognitive components of reading, like goals and plans of reading, or comprehension and monitoring strategies. Large portions of the variance in comprehension and retention performance (40 and 50 percent, respectively) could be accounted for by metacognitive predictors (Hasselhorn & Koerkel, 1986). Among them were knowledge about text processing (metamemory), actual knowledge monitoring and on-line strategy regulation. The relationship between different components of metacognition and memory performance has been examined in a variety of studies; the

average correlation reported in a meta-analysis by Schneider (1985) was .41. A clear developmental trend was observed: the relationship between metacognition and memory becomes stronger with increasing age. However, this trend seems to be mediated by prior knowledge: as Chi (1983) pointed out, there are clear effects of the learner's knowledge base on memory performance and metacognitive skills.

The different aspects of metamemory are assessed using a variety of approaches and research paradigms. Cavanaugh and Perlmutter (1982) classified metamemory assessment into measures taken without concurrent memory assessment (independent measures), on the one hand, and measures collected simultaneously with (concurrent) measures of memory activity, on the other. Independent measures assess knowledge about factual memory knowledge. To this end, adults usually fill out metamemory questionnaires; for children, verbal interview approaches (Kreutzer, Leonard, & Flavell, 1975) and nonverbal techniques are used (for instance, rank ordering of pictured situations or presentation of various memory situations on video). Performance predictions, assessments of recall readiness, "feeling of knowing" judgments or judgments of learning as well as verbal protocol techniques are examples of concurrent measures of metamemory assessment.

1.2.4. Cognitive and metacognitive strategies

A very important component of metacognition is the knowledge of strategies. Schneider and Pressley (1997) define strategies as processes or sequences of processes that are potentially conscious and controllable. The use of strategies is always intentional in the sense that their application is goal directed, and that they aim to improve storage and retrieval of information as well as understanding. Strategies and skills are inextricably interconnected: "what might be a strategy early in one's career must often become a skill, and then an automatized skill, if one is to progress beyond that level" (Kirby, 1988). There are numerous strategies, and the various conceptions of metacognition differ in terms of the strategies they subsume (varying from metacognitive strategies only to cognitive strategies as well).

One way is to differentiate between external strategies (like keeping a shopping list) and internal strategies (like rehearsing a telephone number). Some strategies are partly external and partly internal, like sorting items into different piles to help memorize different groups of material. This classification was used, for instance, in the classic study of children's knowledge about memory by Kreutzer et al. (1975), who

found that kindergarten children reported only external strategies (most suggesting the manipulation of the object) to the skates question ("What could you do to remember to take your skates to kindergarten the next day?"); whereas children in grades 3 and 5 suggested more clear-cut and efficient internal strategies.

Another classification system, related to the depth-of-processing theory (Craik & Lockhart, 1972), was derived from an interview study. Marton and Saljö (1984) included questions on how the kind of learning activities students engaged in and the learning goals they had when required to read comprehension texts relates to levels of processing. A surface-level approach, in which learning focuses on memorizing specific facts and unrelated information with the goal of rehearsal at test-taking, is distinguished from a deep-level approach to learning, where students try to develop an understanding of the message of the text by discovering or constructing relations that exist between the information in the text and prior or world knowledge.

One very detailed system including different ways to classify strategies is the taxonomy of learning strategies by Friedrich and Mandl (1992). They list four major ways to classify learning strategies. The first is to distinguish between primary and supportive strategies, similar to the approach taken by Weinstein and Mayer (1986), who distinguished between strategies for organization of information processing and strategies dealing with the management of time, motivation and emotion. Primary strategies are those which deal directly with the information to be processed, ensuring that the information is understood, retained, remembered or transferred, thereby causing changes in cognitive structures and processes. Examples of primary strategies are summarizing, mnemonic techniques, means-end analysis, testing of hypotheses and planning by abstraction. Supportive strategies, on the other hand, initiate information processing, and maintain and regulate the process. Here self-motivative strategies, strategies that regulate attention and organization of time, metacognitive control of one's own learning process and selection of techniques or procedures that are suitable for the concrete learning situation are of interest. The second way to classify learning strategies relates to the range of applicability of the strategies. General strategies, which are applicable for almost all learning tasks and for all domains, and strategies with a medium grade of generality, applicable in a number of situations, are distinguished from specific strategies, which can be only applied in a specific type of learning situation. Examples of general strategies would be supportive strategies (self-motivation, time management, metacognitive control strategies) and problem-solving strategies (means-

end analyses or generating and testing of hypotheses); text-related strategies aiming at understanding the content are, in principle, applicable to various types of text with different content and would be classified as strategies with a medium grade of generality. Another (third) way to classify between strategies relates to the range of time over which a strategy extends, resulting in micro-, meso- and macro-strategies. The last method of classification that Friedrich and Mandl included in their taxonomy was to group strategies according to their function for information processing. Here memorizing strategies, elaboration strategies (analogies, relating information to prior knowledge, formulating ideas in one's own words), organization strategies (summarizing, grouping or presenting information graphically in forms of diagrams) and control strategies are distinguished.

A few years later, Friedrich (1995) distinguished between cognitive strategies, metacognitive strategies and strategies of resource management. The group of cognitive strategies comprises memorizing strategies like repeatedly reading aloud, memorizing using keywords or the method of places, elaboration strategies (construction, integration and transfer of information) and transformation strategies, that is transferring information to another medium. Planning (goal setting, formulating control questions), monitoring and regulation (adjusting the learning activity to the current state and goal of the learning situation) are regarded as metacognitive strategies. Examples of strategies concerning resource management, further divided into internal and external strategies (Wild & Schiefele, 1992), are the monitoring of effort and attention and systematic use of study time as internal strategies, and the creation of an optimal learning environment as an external strategy.

Another way of differentiating between strategies, which originates from current research on memory development and metamemory, is the categorization proposed by Schneider and Pressley (1997). Their major criteria for classification are content and applicability. The authors distinguish between three different types of strategies: task- and domain-specific strategies, structure-limited memory strategies and general strategies, and assume that there is interaction between them. Task- and domain-specific strategies can only be used in a particular domain, although they are often specific instantiations of more general strategies. An example would be the German first-letter mnemonic "Eine Alte Dumme Gans Hat Eier" (An old stupid goose has eggs) to remember the strings on the guitar, which are E, A, D, G, H, and E. A similar example for an english mnemonic is "Every Good Boy Deserves Favor" to

remember the notes that fall on the lines of the treble clef: E, G, B, D, and F. In contrast, structure-limited memory strategies can be applied to learning in a variety of content areas, but are useful only for material with a particular structure. For instance, rehearsal might work very well to recall items in simple lists, but not for pair-associative learning; here an imagery strategy creating meaningful connections between the two items would produce better results. Rehearsal of text is useful when the goal is to recite a poem in language class, but when the task is to interpret the poem, elaborating on its content and providing a short summary are more suitable for that goal. Other structure-limited memory strategies are elaboration, summarizing and various organizational techniques. Third and finally, general strategies are not specific memory aids, but support the use of memory and other strategies. Among them are monitoring, that is checking one's own performance, being attentive to the task, trying to find similarities between the specific situation and past experiences (activating prior content knowledge or episodic knowledge) and attending to the environment and looking for cues that could help solve the task.

According to Schneider and Pressley (1997), the successful use of all these strategies depends largely on a particular type of metacognition: metacognitive knowledge about specific strategies. This implies a general understanding of the utility of a strategy, how much effort the strategy requires in comparison to other strategies, and how much fun it is to use. This understanding evolves from many experiences in which that strategy was used. It is possible to have this kind of knowledge for any strategy. It is very important to acquire metacognitive knowledge about strategies in order to be able to apply strategies in appropriate situations, where they can help to fulfill learning goals. The general use of a strategy, and especially transfer to other situations, is only possible when learners have understood when and why to use a procedure. This "when and why" is also called conditional knowledge.

O'Sullivan and Pressley (1984), conducted a study with 5^{th} and 6^{th} graders, in which children were given two different pair-associative memory tasks. The practice task was to learn products and the cities in which they were manufactured; the transfer task was to learn Latin word definitions. For the first task, the children were instructed with a mnemonic imagery procedure. Transfer of the strategy to the second task was more likely in those groups of children who were also given conditional information about when and where to use that strategy. Pressley, Borkowski and O'Sullivan (1984) state that there are at least three different ways in which metacognitive knowledge about

strategies can be acquired: First, it can be provided by an external agent, as in the study just described. Second, it is possible that this knowledge evolves through autonomous abstraction with no external help, and that the information "crystallizes" as learners work with strategies. However, both children and adults have great difficulties when expected to do this by themselves, and even if they generate such knowledge there is only little chance that they will use it (Pressley, Ross, Levin, & Ghatala, 1984). The third way is to teach children metacognitive acquisition procedures. That is, to teach them to monitor their use of strategies by gauging how much benefit they gained from using a given strategy compared to other strategies, noting which strategy was more successful, and using this information to guide further strategy use.

There are various difficulties in applying strategies. One major problem is called the *production deficit*. This means that the strategy is not applied (or transferred) without external help, although in cases when it is applied it leads to learning improvements (like in the O'Sullivan and Pressley study, 1984). Another issue is the *utilization deficiency*. This occurs when spontaneous strategy use does not result in performance gains or even leads to a decline in performance.

An explanation for the utilization deficiency is that strategies are only useful once they can be applied routinely with little mental effort and have become at least partly automatic. Conscious strategy usage can consume most or all of the learner's attentional capacity; leaving little or no room for other cognitive demands (execution of the actual task the strategy is supposed to support). With practice, the amount of attention required to execute individual strategies or sequences of strategies decreases; this frees attention for other activities. The development of automaticity is therefore an important goal in strategy education. Many researchers have contributed to the understanding of how strategies and other processes become automatized. For instance, Flavel & Wohlwill (1969) distinguished between the early fragile state of a strategy and its later robust quality. The goal, of course, is to develop a robust, trans-situational form of strategic competence, allowing successful transfer. Schneider & Büttner (1995) assume that there is a transitional stage in the acquisition of strategy knowledge. During this stage, the execution of a strategy consumes so much mental energy (corresponding to working memory, or short-term memory capacity) that there are not yet any positive effects on performance. A theory dealing with the mechanisms of the development of automaticity is that of Shiffrin and Schneider (1977). The authors distinguish between two kinds of processes: controlled and automatic processes.

Processes that are automatic are fast, require no attention and little or no working memory capacity, and take place without conscious reflection. They are difficult to modify once they have become automatic. In contrast, controlled processes require a great deal of activation and attention and take up a large amount, if not all, of the processing capacity. However, those processes have the advantage of being under control, meaning that they can be modified and adapted according to the needs of the situation. Logan (1988) developed a theory of how automation occurs. When a new stimulus is encountered, memory traces or nodes are generated. Repeated use or practice leads to accumulation of knowledge about the stimulus and strong memory traces, allowing the stimulus to be retrieved very quickly. The process has become automatic when it only takes one processing step to retrieve the memory content. Many researchers agree that formerly controlled processes become automatized by repeated practice and then consume less cognitive capacity (for instance Brown, 1984). Baumert and Köller (1996) even defined strategies as goal-directed activities/behavior that are initially consciously applied and over time become automatic, but remain potentially conscious.

The question of how transfer occurs is central to research on learning strategies. In the following, some contributions to this line of research will be discussed. Transfer of strategies to new domains is often dependent upon certain preconditions and is very difficult to achieve because learning is context-bound and situation-specific. All age and ability groups often fail to maintain and generalize strategies. Borkowski and Turner (1990) assume that strategy use is restricted by the specifics of the range of application and the limitations of the available prior (content) knowledge. But it is also assumed that strategies possess elements that are transferable over time, tasks and settings. Sometimes it only takes small cues from the environment for a strategy to be successfully transferred. For instance, Gick and Holyoak (1980) showed that university students who were told that one of the problem-solving stories they were presented with in a study contained a hint that would help them solve another problem, were able to transfer the solution they learned for a “general” problem to another domain. Pressley, Ross et al. (1984) showed that 10- to 13-year-olds were more likely to use strategy utility information gained through strategy practice when given a prompt to do so. Salomon and Globerson (1987) performed a task analysis of different types of transfer and described two different mechanisms: automatic and mindful transfer of strategies. Automatic transfer is believed to be “data-driven” and elicited automatically by

environmental cues. It is assumed that this applies to strategies that have been used and practiced up to the point of automaticity, that it needs only little short-term capacity and often occurs completely unconsciously. In contrast, mindful transfer is fully strategic: the mindful information processor is fully aware of the goal and deliberately seeks a way to accomplish it efficiently. He may consciously inhibit task-inappropriate or task-inefficient strategies and search for more efficient strategies by examining the task, identifying subtle features that may provide cues and thinking of alternatives, thereby relying on metacognitive knowledge to guide selection and deployment. Mindful transfer requires great cognitive effort. This viewpoint is consistent with the *Good Information Processor* model and contains many of the characteristics of the *Good Strategy User* model introduced in 1997 (Schneider & Pressley, 1997). These two models will be presented in the next section, beginning with an outline of the *Good Information Processor* model, of which the *Good Strategy User* is a specific instantiation.

1.2.5. The Good Strategy User

Pressley et al. (1989) used the term *Good Information Processing* to describe effective and self-regulated thinking. The major structural and process components are declarative and procedural knowledge and metacognitive control strategies. The Good Information Processor is, of course, a hypothetical superperson: His cognitive system is characterized by a large knowledge base and rapid access to that knowledge; he possesses a large number of learning strategies and knows how, when and why to use these. His thoughts and actions are reflective and systematic. A few of his motivational and personal characteristics are that he is an intrinsically motivated, task-oriented person with mastery goals, believes in the importance of effort, does not fear failure and views tests as learning opportunities. In short, the good information processor is a person characterized by a good or almost perfect set of cognitive, motivational, personal and situational components that allow him to very effectively process any kind of information (Borkowski & Thorpe, 1994). It is assumed that such motivational characteristics and personal traits are developed by early home experiences and in a supportive environment (e.g. parents, school and society).

Schneider and Pressley (1997) focused on effective, mature strategic functioning when they introduced the model of the *Good Strategy User*. The *Good Strategy User* possesses a variety of memory strategies and conditional knowledge about when and

why to use them. He had many opportunities to use and practice the strategies and many of them have become automatic and habitual. In any particular situation, he is able to select an appropriate strategy to help reach his current goal. But he is also able to monitor his own use of strategies, to determine whether they are correctly applied and helping him to progress towards his learning goal and, if necessary, to switch tactics. In the reading domain, the *Good Strategy User* would be a person who is generally attentive to the material he reads and intends to use strategies in order to remember the text. A few of the reading strategies he possesses would be the ability to re- or paraphrase material, reread text, search for effect-and-cause relationships, summarize, self-question, or form mnemonics for hard material. He would also note when problems of understanding occur and would then monitor his own text learning and strategy use.

It is assumed that much good strategy use occurs automatically and thus beyond conscious control. Learning a new strategy, on the other hand, is done consciously. The problem is that conscious strategy acquisition – e.g., learning a reading strategy - can consume most or all of a learner's attentional capacity, leaving little or no room for other cognitive demands, like elaborating on the text. This can even produce declines in performance as compared to executing a task without using a strategy (Clark, 1990). With repeated use of the strategies (practice), however, the attentional demands decrease, freeing attention for other activities. Yet it takes quite a while before the processes become automatic and seem to be executed effortlessly (Schneider & Shiffrin, 1977).

"In summary, good memory strategy use is complicated, because it involves the coordinated development and use of strategies, metacognition, and the nonstrategic knowledge base, all operating in the confines of limited capacity" (Schneider & Pressley, 1997, p. 237).

1.2.6. Motivation

Up to this point, I have been concerned with characteristics and activities of the reader that are central to the effective processing of information. Yet there are other aspects that must not be overlooked when trying to explain how effective reading can be taught and accomplished, namely, motivational issues. More specifically, intrinsic motivation and interest are conditions for the initiation and maintenance of engagement in learning and the cognitive processing of information.

Motivation is an internal process that activates, guides and maintains behavior over time. A person's desires and needs influence the direction and intensity of behavior. Motivation may come from extrinsic or intrinsic sources (or both). An extrinsic incentive is a reward that is external to the activity, such as recognition, a good grade or money. In contrast, intrinsic incentives are aspects of an activity that people enjoy and find motivating. Intrinsically motivated learning is stimulated "from within"; exploration, curiosity and interest-based learning are examples of intrinsic motivation. An impressive example of behavior motivated by the enjoyment of the activity itself is when flow is experienced: a person is then completely absorbed by and highly concentrated on a specific activity, so that he forgets almost everything around him (Csikszentmihalyi, 1985). Expectancy-value theory views motivation as a product of the perceived probability of success and the value of that success and predicts that motivation will be at its maximum for moderate levels of success.

One major concept in academic and occupational contexts is achievement motivation, defined as the desire to acquire knowledge or skills (Rheinberg, 1986) or as the generalized tendency to strive for success and choose goal-oriented success activities (McClelland & Atkinson, 1948). But achievement motivation is not the only important factor in explaining children's engagement in learning activities; interest also plays a role. In contrast to motivation, which explains the engagement in a momentary learning activity, interest focuses more on stable person-object relations. Interest is defined as a dispositional characteristic of a person or a durable tendency for engagement that is directed towards a specific object or activity (Todt, 1978). Other researchers view interest as a general and durable as well as a specific and temporary characteristic (Krapp, 1992; Schiefele, 1996). Krapp (1992) distinguished two kinds of interest: situational and individual interest. Situational interest is a single, situation-specific motivational state that is aroused by the specific incentive structure of the learning situation. Individual interest, by contrast, is a habitual tendency or dispositional characteristic of a person reflecting a relatively stable preference for a learning object or activity. Both kinds of interest are closely interrelated and are of influence during an interest-based learning activity. When there is little individual interest, the incentive structure of the learning situation is of great importance in directing the learner's attention towards the object of learning and prompting him to engage in the activity for some time. In contrast, when individual interest is high, the structure of the learning situation plays little or no role with regard to engagement.

Over time, situational interest can develop into individual interest; depending on positive learning experiences and the corresponding attributional processes.

(Individual) interest is an especially interesting construct because of its intrinsic character, or self-intentionality. That means that interest-based activities occur in situations when the learner can dispose freely over his time. In this situation, the learner gets the feeling that he is able to do what he wants; the activity is self-determined. Objects and activities of interest are integrated in central regions of the self, and there is a highly positive correlation between identification with objects of interest and the experience of subjective meaningfulness (Krapp, 1999). The development of stable academic interest is an important goal of education (Krapp, 1998; Schiefele, 1986). Instruction should be designed to arouse interest in students and to maintain this interest. Principally, everything that happens in the classroom impacts on students' motivation and interest. Therefore, the question is: how should instruction best be planned to promote students' interest and motivation? A theory that may help to derive some instructional principles fostering motivation is the self-determination theory of Deci and Ryan (1985). Two kinds of motivation are distinguished: (1) motivation that results in activities that are carried out voluntarily, or are self-determined, and (2) motivation resulting in activities that are carried out because the person was forced to, which are not self-determined. In the latter case, the activity is carried out instrumentally to receive a reward or to avoid punishment; in the former case, the reward lies in the activity itself. Intrinsically motivated behavior occurs when a person identifies with the object or activity at hand and when his actions are self-determined, when he has the impression of doing what he wants to do. The basis for the theory are the inherent psychological needs and desires of the organism: the desire for competence, the desire for self-determination (autonomy) and the desire for social relationships. These desires help to explain in which contexts, and under which conditions, intrinsic motivation or interest is developed. First, a person is intrinsically motivated when he or she experiences competence. This is dependent upon feedback in learning situations that leads to judgments of ability and competence. Instructional conditions which allow students to experience competence include a productive working atmosphere with good guidance and few discipline problems, pacing which suits the individual students so that something new is taught every lesson without demanding too much of the students, and activities that encourage students to focus on content matters. Second, people have the desire to experience themselves as acting independently and to determine their goals

and activities themselves. Self-determination means that engagement in an activity is not enforced by external factors. Therefore, extrinsic incentives like money, prizes, avoiding punishment, and deadlines negatively influence the feeling of self-determination, whereas opportunities for choice and acceptance of one's viewpoints and feelings have a positive effect. Assessments based on an individual frame of reference and forms of instruction like groupwork, discussions and individualized instruction, which allow students to develop their own ways of reaching a solution and to state their own opinions, not only help students to feel self-effective and develop a positive self-concept, but also help them to construct knowledge actively rather than acquiring it passively. Last but not least, people seek social security and satisfying social interactions. They identify with other people and their activities, goals and values and are thus motivated to engage in new activities and to examine new areas of knowledge. In this way, activities that are not intrinsically motivated can nonetheless be experienced as personally meaningful because of their value for relations to important people like parents, teachers or friends. The authors also assume a developmental process of internalization and integration of extrinsically motivated behavior: external regulative processes are thought to become integrated into the self and to form the foundation for self-determined, extrinsically motivated behavior (Deci, Vallerand, Pelletier, & Ryan, 1991). This assumption is in accord with the observation that many learning activities are both extrinsically and intrinsically motivated: students learn because of their interest in the topic and enjoyment of the activity itself as well as because they are motivated to succeed in examinations and get good grades (Wild, 2000). Groupwork and discussions are very good examples of instructional methods that foster the development of autonomy as well as the ability to cooperate with and appreciate others, to participate and assume responsibility.

There are many ways to develop student interest and to foster motivation or situational interest in school. For instance, situational interest in textual material can be aroused by a number of textual characteristics: by providing surprising or curious information, by incorporating narrative text passages that allow students to identify with the figure or by integrating personal experiences. In the domain of mathematics and science, the Jasper Series (Vanderbilt, 1992) is an example of the *anchored instruction* teaching concept which incorporates many elements that foster situational interest and motivation. The anchor for instruction is an adventure story with interesting characters who experience an ecologically valid problem. The learners' task is to help the main

characters solve the problem; groupwork is encouraged and different solutions or ways of solving the problem are possible. However, exciting and surprising elements should not be allowed to dominate instruction; when they direct students' attention and processing resources towards less important information, they may even result in less understanding and learning.

1.2.7. School – a social setting for learning

Where and how does learning take place? How are knowledge and skills acquired? Much learning occurs in the institutional educational setting of the school. Children spend an enormous amount of time there – at least half the day during the elementary school years and even more at secondary school (depending on the school program). School is a social setting characterized by complex interactions between teachers and students, as well as among students. Students interact with one another during the lessons as far as the teaching method allows and, of course, during breaks.

Teachers are central agents for students' learning processes; they structure and guide various kinds of learning activities during regular school lessons. The interaction that goes on during lessons is mainly determined by the use of various types of teaching strategies aimed at directing students' learning processes. A few of these are direct instruction, cooperative learning methods and humanistic approaches.

These different types of teaching strategies will be introduced in the following, with a special focus on cooperative learning methods, which are of particular relevance to this thesis. Additionally, and also because of their critical importance for this thesis, learning theories focusing on the role of context and social interaction for learning (anchored instruction and cognitive apprenticeship) will be described. The remainder of the section deals with how learning can be fostered by interaction. A very influential theory in this respect is that of Lev Vygotsky (see Wertsch, 1985). This theory deals with the development of higher mental functions and is especially important for explaining the development of metacognitive knowledge and skills.

1.2.7.1. Instructional methods

Formal education takes place during school lessons. It is characterized by complex interactions between teachers and students – between teaching and learning. The means by which teaching goals can be realized are teaching strategies (analogous to learning strategies), which characterize the individual activities of teachers. But because the teaching and learning processes that occur in the classroom are

bidirectional, it may be more appropriate to speak of instructional methods. Instructional methods differ according to how learning is to be achieved. They can be arranged on a continuum with the following poles: the reception of knowledge by learners, with teachers determining the lesson, at the one extreme, and methods based on the view that learners develop by interacting with their environment and that they themselves can best decide what is useful for them to know and how to achieve it, at the other extreme.

Which instructional method is chosen by the teacher depends on many factors, including the teacher's theoretical knowledge about determinants and laws of learning and teaching, practical experiences of teaching activities and subjective reflections about these activities, routines, the concrete situation in class (including the social climate and students' learning abilities, motivation, behavior, etc.), the learning material and the teacher's personality. Effective teaching depends on the availability of various types of teaching methods, their competent selection and their situation- and goal-adequate application (Lompscher, 2001). A few important instructional approaches will be introduced at this point; emphasizing their strengths and weaknesses. Among them are direct instruction, discussions, cooperative learning methods, humanistic approaches and methods based on constructivist learning theories.

Direct instruction is characterized by teachers transmitting information directly to students, thereby structuring class time to achieve a clearly defined set of objectives as efficiently as possible. The essential events of instruction correspond with the key events in a student's learning process. The teacher's job is to structure the learning situation, select appropriate materials and present them in a well-organized lesson (see also Slavin, 1994). Direct instruction is particularly appropriate when teaching a well-defined set of information (factual knowledge). It can be applied to a whole class and is an economic way of transmitting information to many people (and is often also used in college and university courses). On the part of learners, direct instruction can foster highly intensive analysis of the learning material and lead to high achievement. Over time, however, it is possible that motivation and concentration diminish because learners cannot actively apply their newly acquired knowledge. It is also very difficult, if not impossible, for teachers to adjust to individual student's abilities and needs using this teaching strategy. Therefore, it is recommended that direct instruction should be combined with individualized learning, for instance, with cooperative learning methods (Lompscher, 2001).

Discussions can be useful when there is no simple answer to questions, or when learning objectives with a single correct answer involve difficult concepts or have an affective component. For discussions to be effective, it is important that all students possess factual knowledge so they can participate. Therefore, this method can be combined very easily and fruitfully with direct instruction. There can be whole-class discussions as well as small group discussions. They can be led either by the teacher who introduces the topic or by a student. However, it is important that the leader of the discussion makes sure that everyone participates and (especially when the discussion is led by a student) that the group stays on topic. Applying this teaching method requires the teacher to know how well the students are achieving and to be able to judge their knowledge and skills accurately.

Cooperative learning methods are instructional methods in which students work together in groups, have the opportunity to learn together and to discuss the information provided by the teacher, and to practice their skills. These approaches encourage students to discuss, debate, disagree and teach one another. They emphasize the development of thinking skills and increase higher-order learning (see Slavin, 1991). The groups formed stay together for several weeks or months; they consist of at least two members (see also peer tutoring), but most methods involve groups of about four students. A major characteristic of cooperative learning methods is the reward or goal structure: the group members can only attain their own personal goals if the group is successful. This fosters helping behavior among the teammates and encourages them to exert maximum efforts. In contrast, the competitive grading and informal reward system of traditional classroom practices is the source of peer norms that oppose academic efforts (Slavin, 1996). Some examples of cooperative learning methods are Student Team Learning (see Slavin, 1991), Jigsaw, Learning Together (Johnson & Johnson), Group Investigation (Sharan & Sharan), Reciprocal Teaching (Palincsar & Brown, 1984) and Cooperative Scripting.

Research findings indicate that students learn substantially more with cooperative learning methods than with traditional instructional approaches if two conditions are met: (1) if some reward or recognition is provided and (2) if group success depends on the individual learning of each group member and not on the single group product (see Slavin, 1994). The students who gained most were the ones who gave and received elaborated explanations. Other positive outcomes were higher self-esteem (especially for students who had experienced failures in the past), a more

positive attitude toward school, overall academic improvement, more supportive, friendly and prosocial behavior, acceptance of mainstreamed children with special education needs, etc. Students who cooperate also learn to like one another; it was observed that after working together in mixed groups during school lessons, children also started to spend their spare time (lesson breaks, cafeteria, birthday parties, etc.) with children from other social groups (boys vs. girls; children from families with different socio-economic status or ethnic backgrounds, etc.).

A special form of cooperative learning is *peer tutoring*. The main characteristic of this instructional method is the distribution of directive functions between the teacher and the learners. The investigative setting for studying group processes is typically helping behavior (receiving and giving help) when students work together in groups of four to solve mathematical or computer problems (Webb, 1985). Most of the training units studied were of two weeks duration or less. For help to be effective in this case, it must involve an explanation (rather than just an answer to the math problem), it must be provided in response to a student's needs, and it must be understandable. It has generally been found that providing another student with explanations is positively correlated to achievement. However, no correlations to achievement have been found for low-level help. Most help, of course, is provided by high-attaining students. There is always the fear, mainly communicated by the parents, that high attainers do not profit from working together with low-attaining students. Bennett and Cass (1988) conducted a study designed to contrast the effects of three types of groups – ability groups of high, average and low attainers as well as mixed-ability groups with different proportions of high, average and low attainers – on co-operative decision-making on settlement patterns. The best groups were, of course, those of the high attainers. However, all high-attaining children performed well – irrespective of the type of group they were in. The achievement of low attainers in the mixed ability groups differed considerably between groups: results were far better when two low attainers worked together with only one high-attaining student then when it was the other way around. In the latter structure, low attainers seemed to be largely ignored by the other students. In groups of low-attaining students, the level of instructional talk was very low and only very few explanations were given. These groups apparently do not have the relevant skills and knowledge to give effective explanations; therefore, learning was very limited.

Humanistic approaches focus on self-regulated learning, affective goals and authentic assessments. The key feature of humanistic approaches is that students

choose what, when and how to study. This is assumed to enhance motivation. Some of the affective goals targeted are consideration, cooperativeness, mutual respect and honesty. Grades and standardized tests are avoided; instead, written evaluations or no assessments at all are recommended. The demands on teachers using this method are very high; they have to create conditions for their students to engage in self-regulative and self-determined study activities by providing materials and serving as assistant und counselor. There is little evidence that these teaching strategies are more effective than the methods traditionally used in schools. Humanistic approaches succeed only partially in developing systematic knowledge and skills among all students. However, these methods do foster the development of creativity and engagement (Lompscher, 2001).

Constructivist learning theories assume that students play a far more active role in their own learning. They are based on the idea that teachers cannot simply transmit knowledge to students, but that students must construct knowledge in their own minds. They must individually discover and transform complex information. These approaches draw upon the work of Piaget and Vygotsky, both of whom emphasized that cognitive change only takes place when previous conceptions are reviewed in the light of new information and revised if necessary. Both authors also emphasized the social nature of learning and suggested that mixed-ability learning groups be used to promote conceptual change. Constructivist approaches emphasize top-down rather than bottom-up processing.

Students should begin with complex, complete and authentic problems and then work out or discover the basic skills required to solve these problems. This is exactly what happens in the *anchored instruction* approach, where new information is learned in the context of meaningful activities. These are anchors with inherent ecological validity. In this way, students are more likely to perceive the knowledge acquired as a tool, rather than as a set of facts, and can find out about the conditions in which the new concepts or facts will come in useful. This approach is likely to overcome inert knowledge problems; that is, the problem that students can recall knowledge when explicitly asked for it, but do not use this knowledge spontaneously in problem solving situations (Vanderbilt, 1990). Two examples of programs using anchored instruction are the Jasper Series (Vanderbilt, 1992) and the Young Sherlock Project.

Vygotsky claimed that higher mental functions, such as the ability to direct memory and attention in a purposeful way and to think in symbols, are externally

mediated behaviors that become internalized in learners' minds as tools (see chapter 2.6.2). Teachers serve as cultural agents who guide instruction. *Scaffolding*, a form of assisted learning, involves giving the students more structure at the beginning and gradually turning responsibility over to them as time goes on. Once they have been acquired, internal mediators permit students to engage in more self-mediated learning. This procedure, which underlines the role of the social context, can often be observed in naturalistic learning situations. For instance, Rogoff and Gardner (1984) observed mothers who taught their children about organizing kitchen items, and what would help them remember the location of those items. The mothers began by framing the task or describing it to their children in interesting and understandable terms. Then they modeled strategic approaches, giving both a rationale and specific examples of organization. They attempted to transfer responsibility for solving the memory problem by engaging the child in the organization of kitchen items. Finally, the mothers monitored the children's skill levels and encouraged self-monitoring by asking them about the accuracy of classifications, comparing the classifications with the actual organization of kitchen items at home, and questioning and/or informing them about the usefulness of the strategy.

Cognitive apprenticeship is a form of instruction that has been derived from traditional apprenticeships and transferred to the cognitive domain (Collins, Brown, & Newman, 1989). Traditional apprenticeships were the most common means of learning before schools appeared, from the domains of agriculture and crafts to those of medicine and law. Learning was accomplished by a combination of observation, coaching and practice. Three major phases can be distinguished: (1) observation, (2) attempts to execute the process with guidance and help from the master (scaffolding) and (3) the master reducing his participation once the learner has a grasp of the target skills (fading) until the student is able to do everything by himself. This occurs in a social context; the learner has continual access to several models of experts-in-use. The richness and variety of instruction and experience gained through repeated practice leads to an understanding of multiple ways of carrying out the task.

Cognitive apprenticeship aims primarily at teaching the processes that experts use to handle complex tasks. It focuses on cognitive and metacognitive skills and processes and encourages both a deeper understanding of the meaning of concepts and facts and a rich web of memorable associations between them and problem-solving contexts. It requires the externalization of processes that are usually carried out

internally and extended techniques to encourage the development of self-correction and self-monitoring skills. Two very successful programs using the cognitive apprenticeship approach are Reciprocal Teaching (Palincsar & Brown, 1984) in the domain of reading and Procedural Facilitation of Writing (Scardamalia & Bereiter, 1986). The method can also be observed in naturalistic learning situations, as already described in the study by Rogoff and Gardner (1984), who observed mothers and their children in an everyday learning situation (organizing kitchen items).

1.2.7.2. Vygotsky's theory of the development of higher mental functions

As mentioned above, Vygotsky's theory of the development of higher mental functions is a very influential theory dealing with the development of complex mental skills in a social context as a result of interaction between people. Lev Vygotsky stated that "… this transition from a social influence external to the individual to a social influence internal to the individual … is the center of our research." This view is very important to the development of metacognitive processes. For this thesis, the processes by which reading strategies and skills are acquired in social interaction are of primary importance; therefore this theory will be described in detail at this point.

Vygotsky claimed that mental processes can only be understood by considering how and when they occur in human development. He was interested in the processes by which higher mental functions are established. Biological principles cannot explain psychological phenomena beyond a certain degree, so he focused on social factors. The central argument in his theory is that higher mental processes in the individual have their origin in social processes. This is formulated in his *general genetic law of cultural development*: "Any function in the child's cultural development appears twice, or on two planes. First it appears on the social plane, and then on the psychological plane. First it appears between people as an interpsychological category, and then within the child as an intrapsychological category. […] that internalization transforms the process itself and changes its structure and functions. Social relations or relations among people genetically underlie all higher functions and their relationships."

Vygotsky distinguished between elementary and higher mental functions. Higher mental processes are assumed to represent a qualitatively new level of psychological functioning. Besides the distinction between elementary and higher mental processes, there is also a distinction between levels of development within higher mental functions. Development was seen as representing fundamental "revolutionary" shifts rather than steady quantitative increments. The major transition

points in development are associated with the appearance of some new form of mediation in the form of tools or signs. Vygotsky used four criteria to distinguish between elementary and higher mental functions. (1) Most importantly, higher mental functions have a social origin and are of a social nature. (2) There is a shift of control from the environment to the individual; i.e., voluntary regulation emerges. Whereas elementary functions are totally and directly determined by stimulation from the environment, the immediate cause of higher mental functions are artificial stimuli that are created and used by the individual. (3) Mental processes are consciously realized. This is called *intellectualization*. (4) Psychological tools, or *signs*, are used to control one's own and other's behavior.

The process whereby certain aspects of patterns of activity that had been performed on an external plane (social phenomena) come to be executed on an internal plane (psychological phenomena) is called *internalization*. The product of this process is an internal plane of consciousness that takes on a "quasi-social" nature because of its social origins. Internalization is the mastery of external signs and the mastery of the rules in accordance with which these external signs must be used.

There is a special emphasis on the role that signs play in the process of mediating higher mental processes. Signs are defined as a means of internal activity that are directed inwardly, toward the mastery of humans themselves. Their nature is meaningful and communicative. Examples of human signs that Vygotsky mentioned are language, various systems of counting, mnemonic techniques, algebraic symbol systems, works of art, writings, schemes, diagrams, maps and mechanical drawings. The introduction of psychological tools into a mental function prompts a fundamental transformation of that function. Signs are of a social nature, not only in the sense that there are products of socio-cultural evolution, but also in the sense that they are initially always used for social purposes like influencing others. Only later do they become a means of influencing oneself.

Vygotsky was especially interested in the ontogenesis of speech. He claimed that children induce or infer word meanings and the structure underlying adult speech by interacting with them. That means that children master the existing speech of the adults. Of special interest was inner speech as an example of a higher mental function. According to Vygotsky, inner speech enables humans to plan and regulate their activity and derives from previous participation in verbal social interaction. The transition from external (social) speech is made via egocentric speech.

But when and how is it that the transitions leading to the development of higher mental functions occur? To deal with this issue, the *zone of proximal development* was introduced. This is the dynamic region of sensitivity in which the transition from interpsychological to intrapsychological functioning can be made. The zone of proximal development is defined as the distance between a child's "actual developmental level as determined by independent problem solving" and the higher level of "potential development as determined through problem solving under adult guidance or in collaboration with more capable peers." Vygotsky argued that instruction should take place in the child's zone of proximal development. Then, "… it calls to life, awakens and puts in motion an entire series of internal processes of development in the child. These processes are at the time possible only in the sphere of interaction with those surrounding the child and in collaboration with companions, but in the internal course of development they eventually become the internal property of the child."

A common pattern in children's mastery of the situation definition of a task is the following: First, children participate in the execution of the goal-directed task on the interpsychological plane. Subsequently, they recognize and master the strategic significance of their behavior. Thus, it seems that first they perform the task and only then do they understand it.

Vygotsky tried to measure the potential level of development separately because it may vary independently of the actual level of development. This has also been shown by Brown & Ferrara (1985) in a study with 3rd and 5th graders. The children's task was to identify and continue sequential patterns of letters. For one third of the children, learning speed was not predicted by IQ scores. Similarly, transfer could not be explained by IQ alone. Wertsch and Hickmann (1987) listed the following factors that encourage the transition from interpsychological to intrapsychological functioning: (1) children's cognitive readiness, (2) adults' willingness to transfer strategic responsibility to the children, (3) adults' use of "reflective assessments" to inform children of the significance of their behavior, (4) explicitness of the adult's directives and (5) the possibility of the dialogic structure of the interpsychological functioning being mastered on the intrapsychological plane through the differentiation of language functions.

1.2.8. Developmental aspects

Childhood is characterized by enormous growth and continuous changes – in short: development. This is the case in almost every respect; there is rapid physical growth as well as acquisition of knowledge and skills in many (if not all) domains: in sensomotoric, linguistic, cognitive and socio-emotional domains.

Some developmental aspects have been explained and/or mentioned above: a theory of the (at least partly) sequential emergence of various components of metamemorial knowledge (Borkowski et al., 1988; model of mature metamemory, see section 2.3), an information processing theory which can be used to explain the transition of external stimuli and events into long-term memory content (both declarative and procedural), a model of automatization of processes (Schneider & Shiffrin, 1977) and Vygotsky's theory of the development of higher mental functions.

This section deals with developmental changes that occur in areas which are also relevant for text comprehension: development of various aspects of memory (working memory and long-term memory), metamemory as well as language and reading skills. The goal is not to provide a complete picture, but to illustrate the complexity of the changes that enable students to acquire the skills necessary to become good readers, or hinder them from doing so.

1.2.8.1. Memory

From the first day of their life, children acquire knowledge about objects and people, phenomena that occur in nature, animals and events. This process of knowledge acquisition proceeds until death (unless people suffer from dementia). The knowledge is stored in long-term memory in the form of concepts, episodes, schemata and images.

During childhood, there are also developmental increases in the capacity of short-term memory (see Dempster, 1985), as measured by various memory span tests. For instance, digit span increases from about two digits at age 2 to the "magic" seven digits in adolescence (Miller, 1956). What causes the observed increases in different memory span measures is not clear; there are structural and functional explanations. Structurally, Miller (1956) proposed that there is an increase in the number of neurologically determined "slots" with development. One functional explanation is that the use of strategies determines performance on memory tests. There is evidence to support this: for instance, increased use of rehearsal strategies is observed at the end of elementary school years. With the acquisition of content knowledge, the larger chunks that can be created by the learner explain differences in short-term

capacity. A prominent example is the study by Chi (1978), in which children outperformed adults in memorizing chessboard positions – the children, who were experts in the domain of chess, viewed the whole chess board as the unit of learning, whereas the chunks of the less knowledgeable, novice adults were much smaller. Whether or not the material needs to be recalled in the order it was presented also plays a role: children's performance on memory span test improves (up to about age 6) when serial order is not considered at the time of recall (Schneider & Weinert, 1995). Another functional explanation deals with increases in processing speed: high correlations have been observed ($r = -.74$ in Case, 1978) between reaction time and memory span for children aged 3 to 6. It is likely that the rate of item identification and speech rate or both are important determinants of memory span; both increase during childhood. Case (1985), a neo-Piagetian theorist, assumed that the total processing space (which consists of operating space and storage space) does not change during development and that the developmental increases in functional capacity are due to more efficient processing of stimuli. According to Case, the amount of space required to process stimuli decreases with age, given the more efficient operation of executive actions. This frees up space for the storage of material and accounts for the increases observed in memory span. Another potential explanation for the increase in short-term memory capacity is that it is related to development of cognitive inhibition and resistance of interference (Dempster, 1985). Task-irrelevant cognitions and interference are two phenomena to which children are very susceptible, but which they learn to cope with over the course of development. Increases in the inhibition of task-irrelevant cognitions free up proportionally more space in the working memory; so that more capacity becomes available for children to process task-relevant information as they get older (Bjorklund & Harnishfeger, 1990).

1.2.8.2. Metamemory

How does factual knowledge about memory (declarative metamemory) develop during childhood? In *early childhood* (ages 3 to 5), children begin to understand mental verbs (they realize that "remembering" and "forgetting" require having "known" something at some point) and they begin to recognize the person and task variables that are relevant for memory performance. For instance, only half of the children of this age know that memory improves with age (Kreutzer et al., 1975). When asked to report useful strategies in a natural memory task, for instance the "skates question" ("What can

you do to remember to take your skates to kindergarten the next day?"), they reported only external strategies.

During *preschool and early elementary years* (kindergarten and grades 1 and 2), children develop a somewhat better understanding of person and task variables, but still do not know that taxonomically organized items are easier to learn than unrelated items. They also tend to overestimate their performance; for instance, they are convinced that they never forget anything. What is known about children's procedural knowledge about metamemory? Examples of measures assumed to tap some aspects of procedural metamemory are performance predictions, feeling of knowing, knowledge of recall readiness and of items requiring additional study. Preschool children and children in early elementary grades consistently overestimate their performance on serial memory span tasks and list learning, and even predict performance for items that were not remembered before. In contrast, they underestimate their abilities in tasks requiring recall with categorizable lists, because they lack awareness of the effects of task characteristics. Moreover, the knowledge children have about task difficulty does not result in appropriate self-regulation: 1st and 2nd graders know which items they would probably not answer correctly on a test, but select items that they have previously recalled correctly for additional study. 3rd graders begin to select items they have not previously recalled correctly for additional study, but still spend about as much time on hard as on easy items. Even at this early age, children are able to give accurate feeling of knowing judgments, but only when problems are simply structured and involve highly meaningful material.

One theory concerned with the development of metacognition in early childhood is Wellman's theory of mind (1985). He proposed five partially overlapping classes that are assumed to develop during the preschool years. (1) By the age of 2 or 3 years, children develop knowledge about the existence of an inner world and have a rudimentary understanding of "thinking" or "remembering." (2) Children learn to distinguish between mental processes somewhat later; approximately between age 3 and age 5. Their mental world can then be differentiated into processes like remembering, knowing or guessing. (3) Once this has been accomplished, there is a growing understanding of the similarities between certain mental activities; knowledge becomes integrated. During the late preschool years, (4) knowledge about variables (i.e., tasks, strategies and persons) is acquired and (5) the first cognitive monitoring activities (awareness of one's own mental condition relative to the task demands) appear.

By the *late elementary years* (in grades 5 and 6), children are usually able to differentiate between most person and task variables that are relevant to memory tasks. They have realized that memory skills vary between people and also between situations, and know about the organizational structure of items (that taxonomically organized items are easier to learn). In natural memory situations (e.g., "What can you do when you have lost your jacket?"), they suggest searching systematically and elaborately. Because they are now able to differentiate between effort and ability, they no longer overestimate their performance as much. By the end of elementary school, procedural metamemory has improved greatly.

Although there are clear improvements in monitoring during the elementary years, these skills develop further during *adolescence* and *adulthood*. Factual knowledge about memory also continues to develop – with respect to mental states, for instance: that memory involves conceptual processing, that storage and retrieval are critical processes, and that there is conscious and unconscious memory. Similarly, knowledge about the interaction of memory variables continues to develop beyond the school years. A study on the combined effects of a task variable (number of items to study) and a person variable (amount of effort) by Wellman, Collins and Glieberman (1981) with children of ages 5, 8, and 10 and adults aged 19 showed that, although all age groups were aware of the combined influence of the two variables, they differed in their respective weightings. Younger children placed more emphasis on effort than on task difficulties; this is consistent with their tendency to attribute success to effort (Nicholls, 1978).

1.2.8.3. Reading

One of the most important developmental tasks is the acquisition of language (learning words as well as the rules of word and sentence construction). Reading and writing skills are learned and practiced in the context of formal education. They are the basis for successful learning and performance in almost all school subjects.

Reading is a complex skill which takes a long time to develop, and children do not start learning to read only when they enter school. Already in early childhood, even before children are able to say their first words, they learn a great deal about their native language. Preliminary competences developed during the preschool years are of specific relevance for the acquisition of literary language. Among these are auditive, visual, motoric and linguistic abilities and skills. Recent research has concentrated on phonological information processing; that is, the use of information about phonemic

structure in the processing of spoken and written language. The reason for this is that (in Western cultures) language is, more or less, isomorphically represented by letters. There is widespread agreement that deficits in linguistic coding are the main reason for children's reading problems. Three main components have been the subject of recent research: (1) phonological awareness; that is, identification of larger units like words and syllables and differentiation of small units (phonemes), (2) phonological recoding when accessing the semantic lexicon, and (3) phonetic recoding in working memory; that is, phonetic representation of written symbols in working memory to keep information active for as long as possible. Correlational longitudinal studies have shown that components of phonological information processing during the preschool years are better predictors of reading and spelling skills than general intelligence is (see (Schneider, 2004)). Phonological awareness, in particular, seems to be a vital skill; it was the most important predictor of the acquisition of literary language by the end of grade 2; additional but less important predictors were linguistic information processing speed and working memory (Weinert & Schneider, 1999). However, training programs have only proved to be effective for phonological awareness. First signs of phonological awareness can already be observed in children as young as 3 (Bradley & Bryant, 1985).

When children enter school at age 6 or 7, they are systematically taught decoding skills. The first step is usually to learn to associate abstract signs (i.e., the letters of the alphabet) with sounds, and to learn how the sounds are joined together to form the written word. Because children repeatedly encounter the same words, another learning process sets in implicitly and mainly unconsciously: the identification of words or certain sequences of letters (syllables) by means of their unique morphological appearance. Because all of those processes occur in parallel, decoding can be executed more quickly and with less effort. Repeated practice also results in automatization. The identification of single letters and their joining together to make a word also increasingly occur in parallel. Often, it is no longer necessary to decode all the letters of a word because, on the basis of the first syllable, the meaning of a word may already have been retrieved from long-term memory, or some or all words starting with that sequence of letters or sound may have been activated (priming). In this way, knowledge-driven top-down processes aid the more capacity-consuming bottom-up process of decoding, and reading becomes faster. The knowledge base (long-term memory) is also of importance for reading: lexical access is dependent upon vocabulary

knowledge and phonological decoding. However, the causal relationship between vocabulary and effective text processing is unclear. Lexical access is susceptible to interventions that positively influence vocabulary and phonological decoding; for instance, training in phonological awareness (Schneider, 2004).

Knowledge about text and its structure is also acquired during the school years. One aspect which has often been of interest is students' ability to rate the importance of pieces of information in a text. The results that have been obtained are dependent on the type of texts used and on the task demands (i.e., the rating procedure). For longer and more complicated texts and rating procedures, children from 7th grade on gave importance ratings comparable to those of college students, whereas the ratings of 3rd and 5th graders differed greatly. When the material to be read was shorter and less complicated, and rating procedures were less difficult, 5th graders were also able to differentiate central from less important information in the text, and when simple picture stories were used, even 4- and 6-year-old children showed sensitivity to relative importance levels (see Schneider & Pressley, 1997).

There is a bidirectional relationship between metacognition and text processing skills. Correlational studies have shown that there is a relationship between various types of metacognitive knowledge about text and text processing outcomes (comprehension and recall). The size of these correlations varies from low to moderate (correlation coefficients from .10 to .50, see Schneider & Pressley, 1997). One relationship that has often been explored is that between awareness of the relative importance of textual information and recall. Although even preschoolers tend to remember more central information better than less important information, this does not reflect metacognitively directed differential information processing. In a study with 5th, 7th and 8th graders, Brown and Smiley (1978) demonstrated that only children from grade 7 on are able to direct additional study time towards the more important information in a text; this is assumed to be an indicator of metacognitively directed differential processing of textual information.

1.2.9. What form should an ideal training program take?

In this chapter, many factors that influence how well readers comprehend texts have been described and their complex interplay has been outlined and demonstrated. I have concentrated on factors pertaining to the reader himself: various components of memory (short-term and long-term), metamemory as a special aspect of memory, and

strategies, which are the core of metacognition and a means of goal-directed behavior making information-processing more effective. The main reason for focusing on the reader (his characteristics and some of his activities) and not on textual characteristics or task demands was that, in the majority of situations where reading skills are relevant, the reader himself has no influence on the material to be read or the purpose of reading; he cannot freely choose what or when to read and why to engage in reading; somebody else usually determines both the task and the reading material. The goal should be to improve reading comprehension in a variety of situations and across a wide range of texts. In order to accomplish that goal, it is a good policy to concentrate on the reader: his abilities, skills, motivation and attitudes.

But what are the special merits of a good reader? A good example for the domain of reading is the *Good Strategy User*, a model which has already been outlined in this chapter. A good strategy user is generally attentive to the material he reads, takes a strategic approach to reading tasks (intending to use strategies and remember text), can rephrase or paraphrase text, backtrack, reread, search for cause-and-effect relationships, summarize, self-question and form mnemonics for hard material. He knows and notes when he is having problems understanding a text and when certain aspects are difficult (i.e., he monitors his own understanding of the text and learning progress) and monitors his use of strategies (whether or not the strategies are successfully implemented and help him progress towards his reading goal), switching tactics when necessary. Other characteristics of the good reader are a rich knowledge base, good working memory capacity and fluent decoding skills.

To what extent are the different abilities and skills readers possess susceptible to training? I have already explained the central functions of *working memory* for information processing and storing, and its very limited capacity. The common opinion is that there are interindividual differences in working memory capacity and that this capacity can barely be improved by interventions (see Mähler & Hasselhorn, 2000). *Decoding abilities* and *lexical access* can be influenced by interventions with positive effects on vocabulary and phonological recoding; for instance, by training in phonological awareness (Schneider, 2004). However, the causality of the relationship between a large vocabulary and effective text processing skills is unclear: both are dependent upon the other. *Fluid intelligence* can be improved to some extent by training programs; for instance, training in inductive thinking (Klauer, 1996). *Motivational processes* are associated with levels of processing, and frequency and

amount of reading. For instance, thematic interest is an important predictor of text learning. A variety of approaches have been shown to have positive effects on interest: perception of competence (through informative, positive feedback), self-determination, involvement in a social setting (groupwork) and underlining the importance of the learning content (authenticity and projects). Hope for success and fear of failure, both determining the performance motive, are generally also susceptible to intervention. A person's verbal self-concept is also important for motivation and thus indirectly influences text comprehension. Causal attributions which rely on the ability to differentiate between effort and ability are the basis for the development of the self-concept and are influenced by performance feedback. It is possible to improve children's *prior (content) knowledge*, but this would of course only have an effect for texts on exactly the same topic and would thus not be applicable in the majority of reading situations. In contrast, *metacognitive knowledge* has positive effects on comprehension and effective learning from texts independent of the specific content, and it is plausible to foster reading comprehension by teaching metacognitive knowledge. Strategies representing part of the learner's declarative metacognitive knowledge base are highly susceptible to training (Streblow, 2004). Strategies can influence text processing in all four phases of text processing: (1) selection of information, (2) construction of relations between information in working memory, (3) acquisition or transfer of information to long-term memory and (4) integration of newly acquired information into the (prior) network of knowledge. For instance, cognitive strategies help during the selection and acquisition phases, organizational strategies aid selection, construction and acquisition, and elaborative strategies are particularly useful during the integration phase.

To summarize, the factors most susceptible to training are strategy knowledge, motivational variables and decoding fluency. Additionally, it is very likely that (procedural) metacognitive knowledge can be enhanced by interventions. It is not practicable to promote *prior (content) knowledge* when the goal is to improve reading comprehension in general, because the knowledge acquired would only help with texts on that specific topic. General intelligence and working memory can barely be improved, if at all.

So what should the main focus of training programs be? Concentrating on components that are important for effective information processing would seem a very promising approach when the goal is to help children acquire knowledge and skills for

reading and learning that go beyond the actual situation and allow for transfer. To accomplish that, I would like to emphasize strategy knowledge and metacognitive knowledge. Strategies have proved to be predictive of text comprehension, they are susceptible to training interventions and their use is one of the characteristics in which good and poor readers differ. Substantial associations have also been observed between achievement and reported strategy use (Schneider & Pressley, 1997). Strategies can be seen as powerful tools: they have the potential to help during many stages of processing the information in textual material if they are mindfully applied, and if the learning process and its outcomes are monitored and adjusted. Efficient strategy use can also help to overcome some of the restrictions posed by the memory system: it allows limited processing resources to be used more efficiently, resulting in resources being freed up and becoming available for other processing.

However, before a strategy can be really successfully applied so that it produces benefits and frees resources, it needs to be trained until it becomes automatized and its execution no longer consumes most or all of the learner's cognitive resources. But automatization alone is not sufficient: for strategies to be transferred to new learning situations (for instance, from German to biology class or from the school context to reading for pleasure), readers must also possess metacognitive knowledge. As Borkowski et al. (1988) stated " ... for a newly acquired strategy to be transferred, and for strategy invention to eventually become commonplace, young children must first possess mature, low-level knowledge about a wide range of specific strategies. Then, they must come to appreciate the general importance of strategies in leading to successful performance and must believe in their own capacity to control learning outcomes. Finally, they need executive routines to carry out decisions about strategy selection and monitoring."

Metacognitive activities can be expected to improve learning outcomes, especially when learners work on tasks of medium difficulty (Weinert, 1984). In this case, the chances are good that strategic behavior is able to influence learning activities and the resulting learning outcome positively. In contrast, when tasks are very easy they may be solved more efficiently by applying automatized learning routines. At the other extreme, very difficult tasks, application of metacognitive knowledge and engagement in metacognitive activities might even result in termination of learning activities because of the realization that investing more effort is hopeless. Metacognition has become an integral part of curricula in schools (Hasselhorn, 2001).

Nowadays, metacognition is not merely taught for its own sake, but it is seen as promoting transfer in content-specific cognitive training programs. Many researchers agree that metacognition should be taught in the context of teaching content-specific skills, like strategies, or learning areas (Schneider & Pressley, 1997).

Metacognitive processes are linked bidirectionally with attributive processes: the acquisition of metamnemonic knowledge is influenced by causal attributions of past performance, and attributions also play an important role for the future application of knowledge about metamemory in transfer situations. For instance, in classes in which teachers provided their students with support in the form of feedback on tests and homework, quizzes and sample test questions, more strategy use was reported (Thomas et al., 1993). Thus, monitoring can lead to improved performance which, in turn, is motivating for learners if they attribute their success to strategy use.

In short, metacognitive, strategic and motivational processes constitute elements of self-regulated learning that can barely, if at all, be considered in isolation. Teaching students to become self-regulated learners is a main goal of schooling. Self-regulated - or self-determined – forms of learning are approaches in which the learner himself is enabled to make the major decisions concerning whether, what, when, how and why he learns (Weinert, 1984). But learning in school and guided learning may also be self-regulated. "Students can be described as self-regulated to the degree that they are metacognitively, motivationally and behaviorally active participants in their own learning process" (Zimmerman & Martinez-Pons, 1990; Zimmermann, 1989). According to the authors, self-regulated learners are superior with respect to the following characteristics: (1) motivationally, because of their high degree of self-efficacy and intrinsic motivation, (2) behaviorally, because they select, structure and generate social and physical environments that allow for optimal learning and (3) metacognitively, because of their planning, organizing and assessment.

How can good strategy use be taught? Among the obstacles to instruction in strategy use are that there are very many and different kinds of strategies to learn and that teachers often do not think in information-processing terms (Schneider & Pressley, 1997). But perhaps more important is that it takes a great deal of effort for general, durable use of strategies to be developed. Learners must have the opportunity to practice the procedures until they are executed fluently and without great cognitive effort, and the relevant skills must be practiced in a variety of situations. Explicit efforts must also be made to teach students the critical aspect of when and where to use

a procedure (conditional knowledge). Pressley, Snyder and Cariglia-Bull (1987) argue that only more complete instructional techniques are likely to produce durable and general use of strategies and to promote the complete understanding of strategies. However, the implementation of complete approaches is often difficult because children vary in their abilities and motivation and the approaches place high demands on teachers where diagnostic and instruction-tailoring skills are concerned. However, it seems that "expert" teachers are able to fulfill these demands almost automatically (Berliner, 1996). Pressley and colleagues list two favored approaches. The first is direct explanation of strategies. That is, the teacher explains strategies thoroughly, providing explicit and detailed information about how to carry out processing and about the effects of the strategy, as well as conditional knowledge. Direct explanation includes concrete examples, modeling and teacher-guided practice, during which the difficulties of individual children are taken into account.

The second approach favored by Pressley, Snyder and Cariglia-Bull (1987) is Reciprocal Teaching, a program in which teachers and students take turns executing strategies that are taught during the instructional dialogue. The special feature of this program is that teachers only assume more responsibility for the learning process early in instruction; they gradually transfer control to the students, meaning that the students themselves eventually take on the "teacher" role. In Reciprocal Teaching, instruction occurs in true dialogue, in which strategic processes are made overt. Students are given plenty of exposure to the modeling of strategy use and lots of opportunities to practice the reading strategies. The children discover and the teacher conveys strategy utility information and conditional knowledge. Reciprocal Teaching has proved to be very effective. In most training studies, good or very good effects have been observed with different populations of students and various age groups up to adolescents and adults.

This thesis was undertaken (and an experimental study conducted) to find out more about the mechanisms that are responsible for the success of the Reciprocal Teaching method and to identify major components that contribute to or even cause the large increments in reading comprehension that the method usually produces. Before describing the method and some research results as well as the main research questions and design of the present experimental study, however, I would like to go back to two components on part of the reader that influence reading comprehension and that are susceptible to training in Reciprocal Teaching, strategies and metacognition. One

question has not yet been addressed: from what age on is it possible and efficient to train children in strategy use and to foster their metacognitive knowledge and skills?

Already Piaget argued that there are preconditions for self-regulated learning. In his view, self-regulated learning is not possible until the stage of concrete operations at the earliest. The egocentrism of the child (i.e., the inability to take the viewpoint of someone else, to imagine oneself in the position of somebody else, or to consider one's own perspective as just one among many) and the focus on just one or very few aspects that characterizes the pre-operational stage (lasting from about age 2 until about the beginning of school) must first be overcome. This is the case when children move on to the stage of concrete operations - they are then usually able to consider various relevant dimensions simultaneously and to differentiate between their own perspective and that of other people. However, at this stage their thinking is bound to the information provided, whether it is represented physically (in concrete or visible form) or linguistically. More elaborated self-regulated learning is not possible before the stage of formal operations, however. The major characteristic of this stage is that thinking beyond given information is possible, meaning that judgments and inferences can be made on the basis of the information available as well as information yet to be derived, created or proved. Children become able to mentally combine variables. Moreover, the products of one's own thinking can become the objects of thinking – thus, monitoring and control of one's own progress in thinking on a higher level (or metacognition) is possible (Zimmermann, 1989). This stage usually begins towards the end of the elementary school years at around age 10. In many countries, this is also the age at which scientific school subjects like chemistry and physics requiring the ability to systematically control variables, formulate hypotheses and methodically search for solutions are introduced.

1.3. Reciprocal Teaching – a promising training program

Reciprocal Teaching is a very successful training program encompassing many of the components that have been shown to be influential and predictive for reading comprehension. Therefore, it seems worth investing effort to find out more about the basic mechanisms leading to improved performance and the components of the program that contribute to its success.

1.3.1. Description of the program

The focus of the studies conducted by Palincsar and Brown was to improve students' ability to learn from texts. They view reading comprehension as the product of four factors: considerate texts, compatibility of the reader's knowledge with the text content, use of active strategies to enhance understanding and retention and to circumvent comprehension failures, and decoding fluency. Comprehension failures occur when an expectation about the text has not been confirmed or when unfamiliar concepts are encountered too frequently. Mature readers react to comprehension failures by slowing down their processing rate and allocating time and effort to overcome the problem. Their mental state can be described as deliberate and systematic, and they apply a variety of active processing strategies that focus on the material itself (comprehension-fostering strategies) and monitor mental processes during reading (comprehension-monitoring strategies). Palincsar and Brown describe six of the most important underlying activities necessary for text comprehension: (1) understanding the purpose of reading, (2) activating relevant background knowledge, (3) allocating attention to the major content, (4) critical evaluation of internal consistency and compatibility with prior knowledge, (5) monitoring whether comprehension is occurring and (6) drawing and testing inferences (interpretations, predictions, conclusions).

The program the authors developed, Reciprocal Teaching, is an instructional approach designed to teach students cognitive strategies with the intention of improving reading comprehension. The focus is on teaching students four specific comprehension-fostering strategies: generating questions, summarizing, clarifying word meanings or confusing text and predicting what might come next in the text. All four strategies involve activation of background knowledge (2) through discussions at the beginning of each session, and understanding the purpose of reading (1), which is to be able to answer questions about the text. The strategies summarizing and questioning both help students to learn to allocate attention to the major content (3) and to monitor whether they have understood the text (5). Critical evaluation of the text (4) is accomplished by the strategy of clarifying, and students learn to draw and test inferences (6) by predicting what will happen next.

Teaching takes place in the context of a dialogue between the teacher and the students. Participants are assigned one of two roles: the *teacher* role or the *student* role. Both read a text passage silently and then one of the *students* carries out the four

strategies for that text section. The roles are switched for the next paragraph. When *teacher* and *students* take turns, the *students* practice using the strategies, while being supported by the *teacher*, who gives feedback, hints and explanations and guides the correction of answers. In the early stages of the training program, the classroom teacher plays the *teacher* role. He is responsible for teaching the reading strategies, modeling their use and providing conditional knowledge. He also explains the *teacher* role to the children, helps them to apply it, and corrects children playing the role of *teacher*, if necessary. As the children become more familiar with the strategies and the procedure, the teacher takes a back seat and the responsibility for carrying out the tasks and implementing the procedure is shifted to the students, who now do the majority of tasks and thinking. Later on, the children do everything by themselves and the classroom teacher is merely an observer or leaves the group entirely to its own devices and is only asked for help if the group has a problem that cannot be solved by the children alone.

In their 1984 publication, the authors report the results of a pilot study and two larger studies. In the pilot study, only one 7th grade poor reader was trained with the reciprocal teaching procedure. His performance in answering questions about texts increased from a baseline of 15% correctly answered questions to 85% after training. Even after a 6-month interval, he still performed at 60%, and with only one additional day of training the rate was again 85%. The next study was carried out with 24 students who can be described as adequate decoders (who read at least 80 words per minute) but poor comprehenders; that is, they read at a comprehension level of two years below their grade level and their baseline performance in answering questions about the text was below 40% correct. The materials used in the training were 13 passages of 1500 words each; the assessments were 45 passages of on average some 400 words each, with a total of 10 text-explicit, text-implicit and script-implicit questions. Three experimental groups and one control group were formed. Students in the control group attended their normal reading class and only took the pre- and post-tests. The teacher of the reciprocal teaching group was Annemarie Palincsar, who trained students in groups of two for 20 sessions of 30 minutes each. In another group, students were exposed to an alternative intervention – locating information. They were shown by the teacher where in the text the answer to a question is located, how to combine separate sections to answer text-implicit questions and how to use prior knowledge to answer script-implicit questions. Students in the third group were given the daily assessment tests, but no training in groups of two. Generalization probes were conducted in regular

social studies and science classes at baseline, during the intervention, during a maintenance period and after 8 weeks. The improvement of the reciprocal teaching group was first apparent in the dialogues, and then on the test scores. Throughout the training period, there was an increase in the quality of dialogues of the reciprocal teaching groups: more main idea questions were asked, the summaries contained more main ideas, the number of details in summaries decreased and fewer incorrect or incomplete statements were observed. In a repeated measures analysis of variance in the daily comprehension assessments with group (3) and phase (4) as factors, main effects of group and phase and a significant group x phase interaction were found. Only the reciprocal teaching group improved significantly, and the gap between this group and the other groups widened over the course of training. The same pattern of results was evident when analyzing the generalization probes in the classroom settings; all groups began at the same level but only the reciprocal teaching group showed steady improvement (from 20% to 60% correct answers). Moreover, students trained with the reciprocal procedure maintained their level of performance during the maintenance period and at follow-up after 8 weeks.

In their second study, the same materials and procedure were used as in study 1. However, the training took place in students' regular school reading groups and was directed by classroom teachers who had volunteered to participate in the study. Students worked in groups of 4 to 7 students. A similar improvement in the quality of dialogues was reported as in study 1, but the improvement was less dramatic. This was explained by the bigger group size. Training the students in larger groups also had some advantages, however: the students provided modeling and feedback for each other and thus learned from their peers as well as from their teacher.

The procedure has now been applied in different settings and extended to other age groups and different populations; it has even been adapted for usage with illiterate adults. To provide an overview of these studies and their main results as well as the major dependent variables that are usually used, two important meta-analyses on the effectiveness of Reciprocal Teaching will now be described.

1.3.2. Meataanalyses on the effectiveness of Reciprocal Teaching

(Moore, 1988) was the first to conduct a meta-analysis on the effectiveness of reciprocal teaching. He reported that the procedure produced larger gains in terms of students' performance in daily comprehension tests than modeling and explicit

instruction in strategy use did, although all forms of instruction led to improvements (Brown & Palincsar, 1985). Reciprocal teaching was also effective when peers served as tutors – both tutors and tutees improved significantly in a 12-session training program of 35 minutes each (Palincsar, Brown, & Martin, 1987).

The best overview of studies examining the reciprocal procedure was published by Rosenshine and Meister (1994). Their meta-analysis included 16 studies meeting a number of criteria, including making explicit reference to Palincsar and Brown (1984) and including comparable control groups as well as experimental groups. There seems to be a major difference in the effectiveness of the studies depending on the outcome measure. For standardized tests like the Gates MacGinitie Reading Test, the mean effect size was .32 standard deviation and only significant in two out of nine studies. In contrast, the difference between reciprocal teaching and control groups was almost always significant (in six out of seven studies) when experimenter-developed short-answer or multiple-choice tests were used, with a mean effect size of 1.00 standard deviation. Also, in four of the five studies that collected summarization probes, significant improvements in the reciprocal training (RT) groups were detected (mean effect size of .85). Students' improvements in applying the individual strategies instructed in the context of the RT program were only assessed in a few studies: one study examined predicting and found a significant improvement for the RT group. For generating questions, the picture to emerge was less clear: in six studies there was no difference between reciprocal and control groups in either the level of questions generated or the number and quality of these questions, although significant improvements in the text comprehension of the students who participated in Reciprocal Teaching were observed in all of these studies. Only one study found differences in the questions generated in favor of RT. The program seems to work for different kinds of students: not only poor readers and good decoders/poor comprehenders showed gains in comprehension, good readers also profited from the method. The reciprocal teaching program was also successful for a wide range of ages, with significant improvements in text comprehension being observed from 4^{th} graders to adults. The significance of the training effects seems to be independent of the person providing the instruction (either an experimenter or the regular teacher), the number of reading strategies taught (which varied between 2 and 10, but was generally the original 4 strategies) and the size of the instructional group (range: 2 to 23 students). The number of instructional sessions ranged between 6 and 100; in most studies 10 to 25 training sessions were realized. The

studies can be classified into two groups: reciprocal teaching only (RTO approach), where all modeling and instruction took place during training, or explicit teaching of the strategies before training (ET-RT approach). There was little difference between the results for these two approaches.

1.3.3. Summary and critique

Reciprocal Teaching seems to be a very successful training program which works for a wide range of ages and ability groups. Furthermore, it does not seem necessary to apply all four reading strategies that Palincsar and Brown (1984) used, nor does it seem to do any harm to include more strategies in the training program. However, some questions about the program and the effects it produces remain unanswered. Evidently, whether or not positive effects of the training method are found, and how large these effects are, seems to depend very much on the dependent measures used for evaluation. Positive training effects are almost always observed with experimenter-developed comprehension assessments, and these were very much larger than the effects found for standardized reading tests like the Gates MacGinitie Reading Test. Another criticism is that, in my opinion, the reading comprehension measures that Palincsar and Brown (1984) and many of the other researchers used do not primarily measure the effects of the training program itself; rather, they are measures of transfer. During training, students were taught reading strategies and applied these to texts; however, the dependent measure used by Palincsar and Brown was the number of correctly answered comprehension questions about texts shorter than those used in training. Moreover, the assessment took place after training and did not imply actual application of the reading strategies taught in the training program. Some more interesting questions concerning Reciprocal Teaching will be addressed in more detail in the following section.

1.4. Research questions

Palincsar and Brown provided only few theoretical assumptions and conceptions in their 1984 publication. A central frame of reference for their work was Vygotsky's developmental theory, with the focus on the zone of proximal development. This zone represents the framework within which a child is sensitive to instruction; the assumption of its existence has led to the development of new instructional approaches, one of which is proleptic teaching. Here, the child works on a task at his own rate, at the level

he is capable of and a little beyond with the help of an expert, because it is assumed that with a little instructional help the child can execute new tasks and acquire new competences. The Reciprocal Teaching intervention was designed according to the principles of expert scaffolding, where the child learns a particular activity by first participating as spectator watching an expert performing the task and doing little of the actual work. Then, as the child becomes more and more experienced and capable of carrying out some of the activities, he is given more responsibility by the expert and gradually takes over the major obligation for doing the job. Finally, the child is able to complete the task by himself. Reciprocal Teaching is in line with this approach; it is an intervention that mimics naturally occurring guided learning interactions in which the teacher both models appropriate comprehension-fostering activities and guides the child to participate at an ever-increasing level of performance. Here, too, the teacher's support is gradually withdrawn, with the children taking on responsibility for their own learning. Reciprocal Teaching occurs in a social setting. By modeling strategies, teacher and students explicitly state thoughts and processes that usually remain internal. In this way, communication about thinking processes between people is possible. Vygotsky assumed that the inter-individual processes thus become intra-individual processes (are internalized), leading to metacognitive knowledge and skills being acquired.

However, Palincsar and Brown did not explicitly address the issue of which concrete processes or mechanisms occur during or after Reciprocal Teaching of reading strategies, or of how the large increments observed in comprehension are produced. For me, this is an important question. Another motivation for taking a closer look at Reciprocal Teaching and trying to evaluate the method was that it is very unusual (at least with children of this age) to give students almost complete responsibility for their learning and let them be "teachers" themselves. It is also noteworthy that children of all ability levels are given the "teacher" role. It seems reasonable to assume that the effects of the Reciprocal Teaching method might have something to do with this unusual procedure. Other researchers (for instance, Pressley, Levin, & Ghatala, 1988) have also called for more detailed analyses of strategy training effects, which should be designed to partial out the causal components of training packages (decoupling of effective elements).

1.4.1. The mechanisms in effect

Most certainly, the success of the training program cannot be solely attributed to improvement (in quality and/or frequency) of the reading strategies instructed or to actual use of the reading strategies. The first reason for this is that there was not enough time during the comprehension tests for learners to apply the four time-consuming reading strategies they had learned (summarizing, questioning, clarifying and predicting). In the Palincsar and Brown (1984) study, training success was measured with 10 relatively easy comprehension questions on texts of 200 to 800 words (M = 425). The tests were administered prior to and after training; the dependent measure was the number of questions answered correctly. Palincsar and Brown observed an increase from 15 % (baseline) to 85 % of questions answered correctly. However, they did not assess whether students used these four or other reading strategies during test-taking. Second, in studies where the ability to apply the strategies was assessed, reliable improvement was only observed in one of the four strategies, namely, summarizing (see Rosenshine & Meister, 1994). The other three reading strategies have only been assessed rarely (in the case of predicting) or not at all (clarification), or have not improved as a result of training (asking questions).

If it is not better strategy application, what does cause the observed improvements in reading comprehension? My main hypothesis is that the cue to understanding these effects lies in the overt explanation of thinking processes to someone else. The description provided by Webb (1989) helps to clarify this argument: "… In explaining to someone else, the helper must clarify, organize and possibly re-organize the material (see Bargh & Schul, 1980). In the process of clarifying and re-organizing the material, the helper may discover gaps in his or her own understanding or discrepancies with others' work or previous work. To resolve these discrepancies, the helper may search for new information and subsequently resolve those inconsistencies, thereby learning the material better than before. Furthermore, when an explanation given to a team-mate is not successful (the team-mate does not understand it or does not use it to solve the problem correctly), the helper is forced to try to formulate the explanation in new or different ways. This may include using different language, such as translating unusual or unfamiliar language into familiar language (Noddings, 1985), generating new or different examples, linking examples to the target student's prior knowledge or work completed previously, using alternative symbolic representations of the same material (e.g., pictures vs. diagrams vs. words vs.

numbers vs. symbols), and translating among different representations of the same material. All of these activities will likely expand and solidify the helper's understanding of the material. Giving only the answer or other low-level information on the other hand, would be less likely to cause the helper to clarify or re-organize his or her own thinking."

I would even go further than Webb, in speculating that students' overt explanations of their thinking processes to others result not only in the material being better understood, but also in more knowledge about individual strategies, the conditions of their efficient use and special features, as well as more general metacognitive knowledge being acquired. Metacognitive knowledge and metacognitive skills are promoted by this instructional method, leading to a more efficient and goal-oriented use of strategies, at least some of which have already become automatic. By modeling the application of the strategies, providing instruction on when, why and how to use them (conditional knowledge) and engaging in repeated practice, the children acquire declarative knowledge about the strategies (specific strategy knowledge) and the conditions of their use. Students also have many opportunities to experience the usefulness of the different strategies through repeated practice and receive frequent feedback from their peers and the teacher during training. Moreover, some of the strategies may become automatic because they are executed repeatedly in a routine manner. This can be assumed to result in procedural metacognitive knowledge about the usefulness of the strategies and the way in which they are applied as well as relational strategy knowledge. This knowledge does not necessarily need to be expressible; it can be sub- or unconscious, but it does lead to improved performance and appropriate use of strategies as tools promoting the learning goal of text comprehension. Thus, the Reciprocal Teaching procedure is in line with Borkowski et al.'s model of mature metamemory (1988) and the way in which these authors assume that metacognitive knowledge is acquired.

To summarize, by acquiring and practicing reading strategies in Reciprocal Teaching, students learn more general metacognitive knowledge and metamemory acquisition procedures, resulting not only in better ability to apply the reading strategies that were taught in a routine manner, but in mindful application of strategies in general, and in monitoring and regulating activities. Metaphorically, this can be explained as follows: I assume that the strategies learned in training are cognitive tools. Students get to know the tools (like a hammer, saw, screw-driver and flat-nosed pliers), are

taught and discover the conditions of their use and their special strengths and weaknesses in comparison to other tools (that it is more efficient to use a hammer than flat-nosed pliers when the goal is to drive in a nail) and acquire the skills that actually enable them to use these tools in challenging learning situations. Moreover, this exposure to cognitive tools makes them more alert in other situations in which usage of tools is likely to result in better performance, and more capable of unconsciously and automatically monitoring their performance and adjusting their activities to the learning task (that is, they would automatically start looking for a suitable tool when trying to solve a new task without first thinking that tools might be helpful in solving the task).

1.4.2. The effective features

The second goal of the study is to identify key features of the program that are responsible and/or necessary for the improvement of reading comprehension and metacognitive knowledge and skills. The situation in which training takes place is quite unusual for the students: they "play" the role that the teacher normally assumes: assigning a passage and appointing other students (or the teacher) to carry out the strategies. They also have to evaluate the other participants' performance, give feedback about the accuracy of answers and, if necessary, provide help and guidance for the other students. In order to do this, they not only have to decide if the answer was correct with respect to content, but also have to be aware of how the strategy should be applied and determine if this was done correctly. Moreover (especially in later stages of training), not all of the four strategies need to be carried out for every passage; here the "teacher" has to decide which of the strategies are useful in the particular context. This requires not only declarative knowledge about the different reading strategies, but also conditional knowledge about the strategies and routines (procedural knowledge) used to carry out the strategies. The turn-taking procedure allows each student to get sufficient practice, observe a number of models, including the teacher, and be given frequent feedback. In sum, being assigned the teacher's role involves complex monitoring in addition to understanding the text. It gives students the opportunity to acquire metacognitive knowledge in a very motivating setting: students are responsible for their own learning and work cooperatively on texts. External help is available when needed. Additionally, being assigned the teacher's role and taking responsibility for their own learning should not only be a challenging task for the students, but also a very motivating condition that keeps them "on track."

Identifying the important features of the program is a complicated undertaking for two main reasons: first, the program has many features and, second, it is impossible to view these features as independent factors or to isolate them in order to specify an orthogonal design which would allow a definite interpretation. A few candidate features would be: giving feedback on other students' performance, guiding the other students' learning process, explicitly stating one's own thoughts, applying the strategies in a highly structured social interaction with a lot of external cues and help, deciding which strategy to apply to a given paragraph and, last but not least, being assigned a different role than usual – that of teacher. To illustrate the complex interplay and interdependence of these features I will provide one example: Not assigning students to the teacher's role would not only mean that the teacher no longer has multiple tasks, but also that the interaction is determined by the instructor throughout, and that the children do not have the opportunity of gradually assuming more responsibility for their own learning.

When we look at the complex tasks assigned to the child playing the *teacher* role in Reciprocal Teaching, it becomes clear that there are two types of tasks: tasks that have to do with organizational matters and tasks that deal with content. The content-related tasks are the following: The *teacher* decides which of the strategies can be meaningfully applied to the text passage in question. When the *student* applies a strategy, the *teacher* has to determine whether the student has done so correctly - with respect to both content and strategy application - and provide the *student* with feedback. If the answer was not perfect (which happens most of the time), the *teacher* has to help the *student* to improve the answer and guide him during this process. If all else fails, the *teacher* should model how it could be done, explaining to his *students* how he arrived at this answer. All of these tasks are content-related tasks and serve what I will call a "monitor" function. The nature of the other tasks is organizational: selecting the strategies to be applied by the other students, deciding when to move on in the text (to the next passage), handing over the *teacher* role, maintaining discipline in the group and managing disturbances.

It is assumed that the content-related tasks of the *teachers* are those that lead to acquisition of metacognitive knowledge and cause reading comprehension to improve. By monitoring other students' understanding of the text and application of the reading strategies, the children learn a great deal about the strategies and also about evaluation of understanding. Moreover, because they are also expected to help the *students*, they

must go beyond monitoring and evaluation and acquire knowledge and skills about regulative activities. In summary, they show mature metacognitive behavior on an inter-individual plane. When demonstrating (modeling) a strategy to others, they make their thinking process – rather than just the products of this process – explicit; their thinking thus becomes the subject of discussion and the whole group is able to learn from this thinking process or to help discover mistakes and find alternative or better ways of solving the task. These metacognitive skills (reflecting upon one's own thinking and modifying it, if necessary) first occur inter-individually and can then become intra-individual by internalization.

In contrast, the organizational functions *teachers* fulfill do not have much to do with the strategies or monitoring processes. They are additional tasks that need to be performed, consuming some of the children's limited available processing resources. However, these tasks need not necessarily only be a burden and extra work for the children: it is very likely that assuming the role of the teacher is very motivating for them in that they are proud to be assigned such an important task and role. It is also plausible to assume that responsibility for the group and its ongoing learning activities leads to increased self-efficacy and intrinsic motivation, and strengthens the belief that strategic behavior leads to better performance, because the children have the opportunity to experience themselves as effective and successful, allowing for feelings of competence and self-determination. These effects can be assumed to occur for all children in the group because the *teacher* role is continually rotated between the children. Another positive aspect may concern the interactions that occur in the group: in fact, the group as a whole is given responsibility by the classroom teacher. This may lead to more cooperation between the children and provide opportunities for positive social relationships to develop. For these reasons, in addition to the two main research question another hypothesis concerning motivation is formulated. It is postulated that the Reciprocal Teaching procedure also produces benefits in terms of measures of motivation.

1.5. Design and hypotheses

1.5.1. Design of the experimental study

To answer the two main research questions [see above] an experimental study was conducted with 5th graders. This age level was chosen for several reasons: At this age, children usually understand that memory is influenced by person and task variables

(Flavell & Wellman, 1977) and start to use strategies mindfully and efficiently (Weinert & Schneider, 1987). Moreover, children's knowledge about text and its structure improves, and their ability to differentiate between important and non-important information improves markedly in grade 5. By grade 7, their importance ratings are almost as good as those of college students (Schneider & Pressley, 1997). At the end of the elementary school years, decoding abilities are fairly well developed so that, when provided with easy or moderately difficult texts without time pressure, there should not be many children whose comprehension difficulties are rooted in non-sufficient decoding abilities. Last, but not least, there was also a pragmatic reason for choosing 5th graders: in the federal state of Brandenburg, where the study took place, it is difficult to conduct additional studies in schools in grade 6 because teachers' and students' efforts are directed towards achieving good grades – the grades students receive in the first semester of their final elementary school year are of huge significance for the type of secondary school track they can attend.

A reading strategies training program as designed by Palincsar and Brown (1984) was set up with small mixed-gender and mixed-ability groups of 4 to 6 children each. Each of the groups had one trainer who was present for the whole time. The training program was conducted during a short, but intensive period of time to allow for continuous practice without longer breaks in which students might forget what they had learned. Therefore, 4 one-hour training sessions per week were scheduled for a period of 4 consecutive weeks, resulting in a total of 16 sessions. This was also fairly practicable with respect to the frequent German school breaks.

The remainder of this section describes how the study was set up to address the research questions. To test whether metacognitive competences are responsible for the improvement in text comprehension (Research Question 1) a number of dependent measures concerning strategy knowledge, strategy application, metamemory, etc., were included in the study. An experimental variation of tasks carried out by either the children or the trainer serves to address the question of which features of Reciprocal Teaching make it so effective (Research Question 2).

1.5.1.1. Dependent measures

To test whether metacognitive knowledge is the key factor for the success of reciprocal teaching, a focus was placed on measures of different aspects of metacognition. Students were required to summarize a short text, and measures of specific strategy knowledge about two of the reading strategies taught in the training

program – summarizing and clarifying – were included in the study. Students' relational and conditional strategy knowledge and their planning knowledge were also assessed.

Reading comprehension was assessed prior to and after training: after reading longer expository texts, students were posed open-ended questions which they had to answer without referring back to the text. To answer most of these questions, students had to make inferences based on textual information and a global understanding of the macrostructure of the text requiring deeper levels of understanding, rather than simply identifying single pieces of information.

Additionally, by collecting data on the students' basic cognitive abilities and skills, like their nonverbal reasoning ability, vocabulary, listening and reading ability and decoding speed, it was possible to control the effects of these variables on measures of metacognition and comprehension.

1.5.1.2. Experimental Variation

Based on the assumption that the content-related tasks of the *teacher* are those that lead to the acquisition of metacognitive knowledge and skills, the assignment of the tasks to be performed by either the trainer or the students was experimentally varied. Three different experimental conditions were specified: (1) In the first condition, *Reciprocal Teaching*, as described by Palincsar and Brown (1984), was realized. (2) In the second condition, the *monitor* condition, the content-related monitoring tasks were separated from the organizational tasks usually assumed by the *teacher*. The children were always *students* in the discussion, which was led by the trainer. Before a strategy was applied to a text passage, however, the trainer appointed a second student to monitor the performance of the student applying the reading strategy and asked him to give feedback on it. The *monitor* student was also responsible for helping to correct the answer. The children were also asked which strategies could be meaningfully applied to the text passage and in what order this should be done. (3) The third and final experimental variation, the *student* condition, strongly resembled the traditional instructional setting in schools: here, children only applied the reading strategies while the trainer was responsible for giving feedback, correcting their answers and managing the classroom. Students were neither assigned organizational tasks nor were they charged with evaluating other students' responses or guiding the correction of answers. All they had to do was to apply the strategies when told to do so.

The ET-RT approach (explicit teaching of the strategies before training) was chosen for the study. All of the experimental groups received the same instruction in the four reading strategies prior to training. This was done in the first three training sessions. The strategies were not only explained but also practiced using worksheets. The strategies were introduced following a guide for teachers developed by (Palincsar, David, & Brown, 1989). From the fourth session on, the way in which the strategies were practiced varied between the three experimental groups.

1.5.2. Hypotheses

There are three major hypotheses. The first refers to the role of metacognitive knowledge and skills, the second to specific features of the training program, and the third focuses on motivational issues.

1.5.2.1. Hypothesis 1: The mechanism in effect

The success of Reciprocal Teaching is not merely due to the more frequent use of the strategies instructed, but to the metacognitive knowledge and skills acquired through the procedure. The students learn not only specific strategy knowledge and conditional knowledge about the usefulness of the strategies, but also acquire relational and more general strategy knowledge as well as procedural metacognitive knowledge about the actual use of the strategies. This is mainly accomplished by repeated application of the strategies to texts, the opportunity to learn from the overt thinking processes of experts and other students, and the process of monitoring and regulating the learning processes of other students.

To test this hypothesis, various assessments of students' metacognitive knowledge were implemented in addition to comprehension tests. More specifically, specific strategy knowledge about summarizing and clarifying, relational and conditional strategy knowledge and planning knowledge was assessed. Students were also asked to write a summary of a short text passage. There are not only differences in performance expected but also in the amount of metacognitive knowledge acquired: students who participated in the *strategy training* should outperform *control* children.

There is a specification necessary for this hypothesis: only the *reciprocal* and *monitor* conditions will be considered as "training conditions" and contrasted with the *control* group to estimate the effects of a strategy training on metacognitive knowledge and reading comprehension with the Reciprocal Teaching method, or more specifically

a training incorporating features effectful in Reciprocal Teaching. The reason for excluding children who were taught reading strategies in the *student* condition was that this condition does not involve the features that are hypothesized to cause the mechanisms that are hypothesized to be effectful in Reciprocal Teaching and to be responsible for the large learning gains; that is the student's task of providing feedback and corrective guidance for one another as a main determinant of the success of the Reciprocal Teaching program.

1.5.2.2. Hypothesis 2: Features of the program

I assume that the content-related tasks associated with the *teacher* role are those that lead to the large improvements observed in metacognitive knowledge and skills. By evaluating, monitoring and regulating other children's learning process, children execute metacognitive activities in an inter-individual way. These skills, first executed inter-individually, become internalized and are the basis for true metacognitive (intra-individual) activities and skills.

To test this assumption, a monitor condition was set up in addition to the reciprocal teaching condition. In this condition, children fulfilled the *monitor* function of *teacher*-students, but they did not carry out the organizational tasks of the *teacher*. These two conditions (*reciprocal teaching* and *monitor*) will be contrasted with the *student* condition to test my second hypothesis. I assume that the former two conditions are superior to the latter one with respect to strategy execution, text comprehension and increase in metacognitive knowledge.

1.5.2.3. Hypothesis 3: Motivational effects

In deciding which strategies to apply, giving feedback and guiding the correction of answers, children have the opportunity to feel competent and experience themselves as self-determined. The reciprocal teaching procedure also allows positive social relationships to develop. The fulfillment of the psychological needs of experiences of competence and self-determination and the desire for social relationships should result in increased motivation.

The responsibility that children in *reciprocal teaching* groups assume for their own learning process and for the group as a whole is assumed to be very motivating for these students. Positive, but less dramatic effects on motivational measures are also expected for children in *monitor* groups, who are also responsible for deciding which strategies to apply and for evaluating and guiding the learning of other students. In contrast, the motivation of the children in the *student* condition is expected to be lowest

because of the monotonous pattern of interaction and the students' lack of responsibilities or opportunities to influence the course of the session.

2. METHOD

First, the study participants will be characterized. The administration of the pre- and post-tests and the organization of the training program will then be described, followed by a detailed description of the material used. Finally, the experimental conditions will be characterized in detail.

2.1. Participants

Participants in this study were all 5th graders at four elementary schools in the city of Potsdam who were present at pre- or post-test. The total number of children was 221. However, we only used the data of those children whose parents gave their permission; this is required by law if data related to the person is collected, as was the case in this study. The parents of 169 of the 221 participating students gave their permission (76.47 % of the parents). All other test booklets were destroyed. Table 1 presents means and standard deviations (in parentheses) of the age, nonverbal and verbal reasoning ability, and grade achieved in the native language for all 169 children, as well as for the subset that participated in the reading strategy training program.

Table 1. Main characteristics of participants.

School	Gender	N of students[1]	N of training students[2]	Age	Reasoning ability		Grade in native language
					Nonverbal	Verbal	
1	male	14 [20]	9	11.33 (.48)	47.07 (10.18)	48.50 (8.94)	2.43 (.76)
	female	10 [14]	7	11.32 (.32)	45.00 (7.42)	45.55 (9.56)	2.18 (.60)
	total	24 [34]	16	11.33 (.41)	46.16 (8.95)	47.20 (9.15)	2.32 (.69)
2	male	30 [37]	11 [2]	11.45 (.39)	51.14 (8.59)	43.79 (10.28)	2.37 (.89)
	female	39 [44]	17 [2]	11.55 (.51)	50.44 (10.79)	45.69 (9.37)	2.31 (.95)
	total	69 [81]	28 [4]	11.51 (.46)	50.74 (9.85)	44.90 (9.73)	2.33 (.92)
3	male	26 [33]	9 [2]	11.77 (.59)	53.32 (7.78)	52.24 (8.75)	2.38 (.75)
	female	23 [28]	10	11.64 (.51)	57.43 (8.16)	50.33 (7.86)	1.83 (.78)
	total	49 [61]	19 [2]	11.71 (.55)	55.20 (8.13)	51.37 (8.32)	2.12 (.81)
4	male	21 [24]	13 [4]	11.00 (.56)	49.20 (6.44)	39.81 (9.85)	2.48 (.81)
	female	6 [11]	3 [1]	11.03 (.66)	50.67 (6.06)	42.17 (12.54)	2.33 (1.21)
	total	27 [35]	16 [5]	11.01 (.57)	49.54 (6.27)	40.33 (10.29)	2.44 (.89)
Total	male	91 [114]	42 [8]	11.42 (.57)	50.67 (8.35)	45.70 (10.84)	2.41 (.80)
	female	78 [97]	37 [3]	11.51 (.52)	51.78 (10.02)	46.63 (9.47)	2.15 (.90)
	total	169 [221]	79 [11]	11.46 (.55)	51.18 (9.15)	46.12 (10.21)	2.29 (.85)

[1] N of students with parental permission; total number of 5th graders in the school shown in brackets.
[2] N of students who dropped out early given in brackets.

Of the 169 children, 91 expressed an interest in participating in the training program. 79 children actually participated; 11 of them dropped out early for different reasons (not enough time for homework or play in the afternoon; no further interest in the program; dropout because of being forced to participate in the program by their parents).

2.2. Procedure

After the schools had been contacted and the principals had agreed to participate in the study, letters were sent to the parents to inform them about the tests and the training study. Parents were asked to complete the permission form and to indicate whether or not their child wished to participate in the training study. The training took place in the afternoon after the last regular school session, which was usually at 2 p.m.

First, all 5th graders in the school were tested in two consecutive school lessons (pre-test before training). Within a week, the strategy training started for the children who had volunteered to participate. The duration of the training was between 4 and 5 weeks (due to national holidays and school breaks). About 3 days after completion of the training program, all 5th graders at the school were tested again in their classes in two school lessons (post-test after training).

2.2.1. Pre- and post-tests for all children

Table 2 lists the tests that were administered to all students before and after the training program. Testing took place during two consecutive regular school lessons at both pre-and post-test. To ensure that conditions were similar for all students, the tests were administered during the first two lessons or after the mid-morning break (which is 25 minutes long) in lessons 3 and 4. Besides the experimenter, a teacher was always present to help maintain discipline. Children who finished the tests earlier than their peers were usually assigned additional practice tasks by the teachers; on some occasions they were allowed to start their break early. A detailed description of the tests is provided in the material section.

2.2.2. Training in small groups

The parents of the children who applied to take part in the training program were contacted via telephone and children were assigned to the training groups according to their time schedule. So in general, the groups consisted of children of different classes.

The aim was to have an equal number of girls and boys and children of all attainment levels in every group. All children who volunteered to participate in the training program were included in the study. Group size ranged from 4 to 6 children (originally, 4 children per group were planned). In the first three schools, all experimental conditions were realized. Because of the large number of children who wanted to participate in the second school, two groups for every condition were run in parallel. In the last school, a second round of data collection was necessary to replace the three groups that were excluded from data analysis (please see the results section for more information).

Table 2. Tests administered to entire classes prior to and after training.
Tests marked with * are standardized tests.

Time	Test	Duration in minutes
Pre	Knuspel's Leseaufgaben *	
	subtest 1 (listening comprehension ability)	≈ 5
	subtest 3 (decoding ability)	4.5
	subtest 4 (reading comprehension ability)	8
	KFT *	
	subtest V1 (verbal reasoning ability)	7
	subtest N1 (nonverbal reasoning ability)	9
	---- BREAK ----	
	Reading Speed	5
	Text comprehension 1	
	- read text	≈ 5 to 10
	- answer questions about the text	≈ 10 to 15
	Knowledge about usefulness of reading strategies	≈ 10 to 15
	Index of Reading Awareness (IRA)	≈ 5
	Questionnaire about reading habits	≈ 5
Post	Write a summary of a text (*Popcorn*)	≈ 10
	Knowledge about usefulness of reading strategies	≈ 10 to 15
	Text comprehension 2	
	- read text	≈ 5 to 10
	- answer questions about the text	≈ 10 to 20
	---- BREAK ----	
	Index of Reading Awareness (IRA)	≈ 5
	Verbal self-concept	≈ 1
	Knowledge about reading strategies	
	Declarative knowledge about Summarizing	≈ 5
	Procedural knowledge about Summarizing	≈ 8
	Procedural knowledge about Clarifying	≈ 5

Within two days of the pre-test, training started. 4 sessions per week were scheduled over 4 weeks, resulting in 16 training sessions. Because of national holidays and school breaks, the period in which training took place varied between 3½ and 4½

weeks. It was not possible to conduct 16 separate sessions in every group; therefore the decision was made to conduct double sessions of 1½ hours' duration when necessary.

The training program consisted of two parts. During the first three sessions, the reading strategies were instructed. In sessions 4 to 15, these strategies were then practiced using text paragraphs. In the last session, two tests and a post-training questionnaire were administered.

2.2.2.1. Introduction of reading strategies in sessions 1 to 3

During the first three sessions, the term "strategy" itself and each of the four reading strategies were introduced. The trainers tried to activate students' prior knowledge and to use the examples students generated. The most important information was collected on flipcharts. The knowledge commonly developed by the group was thus accessible for further use throughout the training. Each of the four strategies was practiced by working on examples provided in worksheets.

All of the groups worked in the same way during this phase, irrespective of the experimental condition to which they were assigned. This was done to ensure that all children had roughly the same knowledge at the start of the "real" training, which consisted of intensive practice of the reading strategies.

2.2.2.2. Practicing the reading strategies in sessions 4 to 15

The major part of the strategy training program was devoted to extensive practice of the four reading strategies. Longer expository texts about animals and other phenomena occurring in nature were used as practice material. The units that the groups worked on were text paragraphs – these were meaningful parts of the text that dealt with a common content and consisted of at least three sentences.

The texts were read aloud by the children paragraph by paragraph, starting with the heading. Every student had a blue sheet that was used to cover up the rest of the text. This was done to ensure that all students worked on the same text passage and to prevent them from reading further – *predicting* is, after all, only a fun activity if everybody in the group is naive with respect to the upcoming content.

Each reading strategy that could be meaningfully applied was executed for every paragraph. For all but the last paragraph, all four reading strategies could generally be applied. For the heading of the text, the strategies of *clarifying* of unknown words and – once the meaning of the heading had been understood – *predicting* what the text will be about, were usually applied.

Ideally, the group started by clarifying and then tried to summarize the content of the paragraph. Thinking of questions a teacher might ask is probably the most difficult strategy; this became easier when the text was well understood – a good summary could help students to think of questions. Before starting to work with the next paragraph, predictions about its content were made. An example of children in a *reciprocal teaching* group applying the reading strategies to the heading and the first paragraph of the *Polar Bears* text is available in section 2.2.4.1.

2.2.3. Reading strategies

Students were taught the four reading strategies that Palincsar and Brown used in most of their studies. All of these are comprehension-fostering reading strategies, namely, *Summarizing*, *Questioning*, *Clarifying* and *Predicting*. Each child received a bookmark with the name and symbol of each of the four reading strategies and two or three short notes for each of the strategies as a reminder.

In summarizing and in formulating comprehension questions, the focus of attention is on the most relevant information in the text. The effectiveness of the reading process can thus be tested. When it is clear to the reader that he does not understand or comprehend something, the meaning of the unknown word or difficult passage should be clarified. Understanding of the text and prior knowledge are needed to make predictions about how the text will continue.

2.2.3.1. Summarizing

A good summary of a text has three important features:

1. it is shorter than the text
2. it contains only the most important content and no details
3. the text is paraphrased (formulated in one's own words).

These features were worked out together with the children in the introductory phase and written on a flipchart. Children were shown a number of ways to produce a good summary. Among these were underlining important content, generating titles for paragraphs, or listing and note-taking of the most important content. Generation of topic sentences was also practiced.

2.2.3.2. Questioning

The task for students was to think of difficult questions that tested whether the content of the text had been understood. The only condition was that the answer to the

question had to be provided in the text paragraph. During instruction of Questioning, no emphasis was placed on using interrogative pronouns like *Why*, *When* or *Where* because the content of the question was far more important. Students were made aware that not all questions start with an interrogative pronoun.

Difficult questions are usually ones that ask for the main ideas of the text rather than for details. A metaphor was used to give students an idea what kind of questions were desired: "Ask a question your teacher would ask." Students were also told that the answer to a question often helped to indicate whether it was difficult. Questions with one-word answers are usually easy to answer (e.g., "How much does a polar bear weigh?" – "1600 pounds"). The students were quite well aware of whether or not their questions met these requirements.

2.2.3.3. Clarifying

The strategy *Clarifying* was to be applied when the meaning of a word was unclear or a larger portion of text (a phrase or sentence) was not understood. There are two main ways to clarify in natural learning situations: work with the text itself to see if it provides any explicatory information or use external help. "External help" means asking other people to explain the meaning or consulting other resources. The people that students usually ask for help are their parents and grandparents, teachers, peers and siblings. Resources like dictionaries, reference books and the internet are frequently used to look words up.

In the training program, the children always tried to explain unknown words to one another whenever possible. In cases where they were not sure about the meaning or none of them knew the word, however, they were instructed to use the text itself to clarify, proceeding in the following way: first, they should read the unknown word again and then the whole sentence. If this did not solve the misunderstanding, they should read the preceding and following sentences – in other words, the paragraph - again. The focus should be on information in the text that might help them to understand the unknown word; this might be clues like phrases in commas or parentheses, or pointers like "or," "that means," etc.

In this way, mistakes that may have occurred during the decoding process and retrieval of the meaning from long-term memory (if a word has two or more meanings) can be eliminated. Furthermore, context knowledge and knowledge about language itself – its syntax and grammar – is used to resolve the meaning of unclear words or passages.

2.2.3.4. Predicting

In predicting future text content, students reflect upon the content of the text they have encountered thus far. When asked to think about what will come next, they can either draw upon their prior knowledge or use the text itself to come to a prediction.

Prior knowledge can involve either content knowledge or knowledge about language. For instance, texts about animals usually contain information about their appearance and behaviour, food and mating, the raising of their cubs and protection of the species; often in precisely that order. Content knowledge was used by many students when working on the text *Paper – a Web for Words*: they were fairly familiar with the method that was used in Egypt to produce paper.

Sometimes, however, the text itself provides information about how it will continue: if a problem has been presented but not yet solved, it is probable that the solution will be dealt with in the next passage(s). The most popular example of the reader permanently making predictions, being proven wrong and generating new hypotheses is the detective story – a character is killed early in the story and the rest of the book deals with finding the murderer.

Predicting on the basis of prior knowledge was relatively easy for the children. In contrast, using the text to formulate a hypothesis about future text content was rather difficult. Therefore, instruction in this strategy concentrated on locating clues in the text (for instance, unsolved problems).

2.2.4. Experimental conditions

The groups were randomly assigned to one of the three experimental conditions. Training materials were identical for all groups. Experimental conditions differed only in the assignment of tasks to the students or the trainer and, as a consequence, in the amount of responsibility assumed by the children. Table 3 shows who was responsible for the different tasks in each of the three experimental conditions. When the children were responsible, the name of the role they were fulfilling (“Student”, “Teacher” or “Monitor”) is also given.

Before explaining the experimental conditions in detail, the tasks that were the same for all trainers, irrespective of the experimental condition, will be described. These were especially important during the first practice sessions. Trainers had to instruct students in the use of the reading strategies and teach conditional knowledge about these strategies. They were required to model the use of the strategies and to think aloud and tell students how, when and where to apply the strategies. Students

should be prompted and provided with social support (scaffolded) when trying to apply a strategy, especially early in the training program. Trainers should also make clear to the students that there is no "correct" or " best" way to apply a strategy: there is always more than one way and different children choose different ways. It was also very important to demonstrate that comprehension failures may occur and mistakes may be made during application of the strategies, and to tell the students that such mistakes can be very productive because they provide information about the thinking process and may lead to improvements in strategy use.

Table 3. Assignment of tasks to children and trainer in the three experimental conditions.

		Experimental Condition		
Tasks		Reciprocal Teaching	Monitor	Student
- apply strategy to text		"Student" Child	"Student" Child	"Student" Child
- select strategy to be applied - give feedback on content and application of the strategy - help and guide during correction of answer - "model" answer, if necessary	Monitor Function	"Teacher" Child	"Monitor" Child	Trainer
- assign someone to apply a strategy - decide when to move on to the next text passage - classroom management: maintain discipline	Organizational Function	"Teacher" Child	Trainer	Trainer
- explain and model strategies - teach conditional knowledge about the strategies - correct children when they make mistakes (with respect to content or the application of strategies)		Trainer	Trainer	Trainer

2.2.4.1. Reciprocal Teaching

This experimental condition was characterized by the children taking over the role of the "teacher" themselves and leading the other "students" in application of the strategies. The children took turns in adopting this role. Box 1 provides a transcript of part of a training session.

In this condition, one of the children was assigned to be the "teacher" for each paragraph. The "teacher" monitored the other students' responses and was responsible

for organizational matters as well. His tasks were thus as follows: First, a "student" had to be chosen to read the paragraph aloud. Before assigning someone to be "student" and to apply one of the four reading strategies, the "teacher" had to determine which strategies could and should be applied to the paragraph, and in which order this would occur. After the assigned student had carried out the strategy, it was the "teacher's" task to give feedback on both the content of the answer and correctness of application. If the "student" had not done a perfect job, the "teacher" had to inform the "student" of what could be improved and guide him during correction of the answer. If the "student" was not able to do this alone, the "teacher" was responsible for demonstrating the application of the strategy himself. Finally, the next "teacher" had to be appointed. In addition to all that, the "teacher" was responsible for maintaining discipline, or – if problems occurred – managing disturbances.

The task of the "students" was to apply the strategies or read a paragraph aloud. They were, of course, allowed to volunteer to apply a strategy or help another student.

The trainer was only active in the early sessions: he explained the reciprocal procedure and the tasks of the "student" and the "teacher" to the group. When the children carried out the strategies or adopted the "teacher" role on their own, the trainer made sure they had understood the strategy and the tasks the "teacher" had to perform. It was also important to guide the "teachers" to provide appropriate feedback for the "students" and to correct answers if necessary. The children should be aware that they could always ask the trainer for help; but that they should "take over" responsibility for giving feedback and correcting answers as well as organizing the sessions by themselves. The trainer should only be asked to help, or intervene by himself, when there were severe problems.

Box 1. Dialogue of a Reciprocal Teaching Group

In the following, the transcript of part of a training session in a reciprocal teaching group is reported and briefly commented. Of the 6 children in the group, only Ariane, Jonas and Peter were present that day. Ariane was appointed to be the first "teacher." It was the 7th training session for this group, so the children were quite familiar with the method. They are working on the *Polar Bears* text. The dialogue has been translated into English; students' names have been changed.

Heading

Trainer: We will start with the next text. (*Passes the texts to the children*) Okay, Ariane starts with being "teacher."

Ariane: Jonas, please read the heading.

Jonas: "*Polar Bears*"

Ariane: Peter, what do you think will be in the text? → **Predicting**

Peter: How they live and what they eat.

Ariane: And Jonas?

Jonas: Well, the text will probably describe how they live and eat, how they raise their cubs and the first paragraph will probably deal with their appearance and their way of living.

Ariane: Does everyone know what polar bears are? → **Clarifying**

Peter: Yes.

First paragraph: After the first paragraph of the text has been read aloud, the meaning of unknown words is clarified. Then one of the children tries to summarize the paragraph and the children think of questions that a teacher would ask. Before moving on to the next paragraph, the children make predictions.

Ariane: Peter, please read the text!

Peter: (*Reads aloud*) "The polar bear is one of the largest and strongest of meat-eating animals. He is found only in the Arctic. When grown, he weighs up to 1600 pounds. With his long, heavy body and narrow head he looks clumsy, but he can move very fast and easily. The polar bear can move easily in the water. Most animals have to paddle hard to stay afloat, but the polar bear can lie still in the water for a long time. Air spaces in his fur as well as an oily skin and a thick layer of fat help him to float."

Jonas: (*Turns to Ariane*) I have a question. What does "pound" mean? → **Clarifying**

Ariane: Where is that in the text?

Jonas: Here (*points to the phrase*): "1600 pounds".

Ariane: A pound is, I think, like a kilogram (*looks to the trainer to check*).

Jonas: Does that mean I can say: "1600 kilograms"?

Ariane: I think so.

Peter: I don't think so; otherwise it wouldn't have another name.

Jonas: Maybe it's in the math book. Right, that's in the math book. (*To Ariane*) Can I get the math book? (*Jonas gets up, gets the math book and looks for the right page.)*

Jonas: Pound is a measurement. Here it is. "Connections between measurements.
1 pound = 500 grams = 0.5 kilogram." Does everyone know that? One pound means 500 grams. We could convert this: 1600 pound are (*the boys speculate*)

Trainer: Stop, stop! Once again! (*Jonas looks at the trainer*) How many grams are one pound?

Jonas: 500 grams.

Trainer: And how much is that?

Jonas: Half a kilogram.

Trainer: So what do you have to do now?

Jonas: Divide by two. That's 800 kilograms.

Trainer: Correct.

Ariane: Are there any more unclear words?

Peter: Air spaces, what's that? Maybe spaces of air.

Jonas: It's something similar to the air bladders that fish have. That's also an air space. It's a part of the body where the bear stores air.

Ariane: Jonas, please summarize. → **Summarizing**

Jonas: Me? (*To Ariane*) May I underline the most important content? (*Ariane nods and Jonas underlines in his text*).

Jonas: (*To himself*) Why do I make it so complicated? (*Covers the text with another piece of paper*) Well, it's, it's about the polar bear, he is one of the largest and strongest of meat-eating animals, and he weighs 1600 pounds or 800 kilograms, and he can move easily in the water because he has air spaces and layers of fat, whereas it's difficult for the other animals, and well, that's it, and it was the most important content.

Ariane:	Good.
Jonas:	And? Just good?
Ariane:	Yes, that was good. (*Thinks for a while and looks at her bookmark*) Okay, Peter, ask a question. → **Questioning**
Peter:	A question? (*Ariane nods*)
Peter:	Why does the polar bear have to, well, why does the polar bear have to – Why is the polar bear able to move easily in the water? (*Jonas raises his hand*)
Jonas:	Because he has air spaces and a lot of fat – and light fur and layers of fat.
Ariane:	(*to Peter*) Was that correct?
Peter:	Yes, that was correct.
Ariane:	And now Jonas, another question.
Jonas:	How was his question?
Ariane:	Good. Medium. (*Looks at Jonas again*)
Jonas:	I'm thinking - because he already asked the question that, well that, told the most important content, well asked the question that answers the most important content of the text – I can only think of an easy question: How much does the polar bear weigh?
Peter:	1600 pounds or 800 kilograms.
Jonas:	(*To Peter:*) I would note that here, otherwise you forget it after a while (*notes the conversion formula for himself on the text.*)
Ariane:	That was really an easy question. (*Looks at Jonas*) And Jonas, what do you think will be in the next paragraph? → **Predicting**
Jonas:	What he looks like and what he eats.
Ariane:	And now the next paragraph. (*The boys look at her inquiringly as to who will be "teacher"*) Who wants to be teacher? (*Both boys raise their hands – Ariane counts them off on her fingers*) – Peter.

2.2.4.2. Monitor

In this condition, the children applied the reading strategies ("student" role), and for every strategy that was carried out, one child was assigned by the trainer to give feedback and to guide correction of the answers ("monitor" role). The trainer was responsible for organizing the session as a whole.

As in the other conditions, the "students" applied the strategies and read the text passages aloud when asked by the trainer to do so. But they also had to decide which of the four reading strategies was to be applied to the text and why.

For every strategy applied, one child was assigned to "monitor" the application of the strategy and to give the "student" feedback on this. The "monitor" was also responsible for making suggestions on what could be improved and how this could be done, and for guiding the "student" during the correction.

The trainer "moderated" the session: he appointed one of the students to read the text aloud and asked one of the children or the whole group which strategies could be meaningfully applied to the text. After the children had presented their ideas and given

reasons for them, he chose one of the children to apply a strategy and another child to "monitor" it. When both had completed their tasks, the next two children were appointed to execute the next strategy, and so on. It was also the trainer's task to terminate work on a paragraph and move on to the next one. He was also responsible for discipline. In the first session, he explained to the children what the "monitor" has to do and demonstrated this. Whenever a "monitor" did not fulfil the role requirements, he mentioned this and guided the "monitor" during the corrective feedback procedure.

Box 2. Dialogue of a Monitor Group

This is a transcript of part of the 9th training session in one of the monitor groups. Of the 4 children of the group, only Brady, Matt and Cecile were present that day. They are working on paragraph 6 of the text *Paper – A web for words*.

Trainer: OK, Brady's gonna read aloud. → **Reading the text aloud**

Brady: *(Reads aloud)* "Keller told a friend what he had observed, and he and his friend decided to see whether they could make paper from wood. After many tries, they finally found that they could change tough wood into soft pulp. They produced pulp by grinding pieces of wood and then soaking the wood in water. Out of the pulp they were able to make a rough paper."

Trainer: Which strategies can we apply, Cecile? → **Selecting a strategy**

Cecile: *(Looks at her bookmark and thinks)* Ask questions. *(Looks at the trainer)* → **Questioning**

Trainer: Ask questions? Then try to think of a smart question. The other children can think as well.

Matt: *(To the trainer)* I've got already one.

Trainer: Yes?! *(Looks at Cecile, Matt and Brady are patient and wait for the answer)*

Cecile: What do Keller and his friend produce paper with?

Matt: What with? *(unbelievingly)*

Trainer: What is the answer to the question, Brady?

Brady: *(Shaking his head)* It isn't in the text.

Trainer: *(To Matt)* Do you know the answer? *(Matt shakes his head) (To Cecile)* What is the answer to your question?

Cecile: With... with water and ... p ... pulp *(Looks at the text while giving the answer)*

Trainer: *(To the other children)* Is the answer correct?

Brady: I wouldn't say so.

Trainer: How do we have to change the question a bit for the answer to be correct?

Matt: How, how did they produce the paper?

Trainer: *(To Cecile)* How, or out of what did they produce paper? O.K.? Not what with. But other than that your question was correct. Was that a good question, Matt?

Matt: *(Thinks shortly)* Yes. If there had been an answer to it. But other than that the question was good.

Trainer: Yes, except for the question word. Other than that the question was good. *(To the others)* You should help her – tell her that she meant the right thing even if she didn't find the correct word. O.K.? What was your question, Matt?

Matt: Well, the same as I just said.

Trainer: OK. And what is the answer?

Matt: Out of wood, and water ...

Brady: *(Interrupts)* Out of tough wood

Matt:	Yes; out of tough wood, and then they took water and soaked it and then made pulp out of it.
Trainer:	Uh-hm. Are there other strategies we can apply to this text paragraph? *(To the whole group)*
Matt:	Summarizing. → **Summarizing**
Trainer:	Yes? *(inquiringly)* Do you want to try to summarize? *(Matt thinks shortly and then nods).* Well, then try it.
Matt:	Keller and his friend, ehm, tried, when they made the discovery, well to produce wood pulp out of wood, and to do that they took wood, water and then they put the wood into the water and after a while it became pulp, and then they dried the pulp.
Trainer:	Uh-hm. Cecile. What is your opinion of that summary?
Cecile:	Well, the summary was really good, and it was also correct what he said *(trainer nods)* and, well, I think it was good.
Trainer:	Hm, you think it was good. Is there anything that could be improved?
Cecile:	*(Looks at the text and thinks)* It was tough wood that they ground.
Trainer:	*(To Brady)* that they ground. *(To Cecile)* Is that important? That it was ground?
Cecile:	Well, actually yes ...
Matt:	*(Interrupts)* Yes, because otherwise they would have put big pieces of wood in the water, and that wouldn't have soaked up the water, or it would have taken years to soak it up ...
Trainer:	Yes, correct. *(To the group)* Is that clear? *(After a short pause, holding the bookmark up)* Is there another strategy you can think of?
Cecile:	Predicting. → **Predicting**
Trainer:	Yes, and what would be a prediction? *(All children think)*
Cecile:	Well, that they do this ... do this in a factory.
Brady:	Maybe that they built a machine to do it.

2.2.4.3. Student

The student condition most resembled the classic structure of classroom instruction: the students were only responsible for applying the reading strategies; all other tasks were performed by the trainer.

The trainer carried out the monitor function and was responsible for all organizational matters. He appointed one of the students to read the passage aloud, decided which strategy should be applied and assigned one of the students to do so. When the student had finished, he gave feedback and guided the correction of the answer. He was the one who terminated work on the current paragraph and maintained discipline.

Box 2. Dialogue of a Student Group

This is a transcript of part of a training session in the 6th training session of one of the student groups. There are 3 boys – Dan, Ronald and Nat – and 2 girls in the group – Angela and Diana. They are working on paragraph 5 of the text *Paper – A web for words*.

Trainer:	OK, we'll move on to the next text paragraph. Take your blue paper and (to Dan) only work with the paragraph everybody else is working on, this is paragraph 5.
Nat:	Oh, it's my turn to read aloud.

Trainer:	Yes, please start reading.
Nat:	(*Reads aloud*) "A German named Keller tried to locate more raw material to make into paper. Keller decided to watch the paper-making wasps. He saw the wasps biting little pieces of wood. With their strong jaws they chewed the pieces into a soft, mushy pulp. He saw them spreading this pulp over their nests. The pulp dried into a tough paper that served as a cover for the nests and was a cradle for the young wasps." (*Reading the text aloud, Nat had problems in pronouncing "served as*.")
Trainer:	Are there any words you don't know, or anything else you didn't understand? → **Clarifying**
Dan:	Served as.
Trainer:	O.K., served. Anything else?
Ronald:	Keller.
Nat:	That's a name.
Trainer:	Anything else? (*Short pause*) Served as. Alright. Does anybody have an idea?
Nat:	Functioned.
Trainer:	Functioned. How did you come up with that idea?
Diana:	It sounds similar.
Ronald:	It almost has the same meaning, right?
Nat:	Right. (*To the trainer*) It has the same meaning, right?
Trainer:	You look at me so inquiringly. What strategy are we applying right now? (*holding the bookmark up*)
Group:	Clarifying.
Trainer:	What is the picture for Clarifying?
Group:	The detective.
Trainer:	And what does the detective do?
Group:	He searches with a magnifying glass.
Trainer:	How do we want to proceed when clarifying? – Read again. First the sentence with the unclear word. That's what we will do now. Nat. (Nat reads the sentence aloud again) Do you have an idea?
Nat:	Functioned as.
Trainer:	(*Confirmatively*) Functioned as. Is there another word? (*Short pause*) It is also possible to say "was" twice. (to Nat). Please read the sentence aloud with the verb "was" twice. (*Nat reads the sentence*) Do you have an idea why they wrote "served as"?
Ronald:	Yes. Otherwise it sounds stupid.
Trainer:	Correct. You can also read the sentence with "functioned as." Alright. Anything else to clarify? No? Alright, now Ronald can summarize the text and Diana, please try to think of a question in the meantime. → **Summarizing**
Ronald:	Well, a German man named Keller lay down in a meadow, and he thought about how to produce paper, and he saw a wasp, how she bit a piece of wood with her strong jaws. And when the wasp returned to the tree, and made it into a soft pulp she put it around ...
Nat:	(*Interrupts*) ... and it became
Ronald:	I see, and it became solid.
Nat:	And it functioned like a cradle for the babies ...
Ronald:	That's not important.
Trainer:	OK. What do you think of your summary?
Ronald:	Too long.
Trainer.	Too long. Anything else?
Dan:	Too much read from the text.
Nat:	Not fluently.

Trainer: No, he did not read from the text. And it was fluent. I think it was a good summary. He tried to formulate it in his own words and did not look into the text very often. But it was a bit too long. Alright? Well, it was good but a bit too long. (*To Ronald*) Please remember: was everything really important?! For instance that his name was Keller. Was that important? (*The group discusses*) Not necessarily. See, first there were Egyptians, then it was Chinese people and after that there was a Frenchman.

Nat: Now there was a German.

Trainer: Right, now there was a German. O.K. Now I would like to have a question. Diana. → **Questioning**

Diana: How was the paper produced?

Trainer: Uh-Hm. And what is the answer?

Diana: Well, from wood.

Trainer: (*To the group*) Is that in the paragraph?

Group: No.

Trainer: No? What is in the paragraph?

Angela: That the wasps did it this way.

Dan: I see.

Trainer: What did you think? You just said "I see". Why did you say that?

Dan: How the wasps produced the pulp.

Trainer: O.K. How the wasps produced the pulp. Can you put that into a question?

Dan: But that is a question.

Trainer: (*Repeats*) How the wasps produced the pulp.

Dan: How did the wasps produce the pulp?

Trainer: (*Repeats confirmatively*) How did the wasps produce the pulp? See, he can do it alone when given a little time. Alright? Well, great! (*To Dan*) How did they do it?

Dan: Well, with their strong jaws they bit off pieces of wood, and then carried it to their nests, and they chewed it thoroughly for a long time, and then they spread it on the nest.

Trainer: And then?

Dan: Well, and then they could ... it was like a cradle for the babies.

Trainer: Great. That was a good question.

Ronald: May I?

Trainer: What? Ask a question? (*Ronald nods*) Sure.

Ronald: From ... What did the wasps build their nest for, well how did they produce it.

Trainer: Uh-hm. Please repeat your question.

Ronald: What material did the wasps make their nest from?

Trainer: Uh-hm. You formulated the question differently just now. Can you try it again? (*Ronald thinks*) You asked "What for?"

Ronald: I see. For what ... do the wasps need ... nest (*inquiring and doubtful*) ... no, the wood ...

Trainer: What is the answer to the question: What do the wasps need the wood for?

Angela: For their nest.

Dan: To rock the babies.

Trainer: (*To the group*) To protect the babies. For protection.

Dan: I see.

Trainer: (*To Ronald*) And your other question, can you please repeat it?

Ronald: What material did the wasps make their nest from?

Trainer: And?

Ronald: From wood.

Nat:	Who asked the best question?
Angela:	Well, Dan, undoubtedly.
Trainer:	(*To Ronald*) That was a good question. You thought about that for a while. (*Short pause*) Diana, maybe you were right with your question. But it doesn't say in this paragraph that paper was produced that way. It only says that the wasps did it that way. Alright?
Angela:	May I make a prediction? → **Predicting**
Trainer:	Of course.
Angela:	Maybe it goes on to say how he then produced the paper
Ronald:	Or how he propagated it.
Trainer:	What else? (*Noise in the corridor because another group has just terminated the session*) Can you remember that until tomorrow? Then goodbye for today.
Dan:	(*While clearing the material away*) I've got another prediction. How the wood was changed into pulp.

2.2.5. Instructions for trainers

To ensure that all trainers were well prepared for their work with the students, they took part in a 2-day preparation workshop. During those two days, they were informed about the goals and procedures of the study and received detailed instruction on how to work with the groups. The trainers were familiarized with all training and test materials. A special unit of the workshop was devoted to defining the role of the trainer and explaining the cooperative setting. All trainers were instructed to "create" a learning environment in which students would pay attention and listen to each other and feel accepted so that they were not afraid to give wrong answers or make mistakes. The latter point is very important for the training program: a lot can be learned from mistakes – by both the student trying to execute the strategy and the other group members. To create such a "safe" learning environment, it is essential that the students feel accepted and are not laughed at when giving wrong answers. This cooperative learning environment was to be established by all trainers, irrespective of the experimental condition.

Videos of the pilot study were used to demonstrate the instruction of reading strategies and the different experimental conditions. The trainers were, like the children, assigned to groups according to their time schedules. They were assigned to an experimental condition at random.

Trainers were provided with guidelines on how to teach the reading strategies in sessions 1 to 3. They also received a guide in which their experimental condition was described. Written instructions were handed out for the test in session 16.

To ensure that the groups all worked in the same way and that experimental conditions only differed with respect to the distribution of tasks, trainers were instructed

to organize classrooms in the following way: Two tables were grouped together and the children and the trainer sat in a circle around the tables. One side of the tables was left free if the session was video-taped. To prevent routine communicative styles developing, and children always sitting besides their friends, children were asked to sit in different places in every session. This also ensured that all of the children sat right in front of the camera from time to time.

Another major concern was cooperation between students. Cooperation was enforced in two of the experimental conditions – *reciprocal teaching* and *monitor* – by the interaction that the roles required; but in the *student* condition, tasks were teacher-directed. Trainers in the *student* condition were told to allow cooperation between students if the students initiated it, but not to encourage it. In all experimental conditions, trainers were required to establish one rule: students were not to raise their hands; this is not necessary in such small groups and may lead to competitive behaviour among students.

It was left up to each trainer to establish group rules together with the children at the beginning of the training program. This method was used by five trainers. Common rules were “We will listen to one another” and “We won’t laugh about others.”

2.3. Material

In the following, the tests administered will be described; they have already been listed in table 2. Tests that were only administered to children who participated in the training program are marked with an asterisk (*). Then, the training material will be introduced. The reading comprehension texts and questions, the worksheets for all four reading strategies, the post-training questionnaire, filled-out copies of the video analyses and one of the texts used for training (*Polar Bears*) are presented in the appendix.

2.3.1. Tests

In this section, the reading comprehension tests and measures of metamemory (Metamemory, Index of Reading Awareness) will be described, as will other tests that were administered to all children to control for a number of abilities. These include decoding speed, nonverbal and different kinds of verbal cognitive abilities (Kognitiver Fähigkeitstest and Knuspel’s Leseaufgaben) and verbal self-concept. Finally, the tests

administered only to children who participated in the training program – a post-training questionnaire and a measure of working memory capacity - will be described.

2.3.1.1. Knowledge measures

2.3.1.1.1. Reasoning abilities

Students' *reasoning abilities* were assessed with the *KFT* (Kognitiver Fähigkeitstest by (Heller & Perleth, 2000); one verbal and one nonverbal subtest were administered. The vocabulary subtest (subtest V1) was chosen for *verbal reasoning ability* – here, the children had to choose the word with the same or similar meaning as the target word from five alternatives. *Nonverbal cognitive ability* was measured with a figure classification test (subtest N1). Three or four figures that shared something in common (form, orientation or pattern) were presented in each row. Children had to select a figure that matched them from five alternatives. The number of points children received were converted into T-scores as described in the test manual.

Working memory was measured with the HAWIK (Hamburg-Wechsler Intelligenz-Test für Kinder (Wechsler, 1983) digit span. The experimenter read a list of digits aloud at a pace of one digit per second. The task for each child was to reproduce the digit string (forward: as read by the experimenter; backward: in reverse order). The child had two trials for each digit string length (starting with 2 digits); one of them had to be reproduced correctly to continue the test. The maximum number of digits correctly reproduced by the child is taken as an indicator of working memory capacity. This test was only administered to students who participated in the training program.

2.3.1.1.2. Reading and listening abilities

Knuspel's Leseaufgaben is a standardized reading test battery that was developed by (Marx, 1998) for 1st to 4th graders. Subtests 1, 3 and 4 were administered to all children. Subtest 1 provides a measure of *listening comprehension ability*. The experimenter read the task aloud only once, and children had to complete 14 items of a test form according to these instructions (for example, "Please underline the grade level that you were in last year"). Subtest 4 has a similar design, but here the instructions were printed in the test material, and children had to read and comprehend them themselves, resulting in a test score for *reading comprehension ability*. The last of the subtests administered is a measure of *decoding ability*; here, children read 40 pseudo-words and indicate which ones sound like German words. Children's scores were converted into T-scores as described in the test manual.

To provide a measure of *reading speed,* children were given a text and asked to read it as fast as possible. They had 5 minutes to read the text (a fable by Tolstoy). To prevent them from cheating, there were 12 brackets in the text, each containing 3 words, only one of which fitted the sentence. Subjects were required to underline the word that belonged in the sentence whenever there was a bracket in the text. This allowed the number of words read to be corrected for comprehension failures and a comprehension-corrected reading speed measure to be computed.

2.3.1.1.3. Specific strategy knowledge about Summarizing and Clarifying

Students' declarative knowledge about *Summarizing* was assessed by asking them to "Please note the three features of a good summary." The features that were each awarded one point are (1) contains the most important content and/or no details (2) is shorter than the text and (3) the text is paraphrased.

Students knowledge of *Clarifying* methods was tapped using the following instruction: "Sometimes, when reading a text, it happens that one does not know the meaning of a word or does not understand a sentence. What can be done to understand the text better and discover the meaning of words?" Students were given half a page to write down everything they could think of. Their answers were then classified as either external methods (asking other people for help or using resources like dictionaries or reference books) or text-related methods, that is strategies that use the text itself to discover the meaning (re-reading the word, the sentence or the paragraph again and searching for cues in the text). Each strategy generated was awarded one point.

2.3.1.1.4. Relational strategy knowledge

Relational strategy knowledge was measured by asking students to rank the utility of five different strategies when the task is to comprehend and remember the content of a text. Judgments were made by assigning a grade from 1 (excellent) to 6 (very poor) to each of the five strategies. Each student's ranking was compared with that of experts (professors of educational psychology and teachers). This test, developed by (Schneider & Schlagmüller, 2002), is one of the 6 scenarios of the metamemory test used in the PISA 2000 study for 9th graders. For every comparison of two strategies, 0 to 2 points were awarded: 2 points if the better alternative received a better grade, 1 point if the alternative strategies received equal grades, and 0 points if the student gave the worse alternative a better grade. The score used for analyses was the percentage of alternatives in correct order.

2.3.1.1.5. Conditional strategy knowledge and planning knowledge

In addition, two subscales of the Index of Reading Awareness Questionnaire (Jacobs & Paris, 1987) were administered: *conditional knowledge* and *planning knowledge*, with five items each. For every item, a comprehension-relevant situation was described and three alternative strategies were listed, from which students were required to select the best one. 2 points were awarded for selecting the best alternative; 1 point for the second best alternative, and 0 points for the worst alternative. A maximum of 10 points could thus be scored.

2.3.1.2. Performance measures

2.3.1.2.1. Reading comprehension

At both pre- and post-test, students' reading comprehension abilities were tested. Materials were two expository texts with four comprehension questions each. Table 4 provides information about the texts that were used.

The reading comprehension tests were administered in the following way: First, students were asked to put the booklet away and were given the text with the instruction "Please read the text carefully. Take your time and make sure you understand the text. Please indicate when you are done by raising your hand." As each child indicated that he was finished, the experimenter went over and asked whether he had understood the text and whether her wanted to read it again. Students who wanted to re-read the text were allowed to do so. When the student indicated that he was done reading or re-reading, the text was taken away. Students then continued with answering questions about the text in the booklet.

Four open-ended comprehension questions about the text had to be answered. The questions did not tap single pieces of information about the text, but were designed to cover main ideas of the text and evoke longer, complex answers. For example, instead of asking "How much does the polar bear weigh?" (answer: 1600 pounds), one of the questions would be "Which part of the body does the polar bear have a problem with and why is this the case?" (answer: his black nose, because it is the only part of the body that is visible in the snow). 4 to 5 lines were printed in the booklet for the answer to each question. Students were instructed to read the questions carefully and to write everything down that they could remember.

Table 4. Texts used to assess reading comprehension.

The text marked with * was only administered to children in training groups in session 16.

Time	Reading comprehension texts	N of words	N of paragraphs	N of questions	Source
pre	The Mandan Indians (Mandan-Indianer)	308	6	4	Rich Zelmanowicz (1989)
post	A Brown Wave of Ants (Eine braune Welle von Ameisen)	545	7	4	Brady (1990)
post	Goose Pilot * (Gänseflieger)	879	13	10	("Magazin, Nr. 51," 1999)

Students' answers were categorized by two trained raters (teacher training students). Rating schemes that coded all possible meaningful answers were used to score the answers in two different ways: in terms of their quantity and quality. The tests of children in school 3 (i.e., 29 % of the tests) were double-rated. Inter-rater reliability was very high, with intra-class correlations above .90 (see Appendix C1).

The first rating scheme was designed to measure both the level of understanding of the text and the number of ideas students wrote down. To this end, all meaningful answers were given credit and sorted into three levels of comprehension. Level 1 answers contain only single and rather detailed pieces of information that have a 1:1 correspondence to the original text and were often formulated in exactly the same way as the text. An example of an answer for the question "Where did Joe live and how did he earn his money?" to the text *A brown Wave of Ants* would be "Joe lived in a beautiful valley." Answers at this level correspond to the *Knowledge* category in Bloom's taxonomy of educational objectives (Bloom, Englehart, Furst, Hill, & Krathwohi, 1956), which is the lowest level of that taxonomy. Category level 2 was designed to cover content that is more directed toward comprehension of the content of text and drawing of easy inferences. Children's answers indicate that they understood more central ideas of the text by reproducing them. An example would be "Joe is a farmer." This was not stated explicitly in the text, but is an inference based on the facts that "Joe lives in the country" and "he plants coffee plants". Answers at this level correspond to the second and fourth level of Bloom's educational objectives: *Comprehension* and *Analysis*. Answers were classified as belonging to the highest level, level 3, if they were generalizations of textual information or were judged to be very central, well-elaborated ideas of the text. An example of a generalization and complex inference drawn by a child is that "the farmers were financially ruined because they had nothing more to sell." Bloom described content of this kind as *Synthesis* and *Evaluation*; the two highest educational objectives. From this analysis, four dependent

variables that can be considered rather quantitative measures were derived: one point for each of the three levels of comprehension and a summative score.

Analogous to the analysis of students' summaries, in which the content (quality) was of course more important than the sheer amount of writing produced, students' answers to the comprehension questions were also analyzed using a different procedure: a priori the raters examined the text and judged the importance of the information provided in the text. Important content (main ideas) and less important content (details) were differentiated. Both main ideas and details were then assessed again in terms of their relative importance - central ideas and less central ideas were distinguished, as were details and very unimportant details. Central ideas and very unimportant details were weighted double. To provide a measure of the quality of students' answers that was independent of the total number of ideas that they produced, the ratio of main ideas to details was computed. A ratio of more than 1 means that the children produced relatively more important than unimportant ideas. In contrast, when the ratio is less than 1, the unimportant content outweighed the important content.

2.3.1.2.2. Writing a summary

At post-test, all students were required to write a summary of a short text passage (*Popcorn*, 144 words). The children in the training groups also wrote a summary in their last training session (*Chewing Gum*, 140 words).

The content of the summaries was analyzed in the same way as the content of the reading comprehension texts. The information in the text was classified by the raters to be either important or non-important (main idea or detail) in two grades. Additionally, the number of words produced were counted, and it was recorded whether the students had generated a heading for their summaries and whether or not they had underlined text in the test booklet. Coherence of the text and linguistic style were rated on 5-point Likert scales. Raters were also asked to provide an overall judgment of the summary and to determine which grade "a teacher would give" (from 1-excellent to 6-fail).

2.3.1.2.3. Judgments of learning and performance predictions

Students' metacognitive awareness of their learning and performance in the comprehension tests was measured at both pre- and post-test. After students had read the text, they were told that their understanding would now be tested by questions on the text. The following items had to be answered on 10-point Likert scales ranging

from nothing/very poorly to everything/very well: "How well did you understand the text you just read?", "How much of the text do you still know?" and "How well will you answer the upcoming questions about the text?" The students then turned the page and answered the comprehension questions. When they had finished, they were again asked to indicate: "How much of the text did you still know?", "How well did you answer the questions about the text?" In this way, students' *judgments of learning* and *performance predictions* were tapped.

2.3.2. Post-training questionnaire

In the final training session, students were given 10 minutes to write a *summary* about a short text passage (*Chewing Gum*). After that, *text comprehension* was tested by first giving the children a longer text to read (*Goose Pilot*). When the children had finished reading and had re-read the text if they wanted to, the text was taken away and 10 open-ended comprehension questions were presented. It took the children 5 to 10 minutes to read the text and about 15 minutes to answer the comprehension questions.

Then children answered a number of questions (on 4-point Likert scales) about their motivation and involvement in the training. After a main component factor analysis, two scales were extracted; one measuring how much the children enjoyed the training (how much fun they had, whether it was useful or boring, and whether they tried their best) and another reflecting their enjoyment of working in a group. Two single items are also listed in table 5; these concern the utility of the skills learned during the training program and how exhausting they felt the program to be.

Table 5. Scales of the post-training questionnaire.

Scale	Items	M	SD	α	Sample item
Enjoyment of training	4	3.49	.48	.81	*How much fun did you have?*
Enjoyment of the group	2	3.28	.51	.55	*How much fun was it to work together with other children?*
Belief in utility of skills	1	3.31	.61	--	*Do you believe that the skills you learned will be helpful in school?*
Exhaustion through training	1	2.18	.79	--	*How exhausting was the training?*

In addition, the children produced a sociogram; that is, they stated how much they liked working together with each of the other children in the group. Each child rated every other child in the group (including the children who dropped out early) on 6-point Likert scales ranging from 1-liked him/her very much to 6-disliked him/her.

Finally, the children were given half a page to write down what they felt they had learned during the training program, and were given the chance to indicate what they liked and disliked most about the program.

2.3.3. Pilot studies

The main study was preceded by three pilot studies. The goal was to evaluate all materials and procedures in terms of their length and applicability, as well as their usefulness/necessity and statistical criteria. First, most of the tests were piloted by inviting students to the laboratory of the Max Planck Institute for Human Development in Berlin for one afternoon. In a separate study, the Index of Reading Awareness was tested in one elementary school. Finally, the training procedure and all experimental conditions were tested and recorded on video (also in the laboratory). All finalized test materials were administered to the children who participated in these training groups.

Materials. Participants in this study were 16 5th graders (mean age 10.92 years). They took part in a two-hour testing session in the laboratory, in which the following tests were administered: a test measuring relational strategy knowledge developed by (Schlagmüller, Visé, & Schneider, 2001) for grades 3 and 4, all verbal subtests and one nonverbal subtest (N2) of the KFT, and Knuspel's Leseaufgaben (complete test). Additionally, two tests measuring text comprehension with different answer formats were administered. After reading expository texts, students had to either answer 7 multiple-choice questions or 10 open-ended questions on the text. Half of the students were allowed to refer back to the text when answering the multiple-choice questions; the other half were not allowed access to the text at test-taking.

Results of this study were as follows: For the test of *relational strategy knowledge* aimed at 3rd and 4th graders, massive ceiling effects were observed. Therefore, it was decided to measure relational strategy knowledge with a test originally developed for 9th graders (Schneider & Schlagmüller, 2002). For the main study, it proved to be sufficient to administer only one of the three verbal KFT subtests (V1) to assess *verbal reasoning abilities*; this test showed the highest correlation with the total verbal ability score (r = .89). To make the measure of *nonverbal reasoning ability* more compatible with the verbal measure, and because of students' difficulties understanding the instructions of the KFT nonverbal subtest N2, subtest N1 was administered in the main study. The multiple choice *reading comprehension* tests were too easy for the students; many students answered all questions correctly. The administration format

(answering the questions with vs. without reference to the text) had no effect on this. The comprehension questions with an open-ended answer format differentiated between students; at least where questions that required a deeper understanding of the text and more complex answers were concerned. However, it took a very long time for students to generate their answers. In the main study, the number of questions on each comprehension text was reduced to four, and some questions were combined to elicit more detailed answers.

Index of Reading Awareness. The complete Index of Reading Awareness was administered to all 5th grade students in one elementary school. Then the scores for all four subtests and the summative IRA score were computed and correlated with students' grades in their native language course (mean grade and the grades for reading, orthography, oral language application and text production). The only significant correlations with mean grade in the native language course were found for the two subscales *conditional knowledge* and *planning knowledge* (with correlation coefficients of -.31 and -.36). Therefore, the other two subscales, *evaluation* and *regulation*, were not administered in the main study.

Procedure and experimental conditions. One group of students for each experimental condition participated in a pilot strategy training program lasting 8 sessions. Each of the three pilot training groups consisted of 2 boys and 2 girls. Prior to and after training, they were administered all tests that were also used in the main study. It was noted how long it took students to complete these tests. The procedure in the training groups was the same as in the main study: in the first three sessions, the strategies were instructed the same way in all groups; then, the strategies were practiced with longer expository texts with tasks being distributed according to the different experimental conditions. All sessions were recorded on video.

The experiences with these three training groups proved to be extremely useful for the main study. For example, it emerged that, to be able to apply all reading strategies meaningfully, the paragraphs that the texts were divided into should contain at least 3 or 4 sentences and share a common topic, and that in later stages of the training program, the paragraphs could even be longer and contain more information. It also became clear that the patterns of interactions changed dramatically when one child was missing from the group, especially in the *reciprocal* and *monitoring* conditions. In the main study, then, the groups consisted of up to 6 children. The video material of the

pilot training groups was also used to prepare the trainers and video-raters for their work.

2.3.4. Training material

The *Teacher's Guide* developed by (Palincsar et al., 1989) provided the framework for instruction. It was translated and slightly adapted for use with German students. A bookmark with pictures representing the four reading strategies and 2 to 3 short notes on each of the strategies was available for every student at all times throughout training sessions 4 to 15.

During training, a few tests were administered on the texts students had read during the training sessions (*Geese, Polar Bears, Superstition, Mountains of Fire*). Each of these tests consisted of four open-ended comprehension questions on the text and students usually worked on them for 15 minutes.

2.3.4.1. Worksheets

For each of the reading strategies, a worksheet was prepared. All worksheets were translated from (Brady, 1990) and contained examples of gradually increasing difficulty. The worksheet for *Questioning*, for example, started by giving students a question word ("Who") and a sentence it could be applied to ("In the middle ages only members of the King's family were allowed to own falcons"). The difficult examples contained more sentences providing a lot more information and students had to generate the whole question themselves. The worksheets were adapted and complemented by few more examples by the experimenter. Please see the appendix for complete material.

2.3.4.2. Texts

Most of the training and test materials were originally used by (Brady, 1990; Lonberger, 1989; Rich Zelmanowicz, 1989) in their dissertations. The texts were translated into German and, if necessary, adapted; some of them had already been used in the pilot study.

Texts of between 268 and 770 words in length (M = 505.55, SD = 156.26) were used for training and also to measure text comprehension. All of the texts were visibly divided into paragraphs that allowed for the rest of the text to be covered up (M = 6.91, SD = 2.34, with a range of 3 to 11 paragraphs in each text). Each of the paragraphs consisted of at least three sentences to ensure that a summary could be made and that the paragraph consisted of a meaningful unit of the text (e.g., for the *Birds in Dress*

Suits text, one paragraph was on physical appearance, the next on food and another on territory, etc.). The mean number of words per paragraph was 74.89 (*SD* = 10.91).

Table 6. Reading material: Texts used for training.

Order	Texts used for training	N of words	N of paragraphs	Source
1	Birds in Dress Suits (Vögel im Frack)	511	8	Brady, 1990
2	Paper – A Web for Words (Papier – Ein Gewebe-Netz für Wörter)	481	7	Brady, 1990
3 A	Geese (Gänse)	478	7	(Dossenbach, 1992)
3 B	Polar Bears (Eisbären)	770	8	Lonberger, 1989
4	How Till Eulenspiegel bought ground (Wie Till Eulenspiegel Erde kaufte)	268	3	(Kästner, 1970)
5 A	Migrant Birds (Zugvögel)	326	4	(Das mußt Du wissen, 1990)
5 B	About Dolphins (Delphine)	475	6	Brady, 1990
6	Superstition (Aberglaube)	467	6	Brady, 1990
7	Mountains of Fire (Berge des Feuers)	399	6	Brady, 1990
8	Swamplands (Sümpfe)	685	11	Lonberger, 1989
9	Silk (Seide)	701	10	(Treff-Schülerbuch 1993, 1992)
Order	**Tests: Summaries**	**N of words**	**N of paragraphs**	**Source**
post	Popcorn (Popcorn)	144	3	(Eroberer und Entdecker, 1990)
post	Chewing Gum (Kaugummi)	140	4	(Widmann, 1994)

2.4. Analysis of video-taped sessions

All sessions that were available on video were analyzed in two successive steps: first, the entire sessions were analyzed; this part of the analysis will be called *session analysis*. Then, in a second step, only those parts of the videos showing children practicing the reading strategies verbally were analyzed more closely; this will be termed *paragraph analysis*. The purpose of the *session analysis* was to record the time that was spent on different types of activities and to rate the session as a whole. *Paragraph analysis* served to investigate children's abilities to apply the reading strategies to text paragraphs and to determine whether the experimental conditions were implemented correctly.

Before describing in detail how the video material was rated, the source of the data is of interest; more specifically, the number of sessions that were video-taped and analyzed. Table 7 shows the sessions that were video-taped for each of the groups.

Table 7. Overview of video-taped sessions for each training group.

School	Condition	Trainer	Session[3]														
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	RT	3		▒	▒		■	■	■	■				▒	▒	■		
1	Monitor	6		■	■	■	■		■	■				■	▒	▒	■	
1	Student	1	■			■		■	■		■	■	■		▒	▒		
2	RT	7	■	■		■	■	■		■		■		■		■	■	
2	RT	2			■	■		■	■		■	■		■	■		■	
2	Monitor	5	■	■	■		■		■		■		■	■		■	■	
2	Monitor	1	■	■		■		■		■	■	▒	▒		■		■	
2	Student	4	■		■		■	■				■		■	■			
2	Student	9		■	■			■			■	■	■			■		
3	RT	1	■	■	■	■	■	■	■		▒	▒	■	■	■	■		
3	Monitor	7		■	■	■	■		■	■	■	■		▒	▒	■	■	■
3	Student	8	■		■	■	■	■	■		■		■	▒	▒	■	■	■
4	RT	1	■		■	■	■	■		■		■	▒	▒		■	■	■
4	RT	7	■	■	■	■		■	■	■	■	■	▒	▒		■	■	■
4	Student	10	■	■	■	■	■	■	■					■	■	■		■

The number of training sessions recorded on video ranged between 7 and 13 sessions per group, with a mean of 9.92 (*SD* = 1.68). Regular training sessions (total N = 110) lasted 55 minutes (*SD* = 4.75). 10 double sessions were video-taped; these were of 85 minutes' duration (*SD* = 16.41). There were no systematic differences between experimental conditions in terms of the number of sessions recorded on video or the duration of sessions. For each of the training groups, about 27 text paragraphs were recorded (SD = 8.57, range between 13 and 44). Every session and every paragraph that was available on video was analyzed.

Some of the scales constructed for the longitudinal BIJU study ("Bildungsverläufe und psychozoziale Entwicklung im Jugendalter") were used to construct items and scales for analyzing the present video material. The BIJU study was also carried out by the Max Planck Institute for Human Development, and was designed to investigate the quality of school education and its importance for acquisition of knowledge and identity formation. The scales used are task orientation, quality of answers, cooperation, discipline, regulation of conversation, pace of

[3] ■ - single session; ▒ - double session.

interaction, individual norm orientation and participation. All of the items were developed for video rating purposes (Dann, Diegritz, & Tosenbusch, 1999; Gruehn, 2000); some of them were modified for the present study. In addition, several items were constructed to provide measures of how well the experimental conditions and reading strategies were implemented, how much responsibility the trainer and the children had in determining the course of the sessions, and how successful the children were in carrying out the "teacher" and "monitor" roles. The rating schemes were tested with videos recorded during the pilot study (1 group with 8 training sessions for every experimental condition; completely video-taped). The video raters were also trained with pilot study videos.

Three raters were responsible for session analysis of the 141 training sessions recorded on video; 28 sessions were double-rated. The 408 text paragraphs that were identified in the session analyses were rated by three different raters. 87 of these paragraphs were analyzed twice. All of the raters were students; four of them were majoring in psychology, one in administrative science.

2.4.1. Session analysis

In the first step of analyzing the video material, entire sessions were the unit of analysis. First, a protocol of the session was made, documenting the type and the duration of all activities. Then, the session was rated according to a number of features reflecting aspects like working atmosphere and discipline; this included some scales that made it possible to determine whether the experimental conditions were implemented correctly.

2.4.1.1. Duration of different activities

The start and end times of each of the following activities were noted: (1) organizational matters, (2) summary of last session (repeating the strategies or content of the current text), (3) group work: children working together with the trainer on the chalkboard or flipchart, (4) reading a text paragraph aloud, (5) silent work with the text (underlining in the text, writing a summary or notes; all children writing at the same time), (6) repetition of the reading strategies, (7) verbal application of the strategies (one child at a time), (8) strategy instruction by the trainer, (9) role instruction for "teacher" and "monitor" students by the trainer, (10) trainer giving a sample answer, (11) taking tests (summarizing text paragraphs or answering comprehension questions about a text), (12) trainer dealing with discipline problems (only when the intervention interrupted the

course of the training), (13) group chatting (mostly private matters) and (14) group discussions (not text-related).

If possible, it was noted which text and text paragraph the group was working on, whether there were any discipline problems, whether the trainer intervened and whether there were external interruptions (for instance, a teacher coming into the room). For each of the activities, *working atmosphere* was rated on a 6-point Likert scale (corresponding to the German grading scale; 1-excellent to 6-very poor).

2.4.1.2. Global ratings

First, raters were asked to estimate the *percentage of time* the children and the trainer talked (in steps of 10 %). This was followed by ratings of *noise level*, *discipline*, *adequacy of trainer intervention* and *working atmosphere*. All of these features were rated on 6-point Likert scales, separately for oral work with text paragraphs, silent work with texts, test and organizational matters. It also was of interest how well the trainer implemented the experimental condition.

After finishing the detail ratings, the raters were asked to assess the session as a whole in terms of *global working atmosphere*, *global discipline* and *adequacy of trainer intervention* on 6-point Likert scales. They were instructed to take discipline, discipline management by the trainer, involvement of the children, external interruptions, impact of the day's events (sports day, extremely hot weather, etc.), quality of children's answers, noise level and compliance with social manners and rules of conversation into account when assessing the *global working atmosphere*. When judging *discipline*, they were instructed to pay no attention to noise level or other indicators of *working atmosphere*. For ratings of *adequacy of trainer intervention*, the experimental condition had to be taken into account.

2.4.1.3. Ratings of group work and trainer behavior

Items on the level of the session as a whole were rated on 4- or 5-point Likert scales ("is not true" to "true", sometimes including also "did not occur"). To test whether the item responses correspond to empirical patterns and form the intended scales, they were subjected to main component factor analyses with VARIMAX rotation (criterion eigenvalue > 1). The items pertaining to noise were analyzed separately for oral work on text paragraphs, silent work with texts, tests and organizational matters. The factor loading matrix is available in the appendix.

The results distinguished between scales that concern the behavior of the trainer and those that pertain to the work of the group, as shown in table 8. For each scale, the

number of items, internal consistency of the scale (Cronbach's α), scale mean, standard deviation and number of observed sessions is displayed along with a sample item. The last two columns present the number of sessions that were double-rated by independent raters and the intra-class correlation as a measure of absolute agreement between the raters.

For the work of the group, the factors *feedback* and *co-determination by the children* were extracted. *Feedback* reflects whether or not the children gave each other feedback on the application of strategies, and whether the trainer encouraged or helped them. The scale *co-determination* was designed to provide a measure of whether the children or the trainer decided who should read the text aloud or execute a strategy and when to proceed with the next text paragraph (organizational responsibility). The higher the scale value, the more the session was determined by the children.

Instruction, discipline management, clarity of rules and *effective enforcement of rules* characterize the behavior of the trainer. The *instruction* scale measures the extent to which the trainer asks the children for prior knowledge about reading strategies or explains strategies, demonstrates how they should be executed or asks one of the children to do so, asks for prior knowledge about a topic and asks children to evaluate their own answers. *Discipline management* reflects whether or not the trainer intervened when there were problems with discipline (by directing children's attention to the problem or intervening in another way). The other two scales both dealt with class rules. *Clarity of rules* was designed to measure whether or not the trainer established rules at the beginning of the program, explained consequences for non-compliant behavior or referred to these rules during the training program.

The trainer's management of these rules was judged by the *effective enforcement of rules* scale. The following aspects were taken into account: whether the trainer reacted to a disturbance or ignored it, whether or not he intervened in a way that did not interrupt the course of the session, and the overall extent to which he seemed to be "at the controls."

Table 8. Session analysis scales.
Scales that are marked with * are also available for paragraph analysis.

Scale	Items	M	SD	α	N	Sample item	ICC	N
Trainer								
Instruction *	4	2.58	.84	.68	143	*The trainer demonstrates aloud.*	.20	27
Discipline Management	6	1.43	.60	.69	143	*The trainer rebukes a child.*	.51	27
Clarity of rules *	3	1.05	1.29	.86	143	*The trainer refers to rules that were established.*	.42	27
Effective enforcement of rules	6	1.96	.33	.86	143	*The trainer ignores disturbances.*	.66	27
Group								
Feedback *	4	2.18	1.26	.86	142	*The children give each other feedback.*	.84	27
Co-determination by the children *	7	1.94	1.09	.95	143	*The children terminate work on the paragraph.*	.94	27
Noise – oral work on text paragraphs	9	1.41	.61	.94	143	*It is so noisy that it is barely possible to work.*	.85	27
Noise – silent work with texts	9	1.32	.58	.95	60	*Ambient noise audible.*	--	<10
Noise – tests	8	1.15	.29	.79	33	*It is silent.*	--	<10
Noise – organizational matters	9	1.71	.72	.92	24	*Almost every word is understandable.*	--	<10
Reciprocal: "Teacher" students								
Meeting role requirements *	2	3.21	.93	.78	55	*The "students" are given feedback.*	.71	10
Guidance to correct answer *	3	1.71	.75	.67	55	*The "teacher" recommends how to correct the answer.*	.84	10
Monitor: "Monitor" students								
Meeting role requirements *	2	3.61	.58	.60	36	*The "students" are given feedback.*	--	<10
Guidance to correct answer *	2	1.44	.58	.58	36	*The "monitors" recommend how to correct the answer.*	--	<10

The work of “teacher” and “monitor” children was also rated. It was determined how well the children met the *role requirements* and whether they *guided* the other children *to correct their answer*. The items had to be rated on the level of the whole group, meaning that more than one “teacher” child and a number of “monitor” students had to be taken into account simultaneously.

2.4.2. Paragraph analysis

In a second step, the text paragraphs that were identified by the session analysis were assessed more closely. Most of the time when working on text paragraphs, the children applied the reading strategies orally, one child at a time. Every strategy application was noted and rated. As in the session analysis, the work of the group and the behavior of the trainer were rated. Finally, every child was rated with respect to cognitive and social behavior. Table 9 shows the scales that were extracted and the level (group, trainer or individual children) to which they belong.

2.4.2.1. Strategy application

In a protocol for each paragraph, it was recorded which child executed each strategy. It was noted whether the child applied the strategy properly, whether he or she volunteered or was appointed to do so, and whether the child received help from his peers or the trainer. For the strategies *Summarizing*, *Questioning* and *Predicting*, raters gave a grade ranging from 1 (excellent) to 6 (not acceptable). Raters were given criteria developed a priori to determine the quality of strategy application. *Clarifying* was not rated because it is not possible to determine “how well” the strategy was applied – either something is unclear to a child or not. It was merely noted whether the child brought up the problem, and whether he or she tried to explain the word in question.

2.4.2.2. Global ratings

As in the session analysis, raters estimated the *percentage of time* the children and the trainer spent talking (in steps of 10 %). They were also required to judge *working atmosphere* and *discipline management by the trainer* on 6-point Likert scales.

2.4.2.3. Ratings of group work and trainer behavior

Parallel ratings of *instruction* and *clarity of rules* for trainer behavior, and two scales that characterize the work of the group - *feedback* and *co-determination by the children* - were also made.

Table 9. Paragraph analysis scales.
Scales that are marked with * are also available for paragraph analysis.

Scale	Items	M	SD	α	N	Sample item	ICC	N
Trainer								
Instruction *	9	1.09	.78	.90	408	*The trainer demonstrates aloud.*	.93	66
Clarity of rules *	3	.26	.70	.87	406	*The trainer refers to rules that were established.*	.94	66
Intervention	5	1.15	1.19	.87	407	*The trainer intervenes when discipline is low.*	.99	66
Time pressure	3	1.10	.40	.68	408	*The trainer pushes the children to work faster.*	.88	66
Pacing	3	1.97	.39	.54	408	*The children are given time to think before giving an answer.*	.46	66
Group								
Feedback *	4	1.69	.93	.68	408	*The children give each other feedback.*	.97	66
Co-determination by the children *	7	1.89	1.05	.92	408	*The children terminate work on the paragraph.*	.98	66
Implementation of Summarizing	3	2.75	1.20	.87	418	*If the summary was not good, somebody mentions it.*	.86	58
Implementation of Clarifying	2	3.39	.83	.46	309	*The children try to explain the meaning of unknown words.*	.90	43
Children								
Attentiveness	7	3.71	.38	.69	1665	*The child listens carefully.*	.83	386
Contribution	2	2.63	.89	.74	1543	*The child participates actively.*	.87	359
Feedback to the child	2	.53	.38	.57	858	*The child is praised for his answers.*	.96	207
Reciprocal: "Teacher" students								
Meeting role requirements *	2	2.62	1.10	.95	169	*The "student" is given feedback.*	.96	31
Guidance to correct answer *	2	1.51	.68	.62	170	*The "teacher" recommends how to correct the answer.*	.94	31
Monitor: "Monitor" students								
Meeting role requirements *	2	3.73	.59	.60	258	*The "student" is given feedback.*	.90	44
Guidance to correct answer *	2	1.06	.31	.58	259	*The "monitor" recommends how to correct the answer.*	.94	23

In addition, the trainer's *intervention* in case of discipline problems was judged. It was also observed whether the trainer mentioned the time during training and whether or not he tried to get the children to hurry up (*time pressure*). Independent of this, another time-related aspect was rated: pacing. The *pacing* scale measures the amount of time children were given to "think" before answering, whether the trainer waited for an answer or gave the task to another child, and whether he was concerned about the children really understanding the text or more about hurrying on to the next passage.

Two scales measured how well the strategies were understood by the children in the group and the response of the trainer. *Implementation of Summarizing* reflects whether feedback was given to the "student" in case of summaries of low quality. *Implementation of Clarifying* reflects whether the children tried to explain the meaning of unknown words to one another or whether this was done by the trainer immediately.

2.4.2.4. Ratings of children's behavior

There were also some items that were rated separately for each child in the group. The following factors were extracted: *attentiveness*, *contribution* and *feedback given to the child*. The extent to which a child was alert and listened carefully or was occupied with things other than the training (playing with things or chatting) is reflected by *attentiveness*. *Contribution* measures how actively the child participated in the training sessions by making comments, volunteering for tasks, etc. Whether or not the child was provided with *feedback* on his answers was also recorded. Additionally, the amount of time the child appeared to be "on task" was estimated. Raters were instructed to give their estimation as a percentage of total time in steps of 10 percent.

For children who were appointed to be either "teacher" or "monitor," it was rated whether they *met the role requirements* and how well they *guided the "students" to correct their answers*.

2.4.3. Interpretation of scale characteristics

A variety of measures were constructed to assess aspects of group work, behavior of trainers and individual children, implementation of experimental conditions and reading strategies and success in fulfilling role requirements. Some of these aspects were measured with global ratings; others were measured with several items derived from factor analyses. To complicate matters even further, the video tapes were analyzed in two consecutive steps. First, whole sessions constituted the units of analysis. In the second step of analysis, text paragraphs formed the units of analysis.

Because of the large amount of video material, it was analyzed by several raters. All this raises two questions concerning these measures: How reliable were the different aspects measured? Were they defined well enough so that the raters agreed in their judgments?

The first question can be answered by looking at the internal consistency of aspects measured with more than one item (Cronbach's α of the scales), as shown in tables 8 and 9. The internal consistencies of most of the scales were good or very good with coefficients above .70. In all but one of the other cases, the scales consist of two items only, meaning that the coefficients of between .50 and .70 are still satisfactory. Only the implementation of Clarifying was not reliable – this scale consists of two items only, measuring whether the children tried to resolve the meaning of unclear words themselves or whether the trainer did so immediately without really giving them a chance.

To measure whether the different raters came to similar judgments, the intra-class correlations (ICC) of all scales and all single-item measures such as *working atmosphere*, *adequacy of trainer intervention*, *noise*, *discipline* and estimates of the *percentage of time children talked* were computed. As a measure of absolute agreement between raters, the ICC indicates the amount of variance in the true feature that can be explained by the judgment of one rater or by the mean judgments of several raters. Therefore, it can serve as an indirect measure of reliability for the single-item measures. Absolute agreement between raters with respect to the scales derived from session analysis (see table 8) is acceptable to good in most cases. However, for *instruction* and *clarity of rules*, the coefficients (ICCs of .20 and .42 respectively) reveal that the raters had problems with the items. Both of these scales are also available for paragraph analysis; here, the coefficients were considerably higher, with ICCs of .93 and .94. One reason for this may be the larger number of double-ratings in paragraph analysis (66 vs. 27). All but one of the scales in the paragraph analysis had ICCs above .85; this means that these features were very well-defined so that the raters came to very similar or identical judgments. The exception is *pacing*, with a coefficient of .46. It is also notable that the scales designed to measure the behavior of "teacher" students, *meeting role requirements* and *guidance to correct answer*, not only had acceptable internal consistencies, but also good ICCs.

For single-item measures, only absolute agreement between raters is available. Table 10 shows item characteristics of these measures. All in all, the coefficients are

acceptable to good. In most cases, intra-class correlations are larger in paragraph analysis, but the opposite is the case for adequacy of trainer intervention. One possible explanation for this could be that when raters had access to the broader context of the session as a whole and knew what happened in the paragraphs or testing in the run-up to a discipline problem, they were able to provide a more accurate judgment.

Table 10. Item characteristics of single-item-measures.

Item	Value range	N	M	SD	ICC	N_{ICC}
Session analysis						
working atmosphere	1 – 6	143	1.59	0.81	.56	27
noise	1 – 6	143	1.62	0.83	.70	27
discipline	1 – 6	143	1.96	1.02	.66	27
adequacy of trainer intervention	1 – 6	143	1.74	1.24	.75	27
percentage of time children talked	0 – 100%	141	68.97	14.10	.71	25
Paragraph analysis						
working atmosphere	1 – 6	408	1.85	0.98	.80	66
discipline	1 – 6	408	1.85	0.98	.80	66
adequacy of trainer intervention	1 – 6	408	1.49	0.83	.66	66
percentage of time children talked	0 – 100%	408	64.27	16.97	.93	66

To summarize this interpretation of the scale and item characteristics, I conclude that measurement of the desired aspects was reliable and that the aspects were well defined, so that different raters interpreted them similarly. When scales are available for both levels of video analysis, the observations from paragraph analysis should be used because they are larger in number (*instruction*, *clarity of rules*, *feedback* and *co-determination by the children*). Only *adequacy of trainer intervention* should be viewed in the context of whole session.

3. RESULTS

I will start by checking whether the treatment was implemented correctly and whether the experimental conditions worked as specified. These analyses will provide arguments in support of the decision to exclude three of the fifteen groups of children from all further analyses.

In the second part of this section, I will examine the effects that the training in reading strategies had on measures of knowledge and performance. This is done by comparing the children who participated in the training program in *reciprocal* and *monitor* conditions with the *control* children, who were spent the afternoons occupied with their regular activities: doing homework, playing or participating in sporting activities or art circles.

Next, the three experimental conditions will be compared. Results of these analyses will provide additional evidence to support the hypothesis that metacognition is mainly promoted by the task of providing feedback to other children. Also, the effects of responsibility for organizational tasks can be examined. Analyses of all dependent variables that are also available for the training effect will be reported. Additional analyses that are not possible for the comparisons incorporating *control* children will also be performed, namely, analyses of video data. Also, children's impression of the training will be explored: judgments of usefulness of the training, fun and effort, relationships to the other children, etc.

In the third section, motivational issues will be addressed by analyzing the answers that the children provided in a questionnaire administered in the last training session. The questions children were asked concerned their enjoyment of the training program, their perception of working in a small group, the knowledge they acquired in the training program, what they liked or disliked most, and how well they got along with the other children in the group.

Last, the performance and knowledge of children of the *student* condition was compared with that of *control* children.

An alpha level of $< .05$ was adopted for statistical tests. In all figures, error bars reflect the 95 percent confidence interval based on within-subjects mean squared errors. For bar charts, error bars are only depicted in one direction.

3.1. Treatment implementation

To be able to answer the research questions, it is essential to ensure that the experimental conditions worked as intended (treatment implementation check). This is necessitated by the rather unusual circumstances under which training took place (voluntary training after school, groups of pupils from different classes) and the fact that most of the trainers were teacher training students or psychology students. The preconditions for learning within each group (working atmosphere, noise and discipline management) will also be analyzed. To check treatment implementation, experimental conditions will be compared with respect to variables that should differ according to the definition of conditions and instructions given to the trainers: involvement of the children in designing the sessions, feedback given by the children and the amount of time children talked.

3.1.1. Randomization of subjects

Although it was emphasized that the reading strategy training program is useful for all children and that it was not designed to improve decoding ability (which seemed to be particularly highly valued by parents), it was obvious that many of the children who volunteered to participate had only poor verbal skills. One of the reasons for this is that they were advised to take part in the program by their teachers and parents. Therefore, the number of children with poor verbal skills was disproportionately large in the training group, and there were only few children with very good skills. In the *control* group, it was the other way around. In table 11, the cognitive abilities and pre-test measures of knowledge and performance of the children who participated in the training and the *control* group are displayed. Significant differences between the two groups, derived with analyses of variance, are indicated with stars.

Although the two groups did not differ in terms of age, nonverbal reasoning ability (KFT) or listening comprehension scores, significant differences in favor of the *control* group were found for *verbal reasoning ability* (KFT subtest measuring vocabulary knowledge with $MSe = 95.28$, $F(1,139) = 3.68$ and $p < .05$), *decoding ability* ($MSe = 107.58$, $F(1,135) = 4.69$ and $p = .03$), *reading comprehension* ($MSe = 85.35$, $F(1,135) = 7.00$ and $p < .01$) and *decoding speed* ($MSe = 45152.38$, $F(1,139) = 22.65$ and $p < .001$). In line with this, the *control* children received better *grades in the native language course* at school ($MSe = .68$, $F(1,143) = 9.97$ and $p < .01$).

Table 11. Systematic differences in cognitive abilities and pre-test measures of participants in control and training groups.

			Control (N = 86)		Training (N = 55)	
			M	SD	M	SD
Student Characteristics	Age		11.44	.49	11.49	.58
	KFT nonverbal		52.23	10.19	51.39	6.38
	KFT verbal (vocabulary)		* * 47.91	10.79	44.67	7.88
	Knuspel Listening comprehension		49.96	10.67	47.96	10.74
	Knuspel Decoding ability	*	51.52	11.04	47.59	49.43
	Knuspel Reading comprehension	**	53.70	9.10	49.43	9.44
	Reading speed	**	637.78	228.00	463.20	185.48
	Grade in German	**	2.04	.85	2.49	.79
Pre-Test Measures	Relational strategy knowledge		.69	.23	.64	.22
	Conditional knowledge (IRA)	*	6.47	1.61	5.80	1.77
	Planning knowledge (IRA)		6.14	1.62	6.09	1.55
	Reading comprehension					
	- main ideas	**	11.44	6.41	7.87	4.84
	- details		3.01	2.69	2.22	2.69
	- ratio main ideas / details		4.88	4.17	4.50	3.46
	- total number of points	**	10.31	5.47	7.54	4.41

The two groups also differed in their performance at pre-test: the *control* children outperformed the training children prior to the training program in some, but not all, of the dependent measures: this was the case for the *number of main ideas* that the children produced in the *reading comprehension* test and the *total number of points* awarded for their answers ($MSe = 34.20$, $F(1,139) = 12.50$ and $p < .001$ and $MSe = 25.87$, $F(1,139) = 10.01$ and $p < .01$, respectively). *Control* children also scored significantly higher on the *conditional knowledge* subscale of the Index of Reading Awareness ($MSe = 2.79$, $F(1,139) = 5.31$ and $p = .02$).

To get an impression how important and meaningful these differences are, another comparison may be useful: that between boys and girls. Gender differences in reading abilities are regularly found and often described in the literature (see Stanat & Kunter). More and larger differences were observed between the children who participated in the training and those who served as *control* children than between boys and girls. For example, *control* children could read about 175 words more in 5 minutes (.76 SD) than children who participated in the training program; girls read on average about 55 words more than boys (.24 SD). Apart from *decoding speed*, girls only

* $p \leq .05$; ** $p \leq .01$

outperformed boys in *listening comprehension* and the *grades* they received *in their native language course*.

Table 12. Intercorrelations of cognitive abilities.

	Listening compre-hension	Decoding ability	Reading compre-hension	Grade in native language	Reading speed	Mean verbal ability
KFT verbal (vocabulary)	**.41****	**.45****	**.42****	-.54**	.54**	**.75****
Knuspel Listening comprehension		**.36****	**.40****	-.51**	.33**	**.73****
Knuspel Decoding ability			**.57****	-.48**	.35**	**.79****
Knuspel Reading comprehension				-.47**	.41**	**.78****
Grade in native language					-.48**	-.66**
Reading speed						.52**

It is very likely these large and systematic differences in the children's level of cognitive abilities and scores prior to training had an effect on their learning and performance during the training and on their test scores afterwards. A *mean verbal ability score* was computed, because *control* and training children differed significantly in most of the measures concerning verbal abilities. This score will be used to *control* for level differences in further analyses. It consists of the KFT verbal subtest that measures *vocabulary knowledge* and the three subscales of Knuspel's Leseaufgaben: *listening comprehension*, *decoding ability* and *reading comprehension*. Like all of these measures, the *mean verbal ability* score possesses characteristics of a T-scale (M = 50, SD = 10); the mean is 49.35 with a standard deviation of 7.65. Children who participated in the training program scored significantly lower in *mean verbal ability* than *control* children (*MSe* = 56.32, $F(1,139) = 6.39$ and $p = .01$); the former group had a mean score of 47.35 (SD = 6.78) whereas the mean for the latter group was 50.63 (SD = 7.93). Table 12 shows that the aspects of verbal abilities that form the mean verbal ability score are highly intercorrelated (bold letters). The correlations of the various measures of verbal abilities with *grade in native language*, *reading speed* and the *mean verbal ability* score are also displayed. The better the children's verbal abilities, the better grades they receive in their native language course (the correlations are negative because in Germany the best grade is 1 and the worst is 6).

To illustrate the scope and the relevance of these differences in the level of verbal cognitive abilities, the two groups of children are broken down into quartiles according to their scores in reading speed and mean verbal ability. As shown in table 13, only 5.5 percent of the children in the training group belong to the top quartile with respect to *mean verbal abilities*. In contrast, more than one third of the *control*

children belong to that ability group. At all other performance levels, training group children are over-represented. The picture for children's *reading speed* is as follows: children with very low reading speed are highly over-represented in the training group (41.8% of training group children and only 14% of *control* group children) whereas in the top quartile it is the other way around (9.1% of training group children and 34.9% of *control* group children).

Table 13. Absolute and relative number of control and training children in ability groups.

	Quartile							
	IV		III		II		I	
	Control	Training	Control	Training	Control	Training	Control	Training
Mean verbal ability	19 22.1%	16 29.1%	18 20.9%	18 32.7%	17 19.8%	18 32.7%	32 37.2%	3 5.5%
Reading speed	12 14.0%	23 41.8%	23 26.7%	13 23.6%	21 24.4%	14 25.5%	30 34.9%	5 9.1%

3.1.2. Global preconditions for learning in the training groups

Early in the training period, it became obvious after randomly watching videos of all groups that three of the fifteen trainers (namely trainers 2, 3 and 4) in the first two schools had severe difficulties – the children were very noisy and were not engaged in the training program most of the time, and the trainers did not seem to be able and/or try to intervene in ways that were efficient to restoring a conducive learning environment. So as not to rely on these subjective impressions alone, video analysis will be used to provide data about the working conditions. Data from the content video analysis will be used because these data provide the most accurate assessments of the conditions actually present when the children were working with the texts and practicing the reading strategies. *Global working atmosphere* and *discipline* will be taken as indicators. These were globally rated for each paragraph on 6-point grade-equivalent Likert scales (from 1-excellent to 6-very poor).

The mean rating for *working atmosphere* at the level of text paragraphs (N = 408) for all groups was 1.80 (SD = .95); this means that the video raters judged working conditions overall to be good. The three critical groups received the lowest ratings for working atmosphere (trainer 2: M = 2.21, SD = 1.32; trainer 3: M = 2.67, SD = 1.22; trainer 4: M = 2.61, SD = .98). Low grades, that is ratings below 3, were given in 17.8, 21.7 and 10.0 percent, respectively, of paragraphs observed for these groups.

Similar results emerged for *discipline*: discipline was good to moderate (for all groups M = 1.85, SD = .98). The groups led by trainers 2, 3 and 4 received the lowest ratings for discipline (trainer 2: M = 2.38, SD = 1.39; trainer 3: M = 2.91, SD = 1.08; trainer 4: M = 2.95, SD = 1.03). All three groups were given low grades in more than 20 percent of the sessions observed.

So far, the video raters confirmed the impression that I gained when watching the videos for the purpose of giving feedback to trainers. But do poor discipline and poor working atmosphere alone justify excluding these groups? There are not only children present in the groups, but also the trainer, who is responsible and accountable for managing the situation and, if necessary, for restoring a conducive learning environment. So the question is: how did the trainers behave in critical situations? The variable that can be used to address this question is *adequacy of trainer intervention*, which was rated along with *discipline* and *working atmosphere*. Raters were instructed not to indicate the kind of intervention (nonverbal or verbal, friendly or unfriendly etc.), but instead take the discipline problems into account and to provide a measure of whether the trainer tried to intervene in a manner appropriate to the problem encountered. For instance, if a child is playing with materials or talking to another child, it often is sufficient to intervene nonverbally by looking at the child, whereas if children are hitting one another, the trainer has to address them directly, tell them to stop and make it clear to the whole group this behavior is unacceptable.

The mean ratings for *adequacy of trainer intervention* were generally very high - with a mean of 1.49 (SD = .84). The interventions of trainers 2, 3 and 4 were judged to be the worst of all trainers (means of 1.93, 2.74 and 2.26, respectively). To examine how trainers intervened when discipline problems arose, the two variables were cross-tabulated, focusing on the ratings for paragraphs when discipline was bad (grades 4, 5 and 6). All of the other trainers received good grades for their interventions in these instances. This was not the case for trainers 2, 3 and 4: they did not intervene in ways that would have made it possible to restore discipline. More specifically, in 4 of the 9 paragraphs with bad discipline, trainer 2 received a grade below 3 for his intervention; the same applied to trainer 3 in 5 out of 7 paragraphs, and trainer 4 received bad grades for his interventions in all 5 paragraphs with low discipline ratings.

In summary, the ratings of *working atmosphere*, *discipline* and *adequacy of trainer intervention* indicated that, in three of the fifteen groups (those led by trainers 2, 3 and 4), learning conditions were inadequate for at least 20 percent of the paragraphs

observed. Under such conditions, efficient learning was not possible and the children in those groups were clearly at a disadvantage in comparison to those in the other twelve groups.

The group led by trainer 4 had another disadvantage: one of the four children in the original group dropped out early and another became very ill and was absent for 2½ weeks. This left only two children in this group for most of the training sessions and produced patterns of interaction that differed considerably from all other training groups.

It was decided not to use the data of these three experimental groups in analyses because it could not be ensured that the learning environment was set up well enough to provide all children an equal opportunity to concentrate on learning and acquire new knowledge. Because those problems occurred early in data collection, it was possible to replace these groups. The sample was enlarged by another school (school 4) and three groups were re-sampled. Thus, for all following analyses, data are available for twelve training groups (four in each experimental condition).

3.1.3. Implementation of experimental conditions

Was the variation between the different experimental conditions successfully implemented? Addressing this question is not only interesting, but also necessary in order to answer the research questions. Three variables are of special interest when examining whether or not the three experimental conditions worked as intended: (1) the *percentage of time the children talked*, (2) the *co-determination of sessions by the children* and (3) the *amount of feedback given by the children*. These variables can be used as criteria for treatment implementation because the experimental conditions were designed to differ in precisely these respects.

In the *reciprocal teaching* condition, children are supposed take on increasing responsibility for determining the session over the course of the training program; the trainer is required to withdraw and let the children "take over." Therefore, the children should also be the ones talking most of the time and giving feedback to one another. The *monitor* condition should also be characterized by children providing the feedback, but by the trainer organizing the session and determining which tasks are done by whom – resulting in less determination by children and less speaking time for them relative to *reciprocal* groups. The third condition – *student* – was set up to be completely

determined by the trainer, who also provides the feedback. As such, he should talk more than the trainer of a *reciprocal* group.

In Table 14, video analysis data on these variables are displayed. In addition to means and standard deviations, the number of cases is given; this helps to explain differences in the significance of results between session and paragraph analyses.

3.1.3.1. Amount of time children talked

The *amount of time the children talked* (paragraph analysis) was computed for the beginning (sessions 1 to 4), middle (sessions 5 to 9) and end (sessions 10 to 15) of the training program and subjected to an analysis of variance with condition(3) and time(3) as between-subject factors. To test whether the *reciprocal* and *monitor* conditions differ from the *student* condition in this respect, the contrast specification for the condition factor (later referred to as *setting A*) was the following: (1) *reciprocal* and *monitor* conditions versus *student* condition and (2) *reciprocal* condition versus *monitor* condition. Orthogonal contrasts for time contrasted (1) the first four sessions versus all later sessions and (2) sessions in the middle of training program (5 to 9) versus late sessions (10 to 15).

Table 14. Implementation of experimental conditions – amount of time children talked, co-determination and feedback given by children.

		RT		Monitor		Student	
	N	M	SD	M	SD	M	SD
Children talking - % of time beginning	36	56.92	11.82	56.67	13.71	68.18	12.51
Children talking - % of time middle	39	75.36	12.16	66.43	11.51	67.50	15.45
Children talking - % of time end	37	82.92	9.16	67.73	8.17	73.64	9.97
Co-determination by the children (SA*)	91	2.78	.83	.88	.19	.94	.31
Co-determination by the children (PA*)	338	3.03	.71	1.01	.28	1.24	.36
Feedback given by children (SA)	90	3.41	.84	3.25	.67	1.16	.38
Feedback given by children (PA)	338	2.04	.58	2.57	.82	.92	.34

There were significant main effects for both the first and the second condition contrast and the first time contrast (*MSe* = 137.33, for condition(1) $F(1,103) = 4.35$, $p = .036$, for condition(2) $F(1,103) = 4.52$, $p = .039$ and for time(1) $F(1,103) = 24.54$, $p < .001$). The main effects of time(1) and condition(1) need to be interpreted in the light of the highly reliable condition(1) by time(1) interaction $F(1,103) = 9.81$, $p = .002$). In the early sessions, there was no difference in children's speaking time

* SA – Session analysis; observed units: entire sessions. PA – Paragraph analysis; observed units: paragraphs.

between experimental conditions. During the course of the training program, the amount of time students in the *reciprocal* condition talked increased more than in the other two conditions. Students' speaking time increased again from the middle to late sessions, but this was only the case in the *reciprocal* and the *student* condition and is reflected in the significant time(2) main effect ($F(1,103) = 3.27$ and $p = .074$).

3.1.3.2. Co-determination of sessions by children

The children's co-determination of sessions, assessed in both session and paragraph analyses, was subjected to an analysis of variance. Because it is assumed that the *reciprocal teaching* condition differs from the other conditions in this respect, different orthogonal contrasts were specified: first the *reciprocal* versus the *monitor* and *student* conditions were compared (1); the second contrast tested the difference between the *monitor* and *student* conditions (2). Later, this setting is referred to as *condition contrast setting B.*

In both session and paragraph video measures, there was a highly reliable difference between the *reciprocal* condition and the other two conditions (main effect condition(1) for session analysis $MSe = .30$, $F(1,88) = 249.54$ and $p < .001$ and $MSe = .22$, $F(1,335) = 1102.28$ and $p < .001$ for paragraph analysis). Students in the *reciprocal teaching* group determined the course of the session, whereas sessions in the *monitor* and *student* conditions were mainly determined by the trainer. At approximately 0.2 scale points, the difference between *monitor* and *student* conditions was very small compared to the difference of 1.8 scale points between the *reciprocal* condition and the *student* condition; however, this small difference became statistically significant in paragraph analysis; this was probably due to the large number of paragraphs observed ($F(1,335) = 6.39$, $MSe = .30$ and $p = .012$).

3.1.3.3. Amount of feedback given by children

For the two variables that measure feedback given by the children, analyses of variance with *contrast setting A* for the condition effect were computed. For both levels of video analyses, the main effect of the first condition contrast was highly reliable (for session analysis: $MSe = .45$, $F(1,87) = 200.08$ and $p < .001$; for paragraph analysis: $MSe = .37$, $F(1,335) = 410.00$ and $p < .001$). Children in the *reciprocal* and *monitor* conditions provided more feedback for their peers than children in the *student* condition. In session analysis, the difference between the *reciprocal* and *monitor* conditions was not statistically significant, but paragraph analysis found a very significant difference of about .5 scale points in favor of the *monitor* condition ($F(1,335) = 40.23$, $p = .012$).

This finding is relatively easy to explain: in the *monitor* condition, the children are prompted by the trainer to give each other feedback for every strategy application, whereas in the *reciprocal* groups the child acting as "teacher" for the current paragraph has to remember to provide feedback without prompting by the trainer. Moreover, providing feedback is only one of the many tasks he has to perform to keep the session running: planning the next strategy, appointing someone to apply this strategy and managing discipline. Because of this cognitive overload, "teacher" children forget to give feedback every now and then – this results in somewhat lower scores on that variable.

3.1.3.4. Summary

In line with expectations, the *reciprocal* condition was characterized by children determining the sessions and providing feedback for one another. The amount of time they talked increased steadily, exceeding 80 percent by the end of the training program. This is an indicator of children being able to carry out most or all tasks themselves and taking responsibility for their own learning. Sessions in the *monitor* condition were mainly determined by the trainer. Children provided the feedback, but even in late sessions they only talked for about two thirds of the time. The *student* condition was determined almost completely by the trainer; the children had very few responsibilities. The trainer provided the feedback and talked for approximately 30 percent of the time; students on the other hand for 70 percent.

In general, both levels of video analysis yielded similar results. There were only very small differences in judgments of children's co-determination between the time and content video analysis raters. The two video analysis levels differed considerably more with respect to the feedback ratings, but these findings were still in line with expectations.

3.1.4. Conclusion about treatment implementation

Due to the circumstances under which strategy training took place, it was not possible to randomly assign the children to the *control* and training groups. Children with poor verbal abilities were over-represented in the training group, and only a few children with excellent reading skills volunteered to participate in the training program. Because of these systematic differences in children's cognitive abilities and because of the low performance of training group children in measures of reading comprehension

at pre-test, it is necessary to include children's verbal ability as a control variable in further analyses. A combined measure of children's verbal abilities was computed to accomplish this. Additional to analyses of variance, correlations between cognitive abilities and dependent measures will be reported and cognitive abilities will be controlled by conducting hierarchical multiple regression analyses.

Three experimental groups (two *reciprocal* and one *student* group) had to be excluded from all further analyses because of severe discipline problems and inadequate interventions by the trainers. These three groups were re-sampled. In all other training groups, there were good working conditions, allowing children to acquire new knowledge in an adequate working atmosphere with appropriate discipline and trainers who intervened in a manner appropriate to restoring discipline when necessary.

The experimental conditions worked as intended; the expected differences were found between the experimental groups in the amount of time children talked, children's co-determination of sessions and the feedback they provided for one another.

3.2. Effects of strategy training

Because the main hypothesis of the study is that the students' task of providing feedback for one another is a main determinant of the success of the Reciprocal Teaching program and because two of the experimental conditions were designed to incorporate this feature, only these *reciprocal* and *monitor* conditions will be considered as "training conditions" and contrasted with the *control* group to estimate the "training effect."

Analyses of measures of children's knowledge will be reported before performance measures are analyzed. The available performance measures result from children writing a summary and from a test of their comprehension ability. Several indicators of children's metacognitive knowledge were included in the study. These can be classified into two categories: knowledge about the two specific reading strategies *summarizing* and *clarifying* and more general strategy knowledge (*relational* and *conditional strategy knowledge* and *planning knowledge*).

3.2.1. Knowledge measures

3.2.1.1. Specific strategy knowledge about summarizing and clarifying

At post-test, all children were asked to write down the three characteristics of a good summary. They were also required to write down everything that can be done to clarify unknown words or passages. Please note that the instruction in the booklet did not explicitly state the term "*clarifying*." Student responses were sorted into two categories. The first category covers text-related strategies to discover the meaning of words or text passages, like re-reading the unclear word, sentence or passage and searching for hints in the text that could help provide an explanation (like words in brackets or commas; phrases like "that means," "or," etc.). The second category of strategies can be labeled external strategies. Here, students wrote they would seek the help of other people (teacher, parents, siblings) or use external resources to look up the meaning of the unknown words (reference books and dictionaries or the internet). Table 15 lists the number of characteristics of a good summary noted by children in the *control* and training groups (*reciprocal* and *monitor* conditions) and the number of strategies that were produced by the children for *clarifying*.

Table 15. Declarative knowledge about summarizing and clarifying for control and trained children.

		Control		Training (RT+M)	
Strategy		M	SD	M	SD
Summarizing	1. contains most important content	.51	.50	.95	.23
	2. is shorter than the text	.36	.48	.78	.42
	3. is formulated in own words	.13	.34	.43	.50
	total N of characteristics	1.00	.89	2.16	.69
Clarifying	external: ask other people to help	.89	.66	.86	.95
	other resources (dictionaries, etc.)	1.13	.86	1.05	1.15
	text-related strategies	.87	.94	1.59	.98

For all three characteristics of a good summary, as well as for the total number of characteristics, highly significant effects of training were found (for characteristic 1: $F(1,119) = 25.17$, $MSe = .19$, $p < .001$; for characteristic 2: $F(1,119) = 21.77$, $MSe = .21$, $p < .001$; for characteristic 3: $F(1,119) = 14.90$, $MSe = .16$, $p < .001$; for the total number of characteristics of a good summary: $F(1,119) = 49.72$, $MSe = .70$, $p < .001$). On average, children in the *reciprocal* and *monitor* conditions knew about two of the three characteristics of a good summary; whereas *control* children could only name about one characteristic. Nearly all of the children in the training groups stated

that a good summary contains only the most important content and that details should be left out; this seems to be the "easiest" feature of a summary.

As for declarative knowledge about *clarifying*, no differences were found in students' reports of using external strategies, either asking other people to help or using external resources such as dictionaries. For use of text-related strategies, however, a highly significant main effect of training was observed ($F(1,119) = 14.83$, $MSe = .91$, $p < .001$): children in the training groups reported between one and two text-related strategies, whereas *control* children knew less than one of the strategies that make use of the text itself to discover the meaning of unclear words.

3.2.1.2. Relational strategy knowledge

An analysis of variance of *relational knowledge* about reading strategies was computed with training as between-subjects factor. A highly significant main effect of time was found ($F(1,133) = 20.99$, $MSe = .04$, $p < .001$). As shown in Figure 3, prior to training, children who participated in the training program scored lower on *relational strategy knowledge* than did *control* children. Both groups of children obtained higher scores after four weeks. It looks as if children in the training groups "caught up," but this interaction between training and time failed to reach significance ($F(1,133) = 1.95$, p = .16). The corrected effect size was computed and showed a small effect, with d = .28.

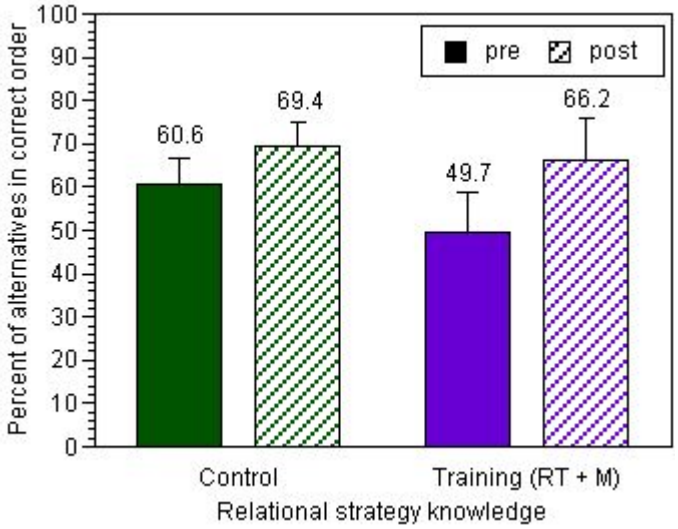


Figure 3. Relational knowledge about reading strategies; comparison between the control group (N = 78) and the two experimental conditions Reciprocal Teaching and Monitor (N = 37).

3.2.1.3. Conditional Strategy Knowledge and Planning Knowledge

Two subscales of the IRA Questionnaire were administered to all children both before and after training; one tapping metacognitive knowledge about *planning* reading activities, and the other measuring *conditional knowledge* about the usefulness of applying strategies in concrete situations. Table 16 shows means and standard deviations of children's scores. No significant effects were observed for either subscale.

Table 16. Conditional strategy knowledge and planning knowledge for children in the control and training groups.

		Control		Training (RT+M)	
		M	SD	M	SD
Conditional knowledge (IRA subscale)	pre	6.50	1.65	6.08	1.79
	post	6.69	1.73	6.54	1.68
Planning knowledge (IRA subscale)	pre	6.17	1.57	6.24	1.71
	post	6.72	1.34	6.11	1.78

3.2.1.4. Control of basic cognitive abilities

Table 17 shows the correlations between measures of metacognitive knowledge (prior to training) and basic cognitive abilities. Both verbal measures, *mean verbal ability* and *reading speed*, were highly correlated with metacognitive knowledge (coefficients between .28 and .45): more able children also possessed more metacognitive knowledge. *Nonverbal reasoning ability* only correlated reliably with *relational knowledge about reading strategies*.

Table 17. Correlations between pre-test metacognitive measures and knowledge about specific reading strategies and cognitive abilities.

	Nonverbal reasoning ability	Mean verbal ability	Reading speed
Relational strategy knowledge	.30 **	.45 **	.32 **
Conditional strategy knowledge (IRA)	.08	.35 **	.28 **
Planning knowledge (IRA)	.13	.28 **	.24 **
Summarizing: total number of characteristics	.28 *	.42 **	.22
Clarifying: external strategies			
- ask other people	-.07	-.01	.10
- use resources	.12	.12	.07
Clarifying: text-related strategies	.03	.26 *	.14

For knowledge about the two reading strategies *summarizing* and *clarifying*, correlations with cognitive abilities were only computed for the *control* group (see the lower part of table 17), because this knowledge was only assessed at post-test. In contrast to the results for the measures that tapped more general metacognitive knowledge, very few correlations with verbal abilities were observed. The total *number of characteristics of a good summary* reported by the children correlated positively with *mean verbal ability* (r = .42) and somewhat less strongly, but still significantly with *nonverbal reasoning ability* (r = .28). As for knowledge about *clarifying*, only the *number of text-related strategies* was positively correlated with *mean verbal ability*, with a coefficient of .26.

Correlation-based techniques as well as standard tests of interaction (ANOVAs) can be used to infer that the effects associated with group membership are greater in one variable than another if the groups being compared differ in the baseline level of performance and/or in other variables that are correlated to performance. This is what was observed for *relational strategy knowledge* and the IRA subscale *conditional strategy knowledge*: the children who participated in the training program showed lower levels of performance at pre-test; the two groups also differ in terms of mean verbal ability and reading speed, and these two variables are significantly correlated to test performance (with correlations between .28 and .45). Hierarchical multiple regression procedures examine the amount of variance related to training by determining the increment in variance associated with training after controlling for the variance in the other variables. If the residual training-related variance is significant, then it can be inferred that the processes were selectively and independently influenced by training.

To test whether there was residual variance related to training after controlling for baseline performance and students' cognitive abilities, I conducted hierarchical multiple regressions. Dependent variables were measures of metacognitive knowledge at post-test. In the first step, the corresponding metacognitive measures at pre-test were entered in the regression equation. Then, cognitive abilities were entered stepwise in the equation as predictors. Finally, training was introduced. The results of the hierarchical regression analyses for all measures of metacognitive knowledge are displayed in Table 18.

Table 18. Hierarchical multiple regression analyses of measures of metacognitive knowledge.

Dependent variable	Method	Predictor	ΔR^2	Δ F	p
Relational strategy knowledge (post)	1. enter	**baseline (pre)**	.18	24.84	< .001
	2. enter	training	.00	.01	.951
Conditional strategy knowledge – IRA (post)	1. enter	**baseline (pre)**	.08	9.25	.003
	2. stepwise	**reading speed**	.13	17.86	< .001
	3. stepwise	**mean verbal ability**	.04	5.58	.020
	4. enter	training	.01	1.15	.285
Planning knowledge – IRA (post)	1. enter	**baseline (pre)**	.09	11.52	.001
	2. stepwise	**mean verbal ability**	.03	4.75	.035
	3. enter	training	.02	3.40	.068
Summarizing: N of features of a good summary	1. enter	**nonverbal reasoning**	.07	8.22	.005
	2. enter	**training**	.30	52.57	< .001
Clarifying: N of text-related strategies	1. stepwise	**mean verbal ability**	.05	5.42	.022
	2. enter	**training**	.15	20.63	< .001

For the more general measures of strategy knowledge, baseline level of performance was a significant predictor, explaining between 8 and 18 percent of the variance in performance: the higher children's scores were at pre-test, the higher they were at post-test. Children's cognitive abilities did not explain additional variance in *relational strategy knowledge*. Also, when controlling for baseline level of performance, *training* failed to have a significant effect. In addition to baseline performance, *reading speed* and *mean verbal ability* explained significant portions of the variance in children's post-test scores for *conditional knowledge* (the better children's verbal abilities and the faster they are able to read, the greater their *conditional knowledge*). *Training* was not a significant predictor. *Planning knowledge* was predicted by baseline level of performance and *mean verbal ability*; again *training* failed to explain additional variance.

Where declarative knowledge about the two reading strategies *summarizing* and *clarifying* was concerned, the findings were different: here, training explained large amounts of performance variance above and beyond cognitive abilities. For the *number of features of a good summary*, there was a very large effect of *training* (Δ F(1,113) = 60.62), explaining 30 percent of the performance variance in addition to the rather moderate influence of *nonverbal reasoning ability* (5% of the variance with Δ F(1,114) = 5.37). When the *number of text-related strategies* was predicted, the portion of variance explained by introducing *training* as the second predictor was three times that explained by *mean verbal ability* (15% vs. 5%, respectively).

Another measure is also very informative with respect to the effects of training: the effect size. Instead of relying on distribution of measures and tests of significance, the difference between treatment and control conditions is calculated relative to the standard deviation of the control group. The resulting measure (d) reflects the size of the treatment effect in units of standard deviations. For all three features of a good summary, there were large to very large effect sizes (1.13, .93 and .70, respectively) and the effect size for the total number of features identified was very large, at d = 1.46. The effect size for the number of text-related strategies produced for *clarifying* was also large, at d = .75.

3.2.2. Performance measures

3.2.2.1. Writing a Summary

At post-test, all students were required to write a summary about a text that did not have a heading. Ratings of the summaries provided a number of variables which will be reported here: number of words in the summary, generation of a title, underlining in the original text (in the test booklet), content of the summary and "grade a teacher would give". The data are presented in table 19.

Table 19. Means and standard deviations for features of the summaries written by children in the control and training groups.

	Control		Training (RT + M)	
Features of the summaries	M	SD	M	SD
Generation of a title (% of students)	2.4	15.3	10.8	31.5
Underlining in text (% of students)	3.6	18.7	21.6	41.7
Number of words	54.21	17.08	40.41	11.27
Linguistic style	2.74	1.18	2.89	1.07
Number of main ideas	7.24	2.40	6.14	2.06
Number of details	13.98	7.01	10.05	5.98
Ratio number of main ideas / details	.74	.84	1.13	1.45
"Grade a teacher would give"	4.41	.96	4.16	.99

The mean number of words children wrote in their summaries is depicted in Figure 4. The effect of training was highly significant in an analysis of variance ($F(1,119) = 20.25$, $MSe = 241.87$, $p < .001$). Children who participated in the training program wrote shorter summaries than children in the *control* group. The size of this effect was very large (d = -.95).

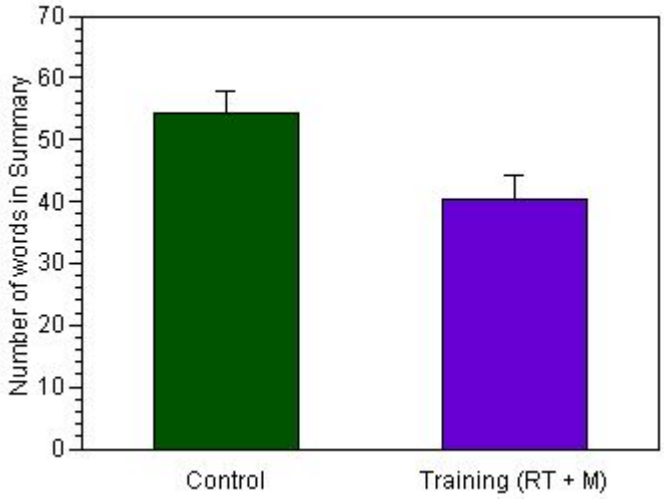


Figure 4. Number of words written in the summary by children in the control and training groups.

The next interesting feature is whether or not children *generated a title* for their summary. Only very few *control* children (less than 3 percent) generated a title; compared to just over 10 percent of the children in the reciprocal and *monitor* groups. This was reflected in a significant training effect in an ANOVA ($F(1,119) = 3.94$, $MSe = .05$, $p < .05$) and a small effect size (d = .34). The raters also noted whether the children *underlined text* in their booklet during the production of their summary. More than 20 percent of the children in the training groups used this strategy, but less than 5 percent of the *control* children. This effect also proved to be highly significant in an analysis of variance ($F(1,119) = 10.87$, $MSe = .08$, $p < .001$). With an effect size of .56, it is of moderate magnitude. No differences between the two groups were found with respect to the *linguistic style* of the summaries.

Thus far, only findings dealing with surface features of the summaries have been reported. But *what* did the children write? The content of the summaries was classified to be either important (a main idea) or unimportant (a detail). Main ideas that were more central were weighted double; as were very unimportant details. *Control* children wrote about one *main idea* more than training group children in their summaries ($F(1,119) = 5.90$, $MSe = 5.30$, $p = .017$). But they also included about three more *details* ($F(1,119) = 8.73$, $MSe = 45.08$, $p < .001$). The most informative measure is thus the *ratio between main ideas and details*. Here, the analysis of variance yielded various results in favor of the training group children: at 1.13, their mean ratio is greater than 1 – this means that they wrote more important than unimportant content. The ratio of the *control* children, in contrast, was less than 1 (.74); they wrote more details than

main ideas. Unfortunately, this effect just failed to reach significance with $F(1,119) = 3.44$, $MSe = 1.13$ and $p = .066$. When the variance due to general ability to comprehend texts was controlled by dividing students into groups of high and low level comprehenders according to the total number of points they scored in the pre-test comprehension assessment (control of text comprehension ability), not only did a MANOVA find a highly significant main effect of text comprehension ability ($F(1,112) = 5.53$, $MSe = 1.09$, $p < .01$) in favor of good comprehenders, whose ratio was higher than that of bad comprehenders, but the effect of training also became significant ($F(1,112) = 7.15$, $p = .02$). Students who participated in the training program wrote more important ideas than details in their summaries relative to *control* children. The larger number of main ideas and details produced by *control* children is also reflected in moderate effect sizes (d = -.49 for main ideas and -.60 for details); the more positive ratio of main ideas to details produced by children in the training group resulted in a small positive effect (d = .33).

Finally, the *grade a teacher would* give, which was defined as a composite grade taking all features of the summary into account, was subjected to an ANOVA. The overall quality of the summary was rated on a grade-equivalent 6-point Likert scale ranging from 1-excellent to 6-fail. The first interesting result is that, overall, the summaries received only poor grades of about 4, which means "acceptable." It looks as if children in the training groups wrote better summaries, but at p = .206 this effect was not significant ($F(1,119 = 1.62$, $MSe = .94$). The size of this effect is rather small, at -.26.

Correlations between students' cognitive abilities and features of the summaries are displayed in table 20. Again, only data from the *control* children (N = 85) were used. Few significant correlations were observed. Children who obtained higher scores in *verbal ability* measures wrote more *words*: their summaries included both more *main ideas* and more *details*. The *number of main ideas* contained in the summaries also correlated significantly with *reading speed* and *nonverbal reasoning ability* (correlation coefficients between .35 and .40).

Table 20. Correlations between features of the summaries and cognitive abilities.

	Nonverbal reasoning ability	Mean verbal ability	Reading speed
Generation of a title	.11	.12	.07
Underlining in text	.17	.17	.12
Number of words	.16	.27 *	.16
Linguistic style	.11	.11	.13
Number of main ideas	.40 **	.39 **	.35 **
Number of details	.22	.31 **	.19
Ratio main ideas / details	.06	-.04	-.06
"Grade a teacher would give"	.03	-.18	-.13

Will the effects of training remain constant when controlling statistically for the influence of cognitive abilities? Table 21 documents the results of hierarchical multiple regression analyses for the features of the summaries that were significantly correlated with cognitive abilities. Students' cognitive abilities were entered stepwise in the equations before training was introduced as a predictor.

Table 21. Hierarchical multiple regression analyses of measures of knowledge about specific reading strategies.

Dependent variable	Method	Predictor	ΔR^2	Δ F	p
Number of words in summary	1. stepwise	**mean verbal ability**	.07	8.82	.004
	2. enter	**training**	.11	15.20	< .001
Number of main ideas	1. stepwise	**mean verbal ability**	.16	21.56	< .001
	2. stepwise	**reading speed**	.04	5.02	.027
	3. enter	training	.01	1.53	.219
Number of details	1. enter	**mean verbal ability**	.04	4.53	.035
	2. enter	**training**	.05	6.47	.012
Ratio main ideas / details	1. enter	mean verbal ability	.01	.12	.732
	2. enter	training	.03	3.71	.057

For the *number of words* in the summaries, the effect of *training* was highly reliable (explaining about 11 percent of variance in performance), even after controlling for the effect of *mean verbal ability*. Mean verbal ability was positively correlated with the number of words (standardized beta coefficients of .20), whereas training was negatively related to the number of words (standardized beta coefficients of -.34). This means that *control* children with better verbal cognitive abilities wrote more words than less able children. In the course of the training program, however, the children in the

training groups were taught that summaries should be short, and they began to apply this knowledge. The effect of training was larger than that of mean verbal ability.

The *number of main ideas* was reliably predicted by *mean verbal ability* and *reading speed*; training did not explain additional variance. For *number of details*, both *mean verbal ability* and *training* were significant predictors of performance; mean verbal ability with a positive standardized beta coefficient of .15 and *training* with a negative standardized beta coefficient of -.23. The corresponding data are presented in figure 5. Lines represent linear regressions.

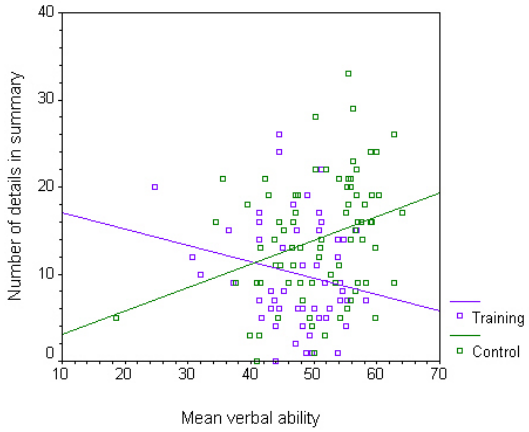


Figure 5. Scatter plot of number of details in summary and mean verbal ability.

The figure shows that the higher the scores of the *control* children on mean verbal ability, the more details they wrote in their summary. For children who participated in the training program, this was not the case: here, the gradient of the line was negative. Participating in the training program suppressed the effect of mean verbal ability and the gradient of the line was reversed. This means that the higher the scores of the children in the training group on mean verbal ability, the fewer details they wrote. In other words, in the training program they learned that details should not be included in a summary.

The result of regression analysis of the ratio of main ideas to details is also shown in table 21. After controlling for the effect of mean verbal ability (which predicted both the number of main ideas and the number of details that the children included in their summaries highly reliably), the training effect just failed to reach the specified significance level (with $p = .057$).

3.2.2.2. Text comprehension

Text comprehension was measured both before and after training with four open-ended questions on a longer text. Students' answers were analyzed in two different ways. First, all answers were categorized as belonging to one of three competence levels. Second, raters determined whether the answers contained important or less important content. The results that will be reported are analyses of variance for the total number of answers and for the answers at each competence level. As in the analysis of content of the summaries, the relation between the main ideas and details was also computed. The data are displayed in table 22.

Table 22. Text comprehension scores for children in the control and training groups.

			Control		Training (RT+M)	
			M	SD	M	SD
Points	level 1	pre	2.94	2.02	1.96	1.63
		post	3.11	2.01	2.81	1.84
	level 2	pre	7.15	3.92	5.14	2.83
		post	13.80	4.93	11.43	4.43
	level 3	pre	.51	.71	.43	.56
		post	1.70	1.39	1.24	1.40
	total	pre	10.60	5.50	7.26	4.09
		post	18.61	6.30	15.49	5.71
Content	main ideas	pre	11.68	6.47	8.43	4.87
		post	13.84	5.78	11.81	4.94
	details	pre	3.11	2.71	2.00	2.16
		post	6.23	4.14	4.84	3.18
	ratio main ideas/details	pre	4.89	4.28	4.98	3.53
		post	3.42	2.86	3.56	2.82

With respect to the *total number of points* scored on the comprehension questions, significant main effects of group and time ($MSe = 50.80$, $F(1,115) = 9.55$ and $p < .01$ for group and $MSe = 12.69$, $F(1,115) = 254.03$ and $p < .001$ for time) were observed. *Control* children obtained more points before and after training than children who participated in the strategy training. All children wrote a lot more at post-test (about double the amount), shown in large increments in the total number of points. This analysis was repeated for each competence level. Significant differences between children in the *control* and training groups emerged for competence levels 1 and 2; for answers on the highest level 3, where children obtained only about ½ point at pre- and between 1 and 2 points at post-test; the group effect was not significant. Highly reliable

improvements in performance were obtained at all levels of competence for all children; the largest increase was found in answers coded as competence level 2.

Analyses of variance of the *number of main ideas* revealed highly reliable main effects of group and time (main effect of group: $MSe = 49.73$, $F(1,115) = 7.06$ and $p < .01$; main effect of time: $MSe = 17.07$, $F(1,115) = 22.75$ and $p < .01$). At both times, *control* children wrote more main ideas than children in the training group. All children produced more important content at post-test than before training. For the *number of details*, the same picture emerged; there were significant effects of group and time ($MSe = 13.00$, $F(1,115) = 6.08$ and $p = .015$ for group and $MSe = 8.46$, $F(1,115) = 52.796$ and $p < .01$ for time). All children wrote more details after the training program than before and *control* children wrote more details at both times than children in the training groups. The most interesting measure is again the relation between important and non-important content: the *ratio of main ideas to details*. ANOVA yielded only a highly reliable main effect of time ($MSe = 12.50$, $F(1,115) = 8.41$ and $p < .01$); surprisingly, the ratio of main ideas to details decreased at post-test. This means that prior to strategy training, the children wrote relatively more important than non-important content, whereas four weeks later, when they wrote a lot more in their answers, the relation shifted to more non-important than important content.

Table 23. Correlations between text comprehension scores at pre-test and cognitive abilities.

		Nonverbal reasoning ability	Mean verbal ability	Reading speed
Points	level 1	.28 **	.45 **	.33 **
	level 2	.36 **	.58 **	.54 **
	level 3	.22 **	.21 *	.12
	total number	.39 **	.60 **	.52 **
Content	main ideas	.35 **	.57 **	.53 **
	details	.19 *	.39 **	.28 **
	ratio main ideas/details	.16	.16	.15

How was text comprehension related to students' cognitive abilities? Table 23 shows the correlations between the text comprehension scores at pre-test and students' nonverbal reasoning ability, their mean verbal ability and reading speed. All measures of reading comprehension apart from the ratio of main ideas to details were significantly

positively correlated with *nonverbal reasoning ability*, *mean verbal ability* and *reading speed*. The highest correlations were obtained for *mean verbal ability*.

Training did not explain additional variance in any of the measures of text comprehension after controlling for the influence of *baseline level of performance* and cognitive abilities (*mean verbal ability*, *reading speed* and *nonverbal reasoning ability*) with hierarchic regression analyses (see table 24). *Baseline level* of performance and *mean verbal ability* predicted the *number of points at comprehension levels 1 and 2*, the *total number of points* and the *number of main ideas and details* that children's answers contained. *Reading speed* explained additional performance variance in the *number of main ideas* produced and served as the only predictor for *number of points at the highest comprehension level*.

Table 24. Hierarchical multiple regression analyses of measures of text comprehension abilities for subjects of the control and training groups.

Dependent variable	Method	Predictor	ΔR^2	Δ F	p
Total number of points (post)	1. enter	**baseline (pre)**	.40	77.44	< .001
	2. stepwise	**mean verbal ability**	.06	13.05	< .001
	3. enter	training	.01	.56	.454
Level 1 – number of points (post)	1. enter	**baseline (pre)**	.06	6.87	.010
	2. stepwise	**mean verbal ability**	.04	4.61	.034
	3. enter	training	.00	.03	.873
Level 2 – number of points (post)	1. enter	**baseline (pre)**	.40	76.86	< .001
	2. stepwise	**mean verbal ability**	.06	12.14	.001
	3. enter	training	.01	.43	.512
Level 3 – number of points (post)	1. enter	baseline (pre)	.00	.01	.947
	2. stepwise	**reading speed**	.12	14.90	< .001
	3. enter	training	.01	.23	.631
Number of main ideas (post)	1. enter	**baseline (pre)**	.26	40.50	< .001
	2. stepwise	**reading speed**	.11	18.80	< .001
	3. stepwise	**mean verbal ability**	.03	4.56	.035
	4. enter	training	.01	.14	.712
Number of details (post)	1. enter	**baseline (pre)**	.07	8.04	.005
	2. stepwise	**mean verbal ability**	.10	12.93	< .001
	2. enter	training	.01	.90	.344
Ratio main ideas / details (post)	1. enter	baseline (pre)	.00	.04	.841
	2. enter	training	.01	.10	.750

3.2.3. Summary of findings

Control children were compared with children who participated in one of the two experimental training conditions that involved the components hypothesized to be effective for producing the desired effects (*reciprocal teaching* condition and *monitor*

condition) with respect to knowledge and performance measures. Some of the measures were administered both before and after training (various measures of more general strategy knowledge and text comprehension); others were only tested after strategy training (knowledge about specific reading strategies taught in the training program and about strategy application).

It is important to bear in mind that there were large differences between the *control* and training groups with respect to verbal cognitive abilities (i.e., vocabulary knowledge, decoding ability, reading comprehension and grade in native language) when considering the methods of analyses that were used to explore the effects of strategy training on knowledge and performance measures. Only few children who scored high on measures of verbal skills participated in the training program; most of the children in the training groups had poor verbal abilities. For *control* children, it was the other way around. Cognitive abilities were significantly correlated with students' knowledge and performance prior to training (baseline levels of knowledge and performance): mean verbal ability and, in most cases, reading speed showed highly reliable positive correlations with *relational* and *conditional strategy knowledge* and *planning knowledge* (correlation coefficients between .30 and .45), as well as with almost all measures of text comprehension (number of main ideas and details; total number of points scored and number of points assigned to competence levels 1 and 2, with correlation coefficients between .30 and .60). To control the influence of students' cognitive abilities, hierarchical multiple regressions were used in addition to analyses of variance to determine which measures were affected by training.

What effects did the training of reading strategies have on students' knowledge? Strategy training had large positive effects on knowledge about two of the reading strategies taught: *summarizing* and *clarifying*. Children who had participated in strategy training knew significantly more characteristics of a good summary than *control* children did; this was true for all of the three characteristics, as well as for the sum of these characteristics. Children in the training groups also reported significantly more text-related strategies when asked what could be done to clarify an unknown word. The effects of training remained stable after controlling for cognitive abilities.

Students' *relational strategy knowledge* improved significantly from pre- to post-test. Although the increment appeared to be larger for children in the training groups, this effect was not significant. For the *conditional strategy knowledge* and *planning knowledge* subscales of the Index of Reading Awareness, no significant effects

were observed. The effects of training remained non-significant after controlling for verbal abilities in addition to baseline performance. In short, no effects of strategy training could be found on these three more general measures of strategy knowledge.

The performance measure directly related to the training program was the production of a summary. Training in reading strategies produced highly reliable effects. Children in the training groups wrote shorter summaries, and more of them generated titles and underlined in the original text when producing the summary than *control* children did. When reviewing the results of all analyses of summary content together, the following picture emerged: the children who participated in the strategy training program wrote shorter summaries containing fewer details than *control* children did. This is a positive effect of the training program, because one of the characteristics of a good summary is that it is shorter than the text. More interesting, the relation between main ideas and details was better for children who participated in the training program than for *control* children: their summaries contained more important than non-important ideas (this just failed to reach the specified significance level when analyzed with analysis of variances and regression analysis, but had an effect size of .33). All in all, training in reading strategies had a large positive impact on students' abilities to write summaries.

For text comprehension, which was not directly related to strategy training, large increments in performance were observed from pre- to post-test for all children. Students wrote a lot more at post-test. In qualitative terms, this was indicated by a reliably higher total number of points due to significantly higher scores for answers on competence levels 1 and 2. In quantitative terms, it was indicated by significantly more main ideas and more details. The ratio of main ideas to details, however, decreased from pre- to post-test. Children who participated in the training program scored lower than *control* children on all comprehension measures apart from the ratio of main ideas to details. This was true at both pre- and post-test. Mean verbal ability predicted performance on almost all comprehension measures (with the exception of the ratio of main ideas to details) in addition to baseline level of performance; training did not explain significant portions of the variance.

3.3. Effects of experimental conditions

Does the way reading strategies are taught really matter for children's learning? By contrasting the experimental conditions, the question of which features of the program are responsible for its success and may be necessary to acquire metacognitive knowledge and competences (hypothesis 2) can be addressed. Large differences between the *reciprocal* and *monitor* conditions on the one hand and the *student* condition on the other hand were predicted.

Before reporting the differences that emerged between the experimental conditions in measures of knowledge and performance, I want to examine the training program in detail: How much of the allotted study time was actually spent on training reading strategies (engaged time)? How many text paragraphs did the groups finish? How often did the children have the opportunity to apply the reading strategies and how successful were they in doing this?

The section will close by considering the effects of the training program on motivation, with data drawn from a post-training questionnaire.

3.3.1. Practicing reading strategies during training

The goal of this section is to examine children's success in applying the reading strategies. Some of the questions I wish to answer are as follows: How well did the children apply the reading strategies during training? Were the reading strategies equally difficult? Was there improvement over time? These questions can be addressed using the data from the video analyses, where the success of every application of the strategies *summarizing*, *questioning* and *predicting* was rated. The remainder of the section considers how successful children in *reciprocal* and *monitor* groups were in adapting the roles of "teachers" and "monitors."

Before turning to these analyses, however, I need to look at the available data more closely, because only part of the sessions (though a large part) were video-recorded, making it possible to explore how the allocated time was actually used and whether children were "on task" during study time. The speed of training and the number of learning opportunities will also be considered because these are important for comparing children's progress in applying the reading strategies.

3.3.1.1. Use of study time

Everything that happened in training was classified as belonging to one of fourteen categories and the time it took was noted (session analysis data). Similar categories were combined, resulting in 10 groups of activities. In figure 6, the mean amount of time spent on each of these activities during the training sessions is depicted separately for the strategy instruction sessions (sessions 1 to 3) and the early and later practice sessions (sessions 4 to 9 and 10 to 15, respectively). Please note that double sessions were not included in this computation of the mean amount of time per activity.

Four of the activities were directly concerned with work on the text paragraphs: (1) reading text paragraphs aloud, (2) verbal application of the strategies (one child at a time), (3) silent work with the text (underlining in the text, writing a summary or notes; all children writing at the same time) and (4) the children working together with the trainer on the chalkboard or flipchart. These activities are depicted in blue or purple. Two other very important tasks that were performed by trainers are strategy instruction (orange) and role instruction for “teacher” and “monitor” students (red). At the beginning of a session, either the reading strategies or the content of the last text paragraph were repeated; the time devoted to this is represented in yellow. Time spent on tests (summaries of text paragraphs or answering comprehension questions about a text) is represented in gray. The category for organizational matters and chatting is shown in dark green. Finally, times when the trainer had to deal with discipline problems and this intervention interrupted the course of the training are shown in light green.

First, strategy instruction sessions and practice sessions will be contrasted before the practice sessions are analyzed in more detail. The changes in time patterns that occurred over the course of the 10 practice sessions are of particular interest. It is for this purpose that the practice sessions were divided into early and later sessions. Both the absolute amount of time and the proportion of total time per session (relative amount of time) spent on the activities was subjected to multivariate analyses of variance. When analyses for both time measures yielded the same results; significances are only reported for the relative amount of time.

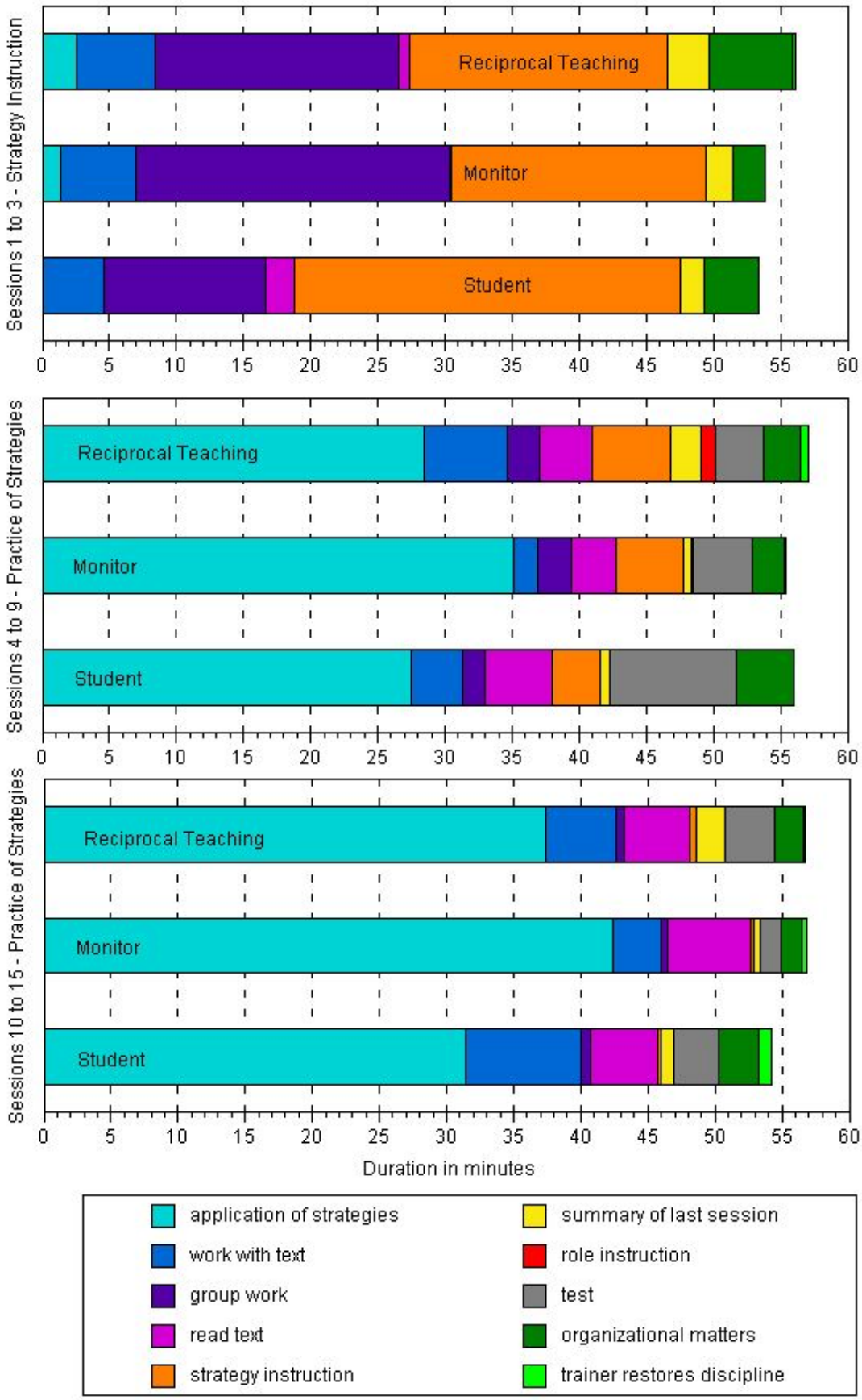


Figure 6. Duration of activities in three phases of the training program (mean number of minutes per session).

As can be seen in figure 6, there were considerable differences in the temporal pattern of activities that occurred in strategy instruction and practice sessions: in the strategy instruction sessions, almost all of the time (85 percent) was spent on teaching the reading strategies and working on text material (worksheets). These two very important activities took about even amounts of time. In contrast, the practice sessions were characterized mainly (70 to 90 percent of the time) by children working with the texts (expository texts divided into paragraphs), and instruction occurred only very rarely. The kind of activities performed with the texts also differed. In the instruction sessions, the children worked mainly together as a group, whereas in practice sessions, work with texts concentrated on oral practice of the reading strategies, one child at a time.

Over the course of the training program, the amount of time spent working on text paragraphs increased significantly from about 40 minutes during the first half of the practice sessions to 47 minutes in later practice sessions ($MSe = .03$, $F(1,77) = 15.35$ and $p < .01$). More specifically, during later practice sessions the children spent more time applying the reading strategies orally ($MSe = .04$, $F(1,77) = 7.58$ and $p < .01$) and reading the texts aloud ($MSe = .01$, $F(1,77) = 8.44$ and $p < .01$; due to increase in length of paragraphs) and less time working as a group on problems with the text ($MSe = .01$, $F(1,77) = 5.72$ and $p = .02$).

On the other hand, less time was spent with strategy instruction by the trainer (decrease from about 5 minutes to less than one minute per session; $MSe = .02$, $F(1,77) = 7.98$ and $p < .01$). The decrease in time spent on instruction of roles, which occurred only in the *reciprocal* and *monitor* conditions, could not be tested for significance because roles were only instructed by the trainers in the earlier practice sessions.

Far fewer differences were observed between the experimental conditions than between the earlier and later sessions. The most important difference between the experimental conditions was, of course, role instruction. Because no roles had to be instructed in the *student* condition, this activity was only observed in the *reciprocal* and *monitor* conditions. Even here, role instruction was only observed during the early practice sessions. In the later sessions, no role instruction was recorded. Mean instruction time in *reciprocal* conditions was about one minute; in the *monitor* conditions it was far less than half a minute. Because of the reduced variance, none of the effects reached the specified significance level.

In *student* groups, organizational matters took about a minute longer than in the sessions observed for *reciprocal and monitor* groups (*MSe* = .01, $F(1,77) = 4.04$ and $p < .05$ with contrast setting A).

Large differences between the experimental conditions were observed when it came to summarizing the last session (repeating the content of the last text paragraph or strategy applied): the *reciprocal* groups spent a little more than two minutes on repetition, whereas in the other two conditions this activity took less than one minute (contrast setting B *MSe* = .01, $F(1,77) = 9.50$ and $p < .01$).

Finally, time devoted to the trainer restoring discipline when this intervention interrupted the momentum of the training is considered. A significant interaction between time and the first condition contrast for setting B was observed: in the *reciprocal* condition the time that was spent on restoring discipline decreased during the training program, whereas in the *monitor* and *student* conditions it increased (*MSe* = 1.43, $F(1,77) = 4.38$ and $p < .05$ for absolute and *MSe* = .01, $F(1,77) = 4.51$ and $p < .05$ for relative amount of time with contrast setting B).

3.3.1.2. Time on task

Time on task was another very informative aspect of the use of study time that was judged by the video raters. The amount of time that the children were either actively or passively engaged in the training program was rated for each child individually in steps of 10 percent.

The total number of individual ratings for time on task was 1336. In only 25 cases was time on task rated to be below 80 percent. Aggregated for every child, mean time on task was generally very high, with a mean of 94.57 percent (SD = 6.73); the minimum was 88.41 percent. In the *reciprocal teaching* condition the mean was 95.84 percent, in the *monitor* condition children were on task for about 94.24 percent of the time, and for the *student* condition mean time on task was 93.66 percent (SDs of 1.80, 3.37 and 1.76, respectively). These differences between experimental conditions were significant ($F(2,52) = 3.78$, *MSe* = 6.11, $p = .029$), but they are so small that they are of no practical importance and thus negligible.

3.3.1.3. Number of learning opportunities

The speed of progression through the training program is important because it is directly associated with children's learning opportunities: the more paragraphs that were read, the more practice was possible for each child and strategy. The most obvious

indicator of 'speed' of training is the number of texts, or more specifically, the number of paragraphs completed by each group during the training program.

As can be seen in Table 25, the four groups in the *reciprocal teaching* condition completed the fewest text paragraphs in the training program, followed by *monitor* condition groups. Children in the *student* condition read the most text paragraphs. However, when this effect was tested with an ANOVA, it failed to reach significance due to the large standard deviation in the *student* condition (SD = 10.1) and the few degrees of freedom (*MSe* = 43.89, $F(2,9) = 2.79$ and $p = .114$).

Table 25. Number of text paragraphs read in training and number of paragraphs recorded on video.

	Reciprocal		Monitor		Student	
	N_{total}	N_{video}	N_{total}	N_{video}	N_{total}	N_{video}
Group 1	37	15	49	28	54	38
Group 2	36	25	47	29	33	12
Group 3	36	19	39	26	54	44
Group 4	39	36	39	28	51	29
Mean	37.0	23.7	43.5	27.7	48.0	30.7
SD	1.4	9.1	5.3	1.3	10.1	13.9

In addition to the total number of text paragraphs read, the number of paragraphs video-taped is also listed in table 25. The percentage of video-taped paragraphs varied substantially between the twelve groups: from 36 to 92 percent. However, there were no differences between the three experimental conditions in terms of the mean percentage of paragraphs recorded on video (*reciprocal* and *monitor* each 64 % and *student* 61 %).

Another interesting variable, which is directly associated with the number of text paragraphs that the entire group of children worked on, but also dependent on the number of children in the group, is the number of opportunities for each individual child to practice the four reading strategies. Table 26 lists the mean number of applications of the three reading strategies *summarizing*, *predicting* and *questioning* that were observed for each child (and hence available on video). The number of students for whom at least one strategy application was observed is also given. Please note that four children were never video-taped trying to predict future text content (one child in the *monitor* and three children in the *student* condition).

Each child was observed about 9 times trying to *summarize* a text, about 11 times asking a question about the paragraph and about 7 times *predicting* the content of

the upcoming text. There were considerable differences between the experimental conditions in the mean number of strategy applications: children in the *student* condition were observed summarizing the paragraph and asking questions more often than children in the *reciprocal* and *monitor* conditions (in an ANOVA with contrast setting A for condition(1) $MSe = 22.80$, $F(1,51) = 7.70$ and $p < .001$ for *summarizing* and $MSe = 33.43$, $F(2,9) = 6.44$ and $p < .05$ for *questioning*). This is in line with the differences in the number of text paragraphs finished by the groups reported above: the fewer paragraphs that were completed, the fewer strategy applications were video-taped and analyzed. For *predicting*, the picture was somewhat different: students in the *reciprocal* condition were observed *predicting* more often than students in the *monitor* and *student* conditions. This was reflected in a highly significant main effect of condition(1) with contrast setting B ($MSe = 19.99$, $F(1,52) = 12.29$ and $p < .01$). At first sight, this finding is somewhat unusual, but it becomes plausible when the videos are explored more carefully: in two of the *reciprocal* groups, the children routinely asked not just one student to make a prediction, but went around the table asking every student to predict the text content.

Table 26. Mean number of observed strategy applications per child.

	Reciprocal (17)			Monitor (20)			Student (18)		
	N	M	SD	N	M	SD	N	M	SD
Summarizing	17	8.4	4.8	20	6.8	2.1	18	11.4	6.6
Questioning	17	10.9	6.1	20	8.9	3.8	18	14.2	7.2
Predicting	17	9.4	3.7	19*	5.6	3.3	15*	5.5	4.2

*

The following analyses of the quality and quantity of application of the reading strategies by the children have to be interpreted very carefully. As mentioned above, the percentage of video-taped sessions varied substantially between training groups. Moreover, the number of observed strategy executions varied systematically between experimental conditions and not all children were observed making predictions of future text content.

3.3.1.4. Quality and improvement of strategy application during training

The question is now how successful the children were in applying the reading strategies. Video raters assessed the quality and voluntariness of three of the reading strategies: *summarizing*, *questioning* and *predicting*. *Clarifying* was not rated because

* ∗ Not all children could be observed executing the strategy *predicting*.

detecting an unknown word or confusing text was not subject to deliberate practice during training. Furthermore, the children usually worked together as a group to discover the meaning so that it was not possible to determine the contribution of a single child to the solution.

On the basis of the low frequencies of strategy applications and an inspection of the videos, I decided to address this issue for two time periods of equal length: the early practice sessions (sessions 4 to 9) and the later practice sessions (session 10 to 15). Figure 7 shows the grades children in the three experimental conditions received for their applications of the reading strategies *summarizing*, *questioning* and *predicting*. Please note that, in accordance with the German grading system, the best grade that could be attained is 1; the lowest grade is 6.

First, I want to take a look at the overall difficulty of these three strategies. *Questioning* appears to be the most difficult strategy because it received the lowest grades (M = 3.01, SD = .52). *Summarizing* was also fairly difficult, with a mean grade of 2.59 (SD = .48); whereas the mean grade given by the raters for *clarifying* was 1.37 (SD = .46). These differences were reflected in a highly significant strategy effect (*MSe* = .20, $F(2,96) = 175.84$ and $p < .01$; MANOVA with strategy(3) as within-subjects factor and experimental condition as between-subjects factor). The interaction between experimental condition and strategy was also significant ($F(4,96) = 3.25$ and $p = .015$).

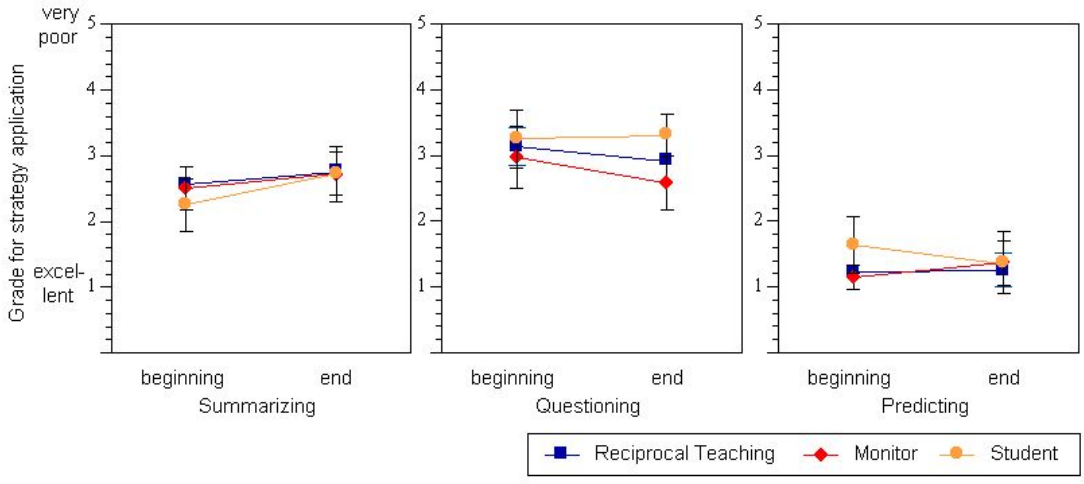


Figure 7. Grades children received for strategy application during training.

Figure 7 makes it clear that the interactions between experimental conditions and strategies are fairly complex; therefore, the grades were analyzed separately for

each strategy with repeated measures analyses of variance with time(2) as within-subjects factor and experimental condition with contrast setting A as between-subjects factor. For *summarizing*, there was only a significant main effect of time; all children received slightly lower grades towards the end of the training (*MSe* = .54, $F(1,49) = 4.35$ and $p < .05$). No differences between the early and later sessions of the training program were observed for *questioning*; but a significant main effect was found for the first condition contrast (*MSe* = .54, $F(1,49) = 4.35$ and $p < .05$). That is, children in the *reciprocal* and *monitor* groups asked better questions than children in the *student* condition. For *predicting*, no significant effects were observed.

The next figure, figure 8, shows how often children volunteered to apply the different strategies. It is easy to see that children volunteered more often for the easiest strategy, *predicting*. Here, the mean for all groups was 59.0 percent (*SD* = 31.7). The fewest voluntary strategy applications were observed for *summarizing* (*M* = 15.6, *SD* =14.2); *questioning* was done voluntarily in somewhat less than half the cases (*M* = 41.8, *SD* =23.7). When these differences were analyzed with a repeated measures analysis of variance with strategy as within-subjects and experimental condition as between-subjects factors, the main effect of experimental condition as well as the strategy main effect and the experimental condition by strategy interaction became highly significant (for experimental condition: *MSe* = .08, $F(2,48) = 7.30$ and $p < .01$; for strategy: *MSe* = .03, $F(2,96) = 84.78$ and $p < .01$; for the interaction between strategy and experimental condition: $F(4,96) = 6.86$ and $p < .01$). Also these differences are were examined in detail by separate MANOVAs for each strategy with contrast setting A for the factor experimental condition.

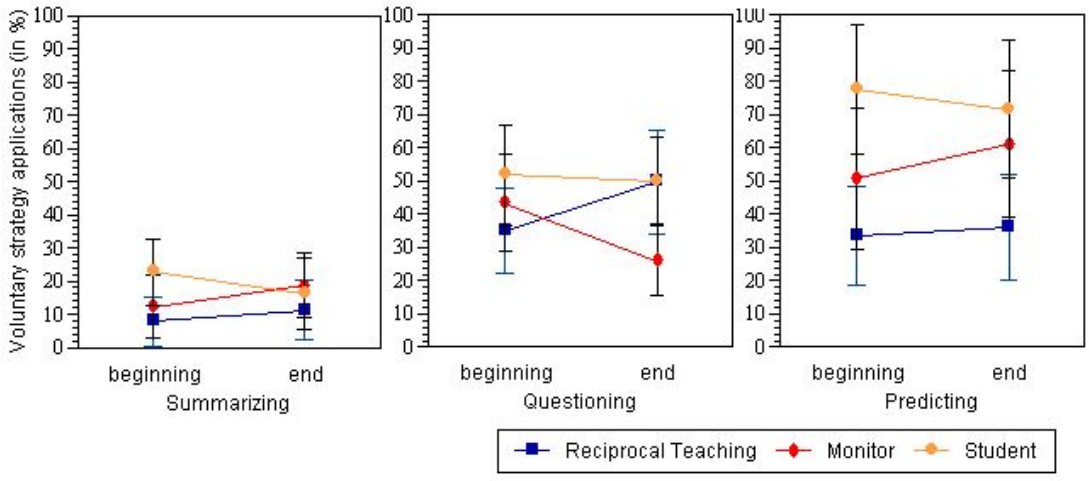


Figure 8. Mean percentage of voluntary strategy applications observed.

For voluntariness of *summarizing*, neither significant differences between experimental conditions nor changes over time were observed. Children in all three experimental conditions rarely volunteered to give a summary of the text paragraph. No significant main effects were found for voluntary *questioning*, but a highly reliable interaction was observed between the second condition contrast and time (*MSe* = .04, $F(1,50) = 10.60$ and $p < .01$). Whereas there were no changes over time in the *student* condition, children in *reciprocal* groups volunteered more often towards the end of the training program, and children in *monitor* groups volunteered less often at the end of the program. Very large differences between the three experimental conditions were found for voluntary *predicting*. Here, the children in the *student* condition volunteered much more frequently than the children in the other two conditions (condition(1) *MSe* = .14, $F(1,39) = 10.10$ and $p < .01$). Moreover, students in the *monitor* groups volunteered more frequently than children in *reciprocal* groups (condition(2) $F(1,39) = 4.62$ and $p < .01$). Time effects were not observed for any of the strategies.

In addition to the analysis of mean grades and mean percentage of voluntary strategy applications observed, the best grades that the children received for each of the strategies were analyzed as another indicator of children's success. Because there were no time effects, the means of all sessions are displayed in table 27.

Table 27. Mean best grade received for strategy application.

	Reciprocal		Monitor		Student	
	M	SD	M	SD	M	SD
Summarizing	1.35	.61	1.40	.68	1.11	.32
Questioning	1.41	.62	1.40	.68	1.39	.85
Predicting	1.00	.00	1.00	.00	1.00	.00

The first interesting finding is that every child was awarded the best possible grade for making a *prediction* at least once. For *summarizing* and *questioning* the mean of best grades was between 1 and 2. There were no significant differences between experimental conditions in the best grades received for strategy applications.

Table 28 shows how mean grade, best grade and voluntariness of executing the reading strategies *summarizing*, *questioning* and *predicting* correlate with children's cognitive abilities. These correlations give an impression of whether the quality of strategy application is determined by children's cognitive abilities. Few significant correlations were observed. The better children's nonverbal reasoning abilities, the better their mean and their best grades for *questioning* (r = -.35), and children with higher verbal ability were more likely to volunteer to execute the *questioning* strategy

(r = .31). It is, however, very important to note that success in applying the reading strategies (mean grade and best grade) did not, in general, depend on children's verbal cognitive abilities; there were no significant correlations with mean verbal ability or reading speed.

Table 28. Correlations between strategy application and cognitive abilities.

		Nonverbal reasoning ability	Mean verbal ability	Reading speed
Mean grade	Summarizing	-.16	-.08	.01
	Questioning	-.35 **	-.13	.02
	Predicting	-.16	-.12	-.08
Best grade	Summarizing	-.02	-.07	.07
	Questioning	-.29 *	-.06	.06
	Predicting	---	---	---
Volun-tariness	Summarizing	.19	.15	.08
	Questioning	.18	.31 **	-.18
	Predicting	.09	.16	-.26

3.3.1.5. Fulfillment of roles in reciprocal and monitor conditions

Were the children in *reciprocal* and *monitor* groups able to fulfill the roles of "teacher" and "monitor" adequately? Items were included in both session and paragraph analysis to provide information about the fulfillment of role requirements and to assess how well the "teacher" or "monitor" guided the correction of answers, if necessary (with scales ranging from a minimum of 1 to a maximum of 4). The data reported in this section were derived from paragraph analysis, because more observations are available at this level (N = 86 for session analysis; N = 207 for paragraph analysis). Data were aggregated for each child; it was not possible to compute developmental trends over the course of the training program because of the limited number of observations for each child (for "teachers" in *reciprocal* groups: M = 5.00 with SD = 2.81, min = 2 and max = 12; for "monitors": M = 10.25 with SD = 3.68, min = 2 and max = 16).

In both *reciprocal* and *monitor* conditions, children met the role requirements well to very well. The mean for "teachers" was 3.08; for "monitors" it was 3.68 (with SDs of .41 and .38, respectively). "Teachers" did not score well on providing "good'" corrective feedback and guiding the other students to correct their answers, however (M = 1.77, SD = .42, N = 17), and "monitors" were almost incapable of accomplishing

this task (M = 1.06, SD = .10, N = 20). Neither of these two tasks correlated significantly with students' cognitive abilities (see Table 29).

Table 29. Correlations between role fulfillment and cognitive abilities.

		Nonverbal reasoning ability	Mean verbal ability	Reading speed
Reciprocal Teaching	meeting role requirements	-.04	-.03	.36
	guidance to correct answer	.46	.45	-.31
Monitor	meeting role requirements	.11	-.23	.02
	guidance to correct answer	.18	-.22	-.09

3.3.1.6. Summary

What can be concluded from this detailed analysis of the type and duration of the different activities that took place during the training sessions? The main reason for investing so much effort in this analysis was to show that time was spent in similar ways in all experimental conditions. After thorough analyses of the available data, it can be concluded that, despite the small differences due to the nature of experimental conditions (instruction of roles, which took on average less than one minute per session), all groups spent similar amounts of time practicing the reading strategies during training. The allocated study time was spent as intended: organizational matters and discipline problems took less than 5 minutes per session and, on average, more than 50 minutes per session were spent on instruction, activities that serve to practice the reading strategies, and test-taking. The reading strategies were introduced in 3 instruction sessions and then practiced for 12 sessions. Additional instruction on the strategies by the trainer was only necessary in the earlier practice sessions. In *reciprocal* and *monitor* groups, it took the trainers an average of less than one minute per session to instruct the students on their roles. One interesting difference between experimental conditions was observed for the time that trainers spent with restoring discipline: whereas in the *reciprocal* condition, the amount of time dedicated to this activity decreased over the course of the training program, in the *monitor* and *student* conditions, it increased.

It is not only the detailed analyses of the different activities taking place during the training sessions that provide support for the efficient use of study time, the raters' judgments of individual children's time on task substantiate these findings. The amount

of time that children were observed to be "on track" was very high, with means of about 95 percent of the time.

There were considerable differences between the three experimental conditions in terms of pace and (associated with this) the number of opportunities to practice the reading strategies; most text paragraphs were completed in the *student* condition (M = 48), followed by the *monitor* condition (with a mean of 43.5). Groups in the *reciprocal* condition finished the fewest text paragraphs (on average about 37). Accordingly, the number of strategy applications recorded on video differed between experimental conditions.

Questioning was the most difficult strategy, being given a mean grade (by the video raters) of just 3.01, followed by *summarizing* with a mean of 2.59. It was obviously very difficult to judge the quality of *predictions* the children made; children received very good grades for this strategy (M = 1.37). There were few to no differences between experimental conditions in the grades that children received for strategy application. Moreover, the grades awarded did not improve with training. Children volunteered most often for *predicting*, somewhat less often for *questioning* and very rarely for *summarizing*. Their success in applying the strategies during training was not dependent upon their verbal abilities; significant correlations were not observed for either mean verbal ability or reading speed.

Children in *reciprocal* and *monitor* groups had special tasks to carry out: they had to fulfill the roles of "teacher" and "monitor." In both cases, they were very successful in meeting the organizational requirements of the roles (organizing the sessions and providing feedbackand guiding the correction of answers). However, "teacher" and "monitor" students did not do very good jobs when it came to guiding the correction of answers with respect to quality of guidance. Again, in role fulfillment did not depend on children's cognitive abilities.

3.3.2. Effects of strategy training on knowledge and performance measures

Turning to the comparison between *control* children and children who participated in the training program, analyses of all indicators that measure aspects of metacognitive knowledge (*relational strategy knowledge*, *planning knowledge* and *conditional knowledge* about reading strategies, and declarative knowledge about the strategies *summarizing* and *clarifying)* will now be reported before analyzing the

differences between experimental conditions on performance measures, i.e., production of a summary and text comprehension.

3.3.2.1. Knowledge measures

3.3.2.1.1. Specific strategy knowledge about summarizing and clarifying

Differences between experimental conditions in terms of knowledge about the two specific reading strategies *summarizing* and *clarifying* were analyzed by computing analyses of variance with contrast setting A for the between-subjects factor experimental condition. This means that, firstly, the *reciprocal* and *monitor* conditions were contrasted with the *student* condition and, secondly, the *reciprocal* and the *monitor* conditions were compared. Table 30 presents means and standard deviations for measures of declarative knowledge about *summarizing* and *clarifying*.

Table 30. Knowledge about summarizing and clarifying by experimental condition.

Strategy		Reciprocal M	Reciprocal SD	Monitor M	Monitor SD	Student M	Student SD
Summarizing	contains most important content	.94	.24	.95	.22	.67	.48
	is shorter than the text	.88	.33	.70	.47	.78	.43
	is formulated in own words	.35	.49	.50	.51	.22	.43
	Total N of characteristics	2.18	.39	2.15	.87	1.67	.84
Clarifying	external: ask other people to help	1.12	1.36	1.00	.97	1.22	.94
	other resources (dictionaries, etc.)	.71	.59	1.00	1.17	.94	.64
	text-related strategies	1.59	.94	1.60	1.05	.94	.87

Where the *number of characteristics of a good summary* identified by the children is concerned, the first condition contrast was highly significant: students in the *reciprocal* and m*onitor* conditions wrote down more than two features that characterize a good summary, whereas children in the *student* condition only produced between one and two characteristics ($F(1,52) = 5.34$, $MSe = .56$, and $p = .025$ for condition contrast(1)). There were no significant differences between the *reciprocal* and *monitor* conditions. When the three characteristics were analyzed separately, the significant condition(1) contrast in favor of the *reciprocal* and *monitor* conditions was only found for the feature "contains the most important content" ($F(1,52) = 8.30$, $MSe = .11$, and $p < .001$). The experimental conditions did not differ with respect to the other two characteristics of good summaries. The effect size for the feature "contains most

important content and no details" was moderate, at .74, as was the effect size for the total number of characteristics of a good summary (d = .64).

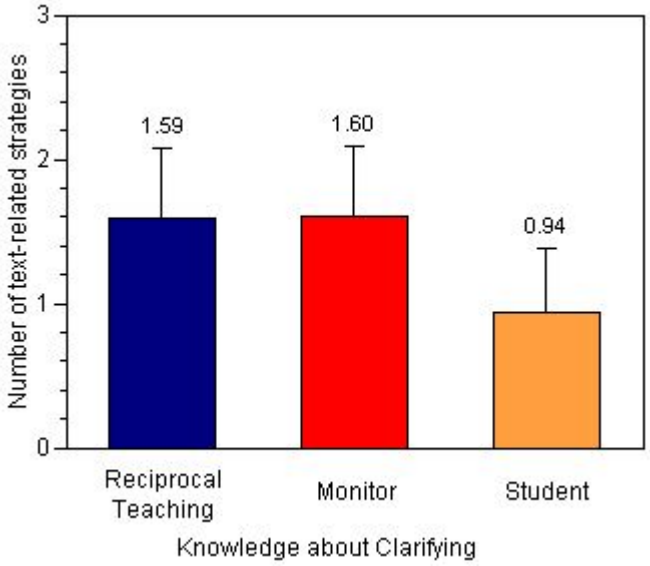


Figure 9. Number of text-related strategies produced for clarifying by experimental condition.

Inspection of the means and standard deviations of the number of strategies produced for *clarifying* shows that there were no reliable differences between experimental conditions in terms of the number of *external strategies* named (all $p > .55$). Highly reliable differences emerged for the use of *text-related strategies*: while children in the *reciprocal teaching* or *monitor* conditions reported about 1.6 strategies that involved using the text to clarify, children in the *student* condition only could think of nearly one text-related strategy ($F(1,52) = 8.30$, $MSe = .11$, and $p < .001$). At .70, the effect was moderate to strong, and is clearly apparent in Figure 9.

3.3.2.1.2. Relational strategy knowledge

When the experimental conditions were compared in terms of students' scores on *relational strategy knowledge*, a significant main effect of time ($F(1,49) = 4.68$, $MSe = .04$, $p = .035$), which needs to be interpreted in the light of the significant interaction between the first condition contrast and time ($F(1,49) = 5.58$, $MSe = .04$, $p = .022$), was found. The corrected effect size was .80. As can be seen in Figure 10, *relational strategy knowledge* about text comprehension increased after training for students in the *reciprocal teaching* and *monitor* conditions. No effect of training was found for children in the *student* condition; in fact, a slight decrease is visible here.

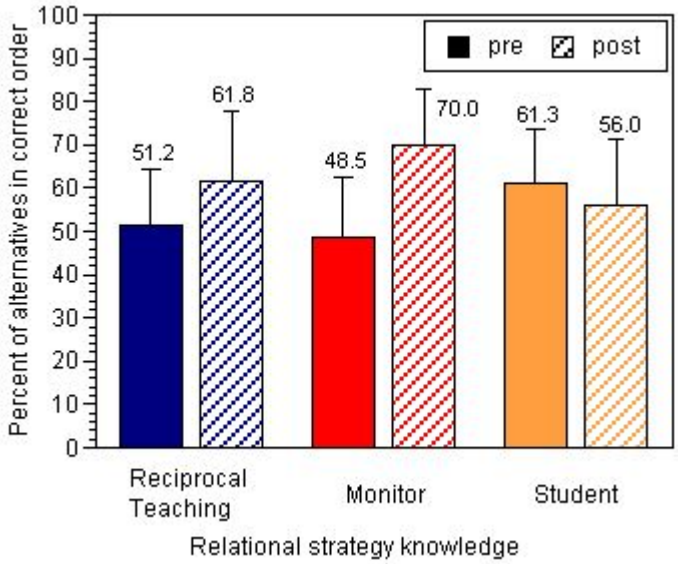


Figure 10. Relational strategy knowledge by experimental condition.

3.3.2.1.3. Conditional strategy knowledge and planning knowledge

The two subscales of the IRA questionnaire were also subjected to multivariate analyses of variance, with time as within-subject factor and contrast setting A for the between-subjects factor condition. Means and standard deviations for the three experimental conditions are listed in table 31.

For *conditional strategy knowledge*, the only effect observed was a main effect of time that just failed the significance level ($F(1,52) = 3.94$, $MSe = 2.72$, $p = .052$). All children scored higher after training than before. The differences between the three conditions were not significant, and there were no interactions between the condition contrasts and time.

Analyses of *planning knowledge* for reading activities did not reveal any significant effects.

Table 31. Scores for conditional strategy knowledge and planning knowledge by experimental condition.

		Reciprocal		Monitor		Student	
		M	SD	M	SD	M	SD
Conditional strategy knowledge	pre	5.82	1.78	6.30	1.81	5.22	1.63
	post	6.00	1.73	7.00	1.52	6.22	1.44
Planning knowledge	pre	6.53	1.94	6.00	1.49	5.78	1.17
	post	6.00	1.70	6.20	1.88	6.61	1.46

3.3.2.2. Performance measures

3.3.2.2.1. Writing a Summary

Ratings data for all features of the summaries written at post-test (means and standard deviations) are shown in table 32. A lot of the children *underlined the original text* when producing their summary. It looks as if there was a significant difference for frequency of underlining between children in the *reciprocal* and *monitor* conditions (29 and 15 percent) and those in the *student* condition (only about 5 percent); however, this effect was not statistically significant ($F(1,49) = 1.86$, $MSe = .14$, $p = .16$). The size of this effect was moderate, at .47.

Results of analyses of variance of the percentage of students who *generated a title* are difficult to interpret because none of the children in the *student* condition generated a title, whereas about 10 percent of students in both the *reciprocal* and the *monitor* conditions generated titles for their summaries. The condition(1) main effect was not significant (p = .15), but because there is no variance in the *student* condition, this finding cannot be interpreted. I will argue that there was a difference in favor of the *reciprocal* and *monitor* conditions. An effect size could not be computed because of the zero-variance in the *student* condition.

Table 32. Means and standard deviations for features of the summaries by experimental condition.

	Reciprocal		Monitor		Student	
	M	SD	M	SD	M	SD
Generation of a title (% of students)	11.84	33.21	10.01	30.85	.00	.00
Underlining in text (% of students)	29.41	47.03	15.03	36.64	5.62	23.64
Number of words	39.41	11.49	41.25	11.31	45.17	17.47
Number of main ideas	6.47	2.35	5.85	1.78	6.67	2.52
Number of details	10.00	5.52	10.10	6.49	10.06	6.52
Ratio between main ideas / details	.88	.65	1.34	1.88	1.04	1.55
“Grade a teacher would give”	4.41	.94	3.95	1.00	4.33	.91

The experimental conditions did not differ in terms of the number of words in the summaries. Similarly, there were no differences between experimental conditions in the “grade a teacher would give.”

The content of students’ summaries was also analyzed. Differences between experimental conditions were not found for any of these measures (number of main ideas, number of details and ratio between main ideas / details).

3.3.2.2.2. Text comprehension

In the following, differences between training and *control* group children in terms of measures of text comprehension will be reported. Table 33 shows text comprehension scores for the children in the three experimental conditions.

First, the number of points awarded is analyzed. Children received more points for their answers at post-test than at pre-test ($F(1,52) = 113.66$, $MSe = 14.90$, $p < .01$). This was true for all competence levels (for level 1: $F(1,52) = 6.09$, $MSe = 2.40$, $p < .05$; for level 2: $F(1,52) = 117.54$, $MSe = 9.33$, $p < .01$; for level 3: $F(1,52) = 15.40$, $MSe = 1.18$, $p < .01$). No differences between experimental conditions were observed.

Table 33. Text comprehension scores for children by experimental condition.

			Reciprocal (17)		Monitor (20)		Student (18)	
			M	SD	M	SD	M	SD
Points	level 1	pre	2.06	1.82	1.88	1.49	2.33	1.91
		post	2.47	1.53	3.10	2.06	2.89	1.63
	level 2	pre	4.59	3.02	5.60	2.64	4.72	3.74
		post	10.18	4.42	12.50	4.25	11.22	6.54
	level 3	pre	.35	.49	.50	.61	.50	.79
		post	1.29	1.27	1.20	1.54	1.28	1.07
	total	pre	7.00	4.65	7.98	3.60	7.56	5.15
		post	13.94	5.28	16.80	5.87	15.39	7.76
Content	main ideas	pre	8.12	4.82	8.70	5.03	6.72	4.69
		post	10.94	5.36	12.55	4.57	11.22	5.53
	details	pre	1.77	2.19	2.20	2.17	2.67	3.58
		post	3.76	2.25	5.75	3.59	6.56	6.17
	ratio main ideas/details	pre	4.99	3.60	4.96	3.56	3.52	3.17
		post	3.79	2.82	3.36	2.87	2.23	1.13

Analyses of the content of children's answers showed that all children included more important content after training than before ($F(1,52) = 28.30$, $MSe = 13.42$, $p < .01$). However, the children also included more details in their answers; this was reflected in a highly significant main effect of time ($F(1,52) = 32.82$, $MSe = 8.26$, $p < .01$). The ratio of main ideas to details is in fact more informative here. The children included relatively more details than main ideas in their answers after training, as reflected in a significant main effect of time ($F(1,52) = 5.37$, $MSe = 9.44$, $p < .05$). The answers of children in the *reciprocal* and *monitor* conditions contained relatively more important content than those of children in the *student* condition (main effect of

condition(1) with contrast setting A $F(1,52) = 5.71$, $MSe = 8.35$, $p < .05$), but these differences were already present at pre-test.

3.3.2.3. Summary of findings

Differences between experimental conditions were observed for many of the various measures of knowledge about reading strategies. Children in the *reciprocal* and *monitor* conditions obtained higher scores (which did not differ significantly from each other) than children in the *student* condition in the following measures: the *total number of characteristics of a good summary* identified, the specific feature *"contains the most important content and no details,"* and the number of *text-related strategies* listed for *clarifying*. For children in the experimental conditions *reciprocal teaching* and *monitor***,** training had also positive effects (increments) with respect to *relational strategy knowledge*; children's knowledge in the *student* condition did not increase after training, however. These effects were moderate to strong.

Summarizing the results for performance measures, it is evident that there were very few to no differences between the three experimental conditions. When analyzing the summaries that the children wrote at post-test, differences were found in the frequency of *underlining* in the original text and *generation of a title* in favor of the *reciprocal* and *monitor* conditions. Nevertheless, these differences failed to reach the significance level of $p > .05$. The reasons for this are the small sample size (lack of statistical power) and, where generation of a title is concerned, lack of variance in the *student* condition. For *text comprehension scores*, differences between the experimental conditions were only observed for one measure, the *ratio of main ideas to details*, where children in the *reciprocal* and *monitor* conditions scored higher than children in the *student* condition at both pre- and post-test.

3.3.3. Children's perception of the training program

This last section will explore how the children themselves perceived the training program. Did they enjoy the program? How much did they like working together as a group? Did they believe that they had learned something important? How exhausting was the training for them? Children were asked about their motivation at the end of the training. Means and standard deviations of the scales are displayed in table 34. Please note that, except for the question of whether the children would participate in such a training program again, values on the response scales could range from 1 to 4.

The data confirmed the trainers' impressions that the children enjoyed the training program; they reported high *enjoyment of the training program* as well as a high endorsement of *working together with other children in a group*. Some of the children judged the strategy training to be *exhausting*; others did not – resulting in an intermediate level of exhaustion. Nearly all of the children thought that the *skills* that they had learned would *be helpful or very helpful in school*. A very high percentage of them also said they would *participate again* in such a training program. There was only one significant difference between the experimental conditions: children in the *reciprocal* and *monitor* groups reported higher enjoyment values than children in the *student* groups (significant main effect of condition(1) for setting A with $MSe = .17$, $F(1,52) = 8.30$ and $p < .01$ with an effect size of .75).

Table 34. Children's perception of the training program.

	Reciprocal		Monitor		Student	
	M	SD	M	SD	M	SD
Enjoyment of the training program (4 items)	3.52	.32	3.74	.26	3.29	.58
Enjoyment of the group (2 items)	3.35	.49	3.30	.52	3.28	.58
How exhausting was the training program?	2.24	.90	2.20	.83	1.17	.62
Belief that skills will be helpful in school	3.41	.51	3.50	.51	3.18	.73
Would you participate in such a training program again?	82.4 %		100 %		72.2 %	

Children were also asked to write down what they thought they had learned during the training program. Half a page in the booklet was left free for their responses, and they were not prompted in any way. Students' answers were then sorted and categories were specified. The most frequent categories are listed in table 35. Other answers were given only very rarely and are therefore are not listed in the table (they concerned discipline in the group, fun, being able to make mistakes and the social atmosphere).

Many of the children noted that they had learned reading strategies, and listed some or all of the four specific strategies. There were no differences between experimental conditions in terms of the percentage of children mentioning reading strategies in general. The *number of specific reading strategies* listed did vary systematically, however: children in the *reciprocal* and *monitor* groups listed more of the four strategies than their peers in the *student* groups ($MSe = 1.94$, $F(2,52) = 4.78$

and $p < .05$, with d = .84). Separate analyses of the four strategies revealed that the largest differences between experimental conditions were observed for the strategies *summarizing* and *predicting* (with effect sizes of 1.02 and .68, respectively).

Improvements in reading abilities and general education were also mentioned, however. Half of the children in *monitor* groups wrote that they *learned to cooperate with other children*; the same was noted by about one quarter of the children in the *reciprocal* condition and only 11 percent of the children in the *student* condition. These differences were significant ($MSe = .30$, $F(2,52) = 3.23$ and $p < .05$). It is also very interesting that some of the students in *reciprocal* and *monitor* groups mentioned that they had learned to look at their learning processes from another perspective.

Table 35. Classification of children's answers to the question "What did you learn in the training program?".

	Reciprocal		Monitor		Student	
Mention of	M	SD	M	SD	M	SD
Specific reading strategies (sum) *	2.59	1.37	3.05	1.40	1.67	1.41
Summarizing *	.82	.39	.85	.37	.39	.50
Questioning	.77	.44	.80	.41	.56	.51
Predicting *	.41	.51	.65	.49	.22	.43
Clarifying	.59	.51	.75	.44	.50	.51
Unspecified reading strategies	.18	.39	.20	.41	.39	.50
Improvement of reading abilities	.53	.51	.50	.51	.50	.51
Improvement of general education	.17	.39	.15	.37	.06	.24
Cooperation with other students **	.24	.44	.40	.50	.11	.32
Change of perspective	.12	.33	.15	.37	.00	.00

At the very end, the children had the opportunity to write down what they liked and disliked most about the training program. Almost half of the children in the *reciprocal teaching* condition wrote that they very much enjoyed acting as teacher; only one child particularly disliked taking the teacher's role. Surprisingly, none of the *monitor* children mentioned anything about evaluating other students' answers. Some of the children mentioned that they particularly liked one of the strategies (*summarizing* - 4 children, *questioning* – 2, *predicting* – 2 and *clarifying* – 3), while some children wrote that they disliked a particular strategy (*summarizing* - 2 children and *predicting* – 1 child). Besides this, one or more children also mentioned that their relations to other children improved, that they liked working in a small group, that the training was fun and that they liked particular children in their group.

The groups consisted of boys and girls of all ability levels. In contrast to their regular lessons with their classmates, the children learned together with students from other classes. I was interested to find out how well the children got along with each other. Therefore, in the post-training questionnaire, they were asked how much they liked working together with the other children in their group, and gave a rating between 1 (*very much*) and 6 (*not at all*) for each of the other children. These ratings are depicted in figure 11.

Interestingly, the ratings varied systematically with sex: children of the same sex were given better ratings than children of the opposite sex (MSe = .84, $F(1,40) = 33.02$ and $p < .01$). This was not the only highly significant two-way interaction, however; the difference between ratings for children of the same sex and the opposite sex was larger in the *student* condition than in the *reciprocal* and *monitor* conditions ($F(1,40) = 8.26$ and $p < .01$; contrast setting A).

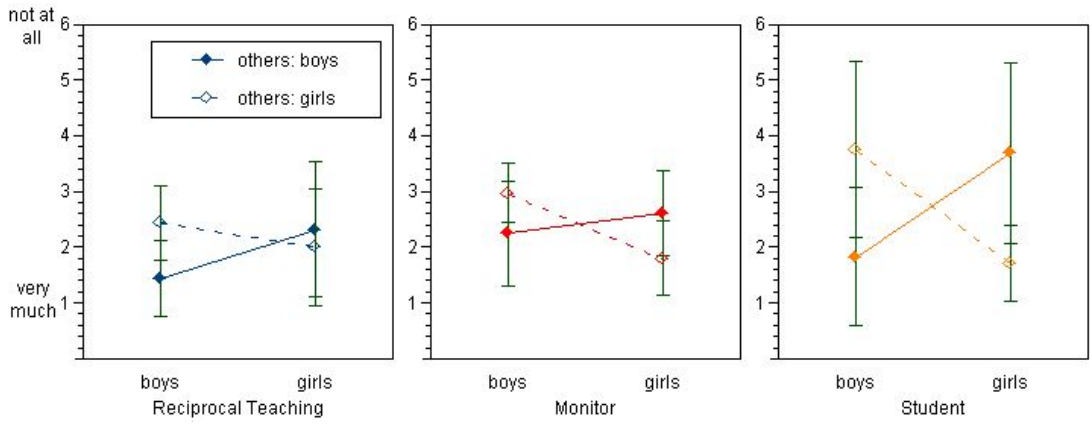


Figure 11. Mean ratings of enjoyment of working with the other children in the group.

3.4. Comparison between control group and student condition

Only two major comparisons have been made for the purpose of testing the hypotheses: the first between the *reciprocal* and *monitor* conditions, on the one hand, and the *control* group, on the other, and the second between the *reciprocal* and *monitor* conditions, on the one hand, and the *student* condition, on the other. The rationale for focusing on the *reciprocal* and *monitor* conditions and considering only these two conditions as "real" training groups was that they were assumed to incorporate the features responsible for the large effects of the Reciprocal Teaching method (e.g., monitoring tasks), whereas the *student* condition did not. It could be shown that training helped to enhance children's metacognitive knowledge and that children who

participated in the training program outperformed *control* children on some of the performance measures. These results hold for the *reciprocal* and *monitor* conditions. It was also demonstrated that training was more beneficial for children in *reciprocal* and *monitor* conditions than for children in the *student* condition. However, the question of whether the *student* condition differed at all from the *control* condition, and – if this was the case – how large the effects were, has not yet been addressed. The answers to this question have important implications for transferring the research results to everyday practice and giving teachers advice as how to design lessons and teach reading strategies. If teaching reading strategies in the way done in the *student* condition produces large positive effects relative to the *control* condition, and the differences between the different methods (experimental conditions) of teaching the strategies are very small, the method used to teach strategies might not be so important. In other words, the important thing would be THAT strategies are taught, and not necessarily HOW they are taught. In order to address these questions, children in the *student* condition were compared with the *control* group with respect to all measures of knowledge and performance. Most of the data for this comparison have already been presented above; therefore, this section will be rather short and concentrate on the important findings for this comparison.

Where the measures of children's knowledge are concerned, significant differences were only found between the children who participated in the training program in *student* groups and the *control* children, with the *student* group children outperforming the *control* children on their knowledge about *summarizing*. Children in the *student* condition could identify more characteristics of a good summary (mean 1.67) than *control* children (mean 1.00; $F(1,100) = 8.45$, $MSe = .78$, and $p < .01$). When the three features of knowledge about a good summary were analyzed separately, this effect was found only for the awareness that a good summary should be shorter than the text ($F(1,100) = 11.71$, $MSe = .22$, and $p < .01$; with 35.7% of *control* children and 77.8% of children in the *student* condition identifying this feature). The two groups did not differ significantly in terms of their awareness that a good summary should only contain important ideas (51.2% of *control* and 66.7% of *student* children) or that it should be formulated in one's own words (13.1% and 22.2%, respectively). Differences between the *control* group and the *student* condition were not found for any of the other knowledge measures; i.e., text-related strategies, use of objects, or asking other people for help to *clarify* the text (.87/.94, .89/.94, and 1.13/1.22 strategies identified by

children in the *control* group and the *student* condition, respectively), *relational strategy knowledge* (see Figures 3 and 10 for the corresponding data), and *planning knowledge* (see Tables 16 and 31). *Control* children outperformed their peers in the *student* condition in *conditional strategy knowledge* at both pre- and post-test ($F(1,96) = 6.03$, $MSe = 3.70$, and $p = .016$).

Moreover, there were very few differences in measures of children's performance. Significant differences, with children in the *student* condition outperforming *control* children, were only found for the number of details in the summaries ($F(1,99) = 4.73$, $MSe = 48.03$, and $p = .032$; with 10.0 vs. 13.9 details). Children in the *student* condition wrote shorter summaries than *control* children (45.2 vs. 53.8 words), but this difference just failed to reach significance ($F(1,99) = 3.87$, $MSe = 282.73$, $p = .052$). No differences were observed for any of the other measures of the summaries written (i.e., grade awarded for summary, percentage of children who underlined text or produced a title, linguistic style of the text, number of main ideas, and ratio of main ideas to details). No differences in favor of the children in the *student* condition were found in the reading comprehension assessments either.

To summarize the findings, few differences were observed between children who were given reading strategy training in a manner similar to traditional forms of classroom instruction (*student* condition) and *control* children, who devoted their afternoons to homework, play, and other leisure activities, as usual. Differences in favor of children in the *student* condition were restricted to a rather unimportant surface feature of a good summary – that it should be shorter than the original text – and were observed on the levels of both knowledge and performance. In other words, relative to the *control* group, more children in the *student* training condition could identify this feature at post-test and, on average, they wrote shorter summaries containing fewer details.

4. DISCUSSION

4.1. Research questions

I would like to start this section by re-stating the research questions. Because the second question draws upon the assumption that the first hypothesis holds, the two questions cannot be answered separately, but need to be considered together.

The first question concerns the mechanisms behind the large effects of the Reciprocal Teaching method on reading comprehension. I assumed that *more general metacognitive knowledge and metamemory acquisition procedures are learned in Reciprocal Teaching, resulting not only in a better ability to apply the reading strategies taught in a routine manner, but in mindful application of strategies in general, and in monitoring and regulating activities.* The second research question is based on the assumption that the first hypothesis holds, and aims to identify the effective features of the Reciprocal Teaching method. *It is assumed that the content-related tasks associated with the "teacher" role are those that lead to the acquisition of metacognitive knowledge and to enhanced reading comprehension. By monitoring other students' understanding of the text and application of reading strategies, children learn a great deal about these strategies as well as about the evaluation of understanding. Moreover, in having to help their peers, they go beyond monitoring and evaluation and acquire knowledge and skills about regulative activities. ... These metacognitive skills (reflecting on one's own thinking and, if necessary, modifying it) first occur inter-individually and can then become intra-individual by internalization.* In short, the mechanism in effect was hypothesized to be the acquisition of metacognitive knowledge and skills, and – provided that this assumption holds – the content-related tasks of the "teacher" role are hypothesized to lead to the activation of that mechanism by practicing inter-individually what later becomes intra-individual by internalization.

4.2. Interpretation of results

4.2.1. Hypothesis 1: The mechanisms in effect

The success of Reciprocal Teaching is not merely due to the more frequent use of the strategies instructed, but to the metacognitive knowledge and skills acquired through the procedure. Differences are expected not only in performance, but also in

the amount of metacognitive knowledge acquired: students who participate in strategy training are expected to outperform children in the control group.

To test this hypothesis, the *control* group was compared with children who were trained to apply reading strategies in either the *reciprocal teaching* or the *monitor* condition with respect to measures of metacognitive knowledge and various performance measures. Dependent variables were subjected to analyses of variance and multiple hierarchical regression analyses; both types of analysis yielded similar results.

Findings were as follows: There were large to very large effects (with effect sizes of between .70 and 1.13) of strategy training on knowledge about the *summarizing* and *clarifying* reading strategies that were taught in training. Children in the training groups could name more characteristics of a good *summary* than could control children. Where knowledge about *clarifying* was concerned, trained children only outperformed their peers in the control group on the number of text-related strategies identified, and not on the number of external strategies (e.g., using objects, help of others). For more general measures of metacognitive knowledge (i.e., *relational* and *conditional strategy knowledge*, *planning knowledge*) no differences were observed between the *control* group and the *reciprocal teaching* and *monitor* training conditions. For *relational strategy knowledge*, a significant effect of time was found: all children knew more about the relative usefulness of reading strategies at post-test. It also looked as if the trained children "caught up"; their performance was lower than that of *control* children prior to training, but at about the same level as the *control* group after training. However, this effect (an interaction between time and training) with a corrected effect size of .28 just failed to reach significance.

Medium to large effects of training were observed for the *summary* of a short text passage that students were required to write at post-test, with effect sizes of between .34 and .95. Children who had participated in strategy training (in the *reciprocal* and *monitor* conditions) wrote shorter *summaries* containing relatively fewer details. When main ideas and details were analyzed in relation to one another (ratio), the finding that training group children wrote relatively more important than non-important content in their *summaries* ($p = .066$ in an analysis of variance and $p = .057$ in a hierarchical multiple regression analysis; corrected effect size of .33) was almost statistically significant. In addition to the content of the *summaries*, more "overt signs" of students' metacognitive activity were analyzed: it was noted whether children had underlined text in the test booklet while producing their *summary* and whether they had

generated a title. Indeed, more children who participated in *training* underlined text, and more of them generated titles for their *summaries* than did *control* children.

However, training failed to have significant effects on children's performance in the *text comprehension* assessments. *Control* children outperformed children in the *reciprocal* and *monitor* conditions at both times of testing. At post-test, all children wrote a lot more than prior to training, irrespective of whether they had participated. Where measures of children's knowledge and performance were concerned, the effects remained the same, even after controlling for pre-test performance and cognitive abilities. Pre-test scores were always predictive of performance at post-test. Mean verbal ability and reading speed explained significant portions of variance.

Hypothesis 1 was thus confirmed for measures of metacognitive knowledge more closely related to the training program, but not for more general metacognitive knowledge measures. The same holds for performance measures: positive effects of training were found for a measure directly related to training, but the hypothesis has to be rejected where the transfer measure of reading comprehension is concerned.

4.2.2. Hypothesis 2: Features of the program

It is assumed that the content-related tasks associated with the "teacher" role are those that lead to the large improvement observed in metacognitive knowledge and skills in Reciprocal Teaching programs. The reciprocal and monitor conditions implemented in the present study are assumed to be superior to the student condition with respect to strategy execution, increase in metacognitive knowledge, and text comprehension.

Children in the *reciprocal* and *monitor* conditions acquired more knowledge about the reading strategies *summarizing* and *clarifying* than did children in the *student* condition: they could identify more features of a good summary and think of more text-related strategies to clarify the meaning of unknown words or sentences that were not yet understood. Large effect sizes of about .70 were observed for these measures. The fact that students in *reciprocal* and *monitor* conditions acquired more knowledge about the reading strategies instructed was also reflected in their answers to the question "What did you learn in the training program?" included in a post-training questionnaire. Here, they mentioned significantly more of the specific reading strategies taught in the training program than did children in the *student* condition (effect size .84). Another difference in favor of the *reciprocal* and *monitor* conditions was found for *relational strategy knowledge* (with an effect size of .80): an increase was observed from pre- to

post-test for children in the *reciprocal* and *monitor* conditions, but a slight decrease in *relational strategy knowledge* was recorded for those in the *student* condition. No differences between the three experimental conditions were observed for *conditional strategy knowledge* or *planning knowledge*.

Differences between the *reciprocal* and *monitor* conditions, on the one hand, and the *student* condition, on the other, were observed as soon as the children practiced the reading strategies during training: the quality of the teacher-like questions posed by the children increased over the course of the training in the *reciprocal* and *monitor* conditions, but stayed the same in the *student* condition. Other signs of superior metacognitive activities occurring in the *reciprocal* and *monitor* conditions during production of the summaries were that more children underlined text in their test booklets (with an effect size of .47, though not statistically significant) and that more children in these groups produced a title for their summary (about 10% of the students in the *reciprocal* and the *monitor* condition compared with none of the children in the *student* condition).

The three experimental conditions did not differ significantly in terms of the performance measures assessed after training; this held for both the content of the children's summaries and the comprehension assessments.

To summarize, the effects of the *reciprocal* and *monitor* conditions were very similar, and very different to the effects of the *student* condition. The *reciprocal* and *monitor* conditions did not differ significantly in any of the knowledge or performance measures. This was true even though children in the *student* condition had the most learning opportunities: they were able to carry out each of the strategies and observe other models more often because they worked through more text paragraphs during training. More differences were observed for measures of knowledge than for performance measures. In sum, hypothesis 2 can be confirmed for measures of children's metacognitive knowledge, but has to be rejected for performance measures.

4.2.3. Hypothesis 3: Motivational effects

The fulfillment of the psychological needs of experiences of competence and self-determination and the desire for social relationships were expected to result in increased motivation. The motivation of children should be highest in the reciprocal teaching condition, followed by the monitor condition. The lowest motivational effects were expected in the student condition.

There were few, but interesting differences between the three experimental conditions in terms of motivational measures. Statistically reliable differences in favor of children in the *reciprocal* and *monitor* conditions were found for enjoyment of the training program, which they rated higher than did children in the *student* condition. Another difference concerned the question "What did you learn in the training program?" Half of the children in the *monitor* condition wrote that they had learned to cooperate with other children; the same held for about a quarter of children in the *reciprocal* condition. None of the children in the *student* condition made similar comments. Another interesting result was that, although all children preferred working which children of their own sex than with children of the opposite sex, the differences between these ratings were less pronounced in the *reciprocal* and *monitor* conditions than in the *student* condition.

Hypothesis 3 has to be rejected for all but one measure of motivation ("learned to cooperate"). Whenever any differences in motivational measures were observed, children in both the *reciprocal* and the *monitor* conditions reported higher ratings than did children in the *student* condition. However, although children's motivation in the *reciprocal* condition was predicted to be superior to that of children in the *monitor* condition, there was only one reliable difference; in all other motivational measures (i.e., enjoyment of the training program and ratings of liking) the ratings of children in those two conditions were very similar and did not differ statistically significantly.

4.3. Answering the research questions

Because "metacognition in action" is not directly observable, we have to rely on performance indicators to address the question of whether metacognition is in fact the mechanism causing the large improvements in reading comprehension that are usually observed with the Reciprocal Teaching method and, even more indirectly, to test the second assumption that it is the "teacher" role that produces the observed benefits by comparing the experimental conditions. The results that are relevant for answering these questions are reviewed at this point. In measures of knowledge and performance more closely related to the content of the training program (knowledge about specific reading strategies; application of the strategies taught), significant differences emerged in favor of children who participated in strategy training in the *reciprocal* or *monitor* conditions relative to the *control* group. No significant differences were observed for more distal measures, but some of the findings were in the predicted direction

(i.e., children in the *reciprocal* and *monitor* conditions "caught up" with control children in terms of *relational strategy knowledge* and the *ratio of important ideas to details* in their summaries). Fewer differences were observed for the comparison between the three experimental conditions. Very interestingly, the results for children's knowledge and performance measures in the *reciprocal* and *monitor* conditions were very alike and did not differ significantly; children in both of these conditions outperformed their peers in the *student* condition, even though the number of learning opportunities for every child in this condition was higher because the *student* groups worked through more text paragraphs. This was true for strategy knowledge on both *summarizing* and *clarifying*, for underlining text in the test booklets during production of the summary, and for the percentage of children who gave their summary a title.

Another finding was that children who participated in strategy training in the *student* condition, which was most comparable to the regular classroom setting, performed only marginally better at post-test than did *control* children, who spent the same time on homework, watching television, doing sports, or playing with friends. Children in the *student* condition only outperformed children in the *control* group on their knowledge about the least important feature of a good summary (that it should be shorter than the text), and by including fewer details in their summaries.

What are the conclusions to be drawn from these results with respect to my research questions? Although significant results were only obtained for measures closely related to strategy training – strategy knowledge and strategy application – and not (yet) for more general components of metacognitive knowledge, the comparison between the experimental conditions provided strong support for the assumption that metacognitive knowledge and skills are acquired when the Reciprocal Teaching method is used to teach reading strategies. Both conditions that involved children giving each other feedback on performance, the *reciprocal* and *monitor* conditions, produced similar results: not only did students in these conditions acquire more knowledge about the reading strategies *summarizing* and *clarifying*, they also applied the *summarizing* strategy better than the *control children* and the children in the *student* condition.

These findings are consistent with Borkowski, Milstead and Hale's componential theory of metamemory (1988). The authors propose that specific strategy knowledge, which is at the center of their model, is a prerequisite for higher order components that, in turn, aid further acquisition of strategy knowledge. In this study large effects of training were found for specific strategy knowledge, but only small (and

not quite significant) effects for *relational strategy knowledge*. In Borkowski's model, there is a triangle of bidirectional relationships between actual strategy use (practice), specific strategy knowledge, and metamemory acquisition procedures. In monitoring other students' performance, students learn and have the opportunity to practice metamemory acquisition procedures. This assumption is supported by the finding that the *student* condition produced little to no effects in relation to the *reciprocal* and *monitor* conditions for either strategy knowledge or the ability to apply a reading strategy.

In these respects, the cooperative setting of the Reciprocal Teaching method seems to promote internalization of inter-individual social processes; the dialogues that occur in the group help the students to acquire metamemory acquisition procedures and strategy knowledge. Internalization is the basis of real "meta-cognition," which enables humans to plan and regulate their activity and derives from previous participation in social settings (Vygotsky, see Wertsch, 1978, 1985). By adopting the role of the teacher (or monitor) the children have the chance to monitor, evaluate, and regulate (other children's) cognition themselves. In traditional instructional settings, teachers carry out these tasks. Few difference in results were detected between the *reciprocal* and the *monitor* conditions; both experimental settings produced virtually the same effects, thus supporting the assumption that it is not the adoption of the teacher's role, but the tasks of monitoring, evaluating, and regulating other students' strategy execution that make the Reciprocal Teaching method so effective. Because the method incorporates instruction in both reading strategies and metamemory acquisition procedures and allows for prolonged practice in a social setting with continuous switching of roles and much corrective feedback from experts and peers, Reciprocal Teaching is one of the more complete instructional techniques that, according to Pressley, Snyder and Cariglia-Bull (1987), are likely to prompt durable and general implementation of strategies. Thus, it is a prime example of cognitive apprenticeship (Collins, Brown, & Newman, 1989).

To summarize, it is not only the application of strategies that leads to the acquisition of knowledge about that strategy, but the combination of strategy application with concurrent teaching and learning of metacognitive acquisition procedures (analysis, monitoring, evaluation, and regulation) in an inter-individual way as the precedent of these processes occurring intra-individually that seems to be an efficient way of acquiring metacognitive knowledge and skills. The study has shown

that the Reciprocal Teaching method leads to the acquisition of metacognitive knowledge and skills by having students give each other feedback about the content of the text and the application of the reading strategy, guiding others to correct their answer, or modeling the correct use of the strategy.

4.4. Implications of main findings

The traditional mode of instruction, with teachers presenting and maybe briefly modeling new strategies, followed by children practicing these strategies with only teachers giving feedback – i.e., the model realized in the *student* condition – produced only very few benefits in terms of children's knowledge and ability to apply the strategies. It seems that the children could just as well have spent their time on leisure activities instead of participating in a one-hour strategy training program four times a week over a four-week period. This is especially surprising given that the children in this experimental condition had the most opportunities to apply the reading strategies. In contrast, despite having fewer opportunities to practice, children in the *monitor* and *reciprocal* conditions were expected to monitor, evaluate, and regulate other children's performance, and consequently acquired significantly more knowledge and skills during strategy training. It made no difference whether they were given complete responsibility for their own learning process or not; it was only necessary for children to carry out the content-related monitoring tasks of the "teacher" for them to produce better results than in the traditional mode of instruction (*student* condition).

Therefore, I conclude that it is most important HOW strategies are taught, and not THAT strategies are taught. In order to learn more about strategies and to be able to apply them, concurrent monitoring, evaluation, and regulation of a strategy's application and effectiveness is needed. This is in line with Borkowski et al. (1988), who argued that strategy instruction and practice needs to be accompanied by instruction and application of metamemory acquisition procedures. These metacognitive skills (e.g., planning, monitoring, evaluation, and regulation) are also necessary for the knowledge and skills acquired to be successfully transferred to new settings and learning tasks.

I would like to mention another very important incidental result because of its great relevance to the applicability of the main findings. *All* children were able to learn reading strategies and to implement the Reciprocal Teaching or Monitoring procedures,

adopting the roles of "teacher" or "monitor" – independently of their ability level or performance in the verbal domain. Evidently, every child is able to learn strategies and metamemory acquisition procedures and to profit from this. Of course, although this result holds for the age group of 5th graders who participated in the study, it cannot necessarily be generalized to other age groups.

This study, like other research dealing with Reciprocal Teaching, has found that this cooperative learning method is beneficial for children's learning. In this procedure, children have the opportunity to learn metacognitive skills by interacting with others before their own thoughts and actions become the subject of evaluation and regulation (really "metacognitively"). They not only get feedback on their performance from an "expert" model (the classroom teacher, in this study referred to as the "trainer"), but also from people of about the same level of competence, their peers. They are able to learn from a number of models, which also gives them the opportunity to observe multiple ways of solving a task. All this contributes to their chances of internalizing the new knowledge and skills and transferring these to other domains. Participants have the chance to realize that strategies are cognitive tools that can help to solve tasks if they are applied in a planned, reflective manner and adapted to the learning goals in question.

But does it have to be *Reciprocal Teaching* for reading strategies to be taught effectively? Is it really necessary to have children work exactly as is done in the Reciprocal Teaching method when learning and practicing reading strategies; i.e., in small groups of at least three, optimally four or five, and at most six children, who alternate in adopting the role of teacher? The answer that can be given based on the findings of this study is that training does not necessarily have to be set up in exactly the same way as Reciprocal Teaching. Only some of the components inherent in the Reciprocal Teaching method are necessary for the acquisition of metacognitive knowledge such as strategy knowledge and metamemory acquisition procedures, namely, planning, monitoring, evaluating, and regulating other children's learning processes in a structured and well-defined manner. In Reciprocal Teaching, this is done by having the children themselves acting as "teacher." Yet the very good results obtained by children in another experimental training condition, the *monitor* condition, which were comparable, if not indistinguishable from those produced by the *reciprocal teaching* condition, show that other approaches can be just as successful. Children do

not have to act as teachers themselves, but they do have to carry out some of the teacher's tasks.

To summarize, Reciprocal Teaching seems to support the acquisition of metacognitive knowledge and skills by having students carry out tasks for which teachers are usually responsible: planning, monitoring, evaluating, and regulating the learning processes of other students. By working cooperatively with their peers and an expert, the children can execute metacognitive processes inter-individually. These processes are the basis for metacognitive knowledge and skills to develop, and provide an excellent basis for intra-individual self-regulative thoughts and actions (Vygotsky, see Wertsch, 1978, 1985).

4.5. Limitations and shortcomings of the study

4.5.1. Why did training fail to have an effect on reading comprehension?

The present study attempted to identify the mechanisms causing the large training effects on reading comprehension measures reported in the literature on Reciprocal Teaching (Palincsar & Brown, 1984; Rosenshine & Meister, 1994). Although no hypothesis explicitly addressed this question, I assumed that the training programs implemented would help to improve children's reading comprehension and that this could be tested and established with the comprehension assessments used in this study. Why did the training fail to produce significant improvements on comprehension measures? There are various possible explanations, which will be addressed shortly below.

It has to be stated explicitly that any possible effects of Reciprocal Teaching on reading comprehension are transfer effects. Reciprocal Teaching was not developed to directly improve children's comprehension abilities, but to teach them reading strategies assumed to be beneficial for comprehension abilities and skills. The training procedure implemented in this study focused on practicing the strategies, not on independently reading and comprehending texts and answering comprehension questions. Comprehension assessments were assigned only rarely during training and had no consequences for the children, neither were the children provided with feedback on these assessments. During the brief assessments, the children were not instructed to remember the strategies they had learned. The test-taking conditions were, in these respects, very similar to the conditions implemented by Palincsar and Brown (1984) and

by many of the other researchers whose studies were included in Rosenshine and Meister's (1994) meta-analysis.

The number of training sessions was only moderate, with twelve "real" training sessions. When taking into account the number of sessions the children missed for various reasons (illness or other appointments) as well as the lack of opportunity to experience the effects of applying the strategies in settings other than the training program, it becomes clear that transfer is not likely to occur easily (Borkowski et al., 1988). A kind of transfer that is more likely to occur is that of simpler routines that become partly automatic. Mindful transfer (Salomon & Globerson, 1987) in the sense that the Reciprocal Teaching strategies (*summarizing*, *questioning*, *clarifying*, and *predicting*) are applied deliberately during the comprehension test, on the other hand, would consume a lot, if not almost all, of the children's processing resources (which may lead to even weaker performance than without strategy use, i.e., the transitional stage proposed by Schneider & Büttner, 1995) and take a lot of time. Studies have shown that, even if children already possess the relevant strategy knowledge at the age in question, this does not necessarily mean that they will engage in metacognitively directed information processing (Brown & Smiley, 1978). It is argued that an increase in text comprehension is thus more likely to occur when comprehension assessments make similar demands to the strategies learned during training (e.g., when children are asked to summarize, generate questions, predict, or clarify) than with more distal measures such as those applied in this study (i.e., having to read a longer expository text and answer complex comprehension questions requiring inferences to be drawn).

Other reasons for any transfer that occurred at all being more likely to be automatic than mindful are that the children were not instructed to use the strategies (or reminded of them) at the time of the comprehension assessment and that there was not enough time to implement the strategies (children read the text once or, if they wanted to, twice; it then took them on average 15 to 20 minutes to answer the four comprehension questions). For either automatic or mindful transfer of the strategies to occur, it would have helped for the strategies to have been mindfully and successfully applied in a variety of settings. This was not the case in the present study.

In sum, with a maximum of just twelve training sessions, it is very unlikely that the automatization of time- and energy-consuming reading strategies usually requiring a great deal of practice (Logan, 1988) had already taken place or that these strategies could be transferred from the training context to the test-taking setting, a phenomenon

which is also referred to as production deficit (O'Sullivan & Pressley, 1984; Schneider & Pressley, 1997).

There are more arguments to consider: one is the very young age of the participating children in terms of metacognitive knowledge and skills. According to Piaget's developmental theory, metacognition – where one's own thinking and its possible products become the object of reflection – is bound to the stage of formal operation, which does not occur before the late elementary school years (being first observed around age 10), and continues to develop until and beyond adolescence. It thus seems likely that the 5th graders who participated in the present study possessed only little knowledge about specific reading strategies, knew few strategies, and did not have much general strategy knowledge, experience with strategy application or the resource management necessary for successful application. All of these components necessary for successful transfer were still developing rather than firmly established. The assumption that many components of metacognition develop during the late elementary years is in line with Wellman's theory of mind (1985). Wellman proposed that although, by the end of elementary schooling, children are usually able to differentiate between most person and task variables, can distinguish effort and ability, and have learned about the organizational structure of items, many capacities continue to develop or appear during adolescence or adulthood. These include knowledge about mental states and interaction of memory variables and a great deal of factual knowledge. Another problem with the relatively young age of the children in this study is that their performance on such a difficult test requiring a lot of skills, knowledge, and competence is more dependent on motivation, mood, and physical state than is the case for older children. These factors may have interfered with the relatively small effects of training.

Another issue worth discussing is the setting of the test. This might also explain another interesting finding: the large improvement in comprehension scores for all children from pre- to post-test. The pre-test was administered by an external experimenter, a person unknown to the children. During training, the experimenter (the author herself) was then present at the school every day for a period of at least 4 weeks. By the time of post-test, she not only knew all of the children who participated in the training, but most of their classmates as well. Her status was like that of any other teacher in the school. Additionally, the children were very familiar with the testing procedure and the demands of the test (pre-test effect). It is highly likely that, being

generally more familiar with the demands and format of the test, students read the text differently than at pre-test and/or answered the comprehension questions more freely (Willson & Putnam, 1982). The effects that training may have had on comprehension were, if there were any at all, very small relative to the pre-test effects and may have not been detected for this reason.

Apart from the setting and the method of testing, the reading comprehension assessments used may have been a problem in themselves. As already described, the effect size that other researchers have found for the Reciprocal Teaching method varied dramatically depending on the kind of test used (Rosenshine & Meister, 1994). For standardized tests of comprehension (the Gates MacGinitie Reading Test), significant effects were only reported in two of the nine studies included in Rosenshine and Meister's meta-analysis (1994), with a mean effect size of .32. In contrast, significant effects have almost always been obtained when summarization tests (4 of 5 studies) or experimenter-developed comprehension tests (6 of 7 studies) are used as outcome measures (with mean effect sizes of .85 and 1.00 standard deviations, respectively). In the summarization tests, the independent clauses of 200- to 400-word passages are rated with respect to the importance of the ideas they contain. Experimenter-developed tests use 200- to 800-word passages and usually require students to answer 5 to 10 short-answer questions, half concerning facts and half requiring inferences to be made. However, the questions usually make few demands on children's understanding and ask for single facts or very easy inferences (corresponding to level 1 and 2 answers in the comprehension assessments used in the present study). Additionally, students who participated in some of the training studies covered in this meta-analysis were required to answer the same kind of short comprehension questions practically daily during the training program. Relative to the *control* group, they were thus well acquainted with the post-test assessments. They knew what kind of questions (requiring little or no deeper understanding) were to be expected and could easily pay more attention to textual information of this kind (e.g., how much the polar bear weighs).

The present study, in contrast, used longer expository texts of about 500 words with few, relatively difficult comprehension questions. These questions did not ask for single pieces of information, but required a deeper understanding of the text, the drawing of text-inherent inferences and the combination of textual information with prior knowledge. At pre-test, the students were not familiar with either kind of questions or with the procedure of assessment. The text comprehension assessments

used in this study thus required students to transfer the reading strategies independently, without explicitly being told to do so or to apply any strategies at all from another setting to a test-taking setting involving silently reading a text by oneself and silently answering complex comprehension questions by writing down the answers in a test booklet. During strategy training, strategies were only applied in a context where a short paragraph was read aloud by one child and four or five children orally applied the reading strategies more or less cooperatively. Successfully applying the strategies during the comprehension assessment would have meant applying the reading strategies in one's own mind without overtly discussing an entire text of about one page in length. In training, children were never taught or shown how to apply the reading strategies in a setting other than the training program. The reading comprehension assessments included only very few questions (four at pre- and post-test); thus, the reliability of this measure is restricted. Taking all of these factors into account, it is very unlikely that transfer would occur, a phenomenon which is called production deficit (Renkl, Mandl, & Gruber, 1996).

All of these factors - the unusual setting, the difficult comprehension assessments requiring deep textual understanding, the relative lack of metacognitive knowledge of children of this age, and the large pre-test effects (to name but a few) - impact on students' scoring on the comprehension assessments and may help to explain why training failed to have an effect on reading comprehension. It is argued that, under the present circumstances, it was difficult to find an effect of strategy training on a transfer measure not directly related to the training itself.

4.5.2. What went wrong in the excluded groups?

Three of the experimental groups were excluded from all further analyses, the main reasons being a poor working atmosphere, poor discipline, and inadequate interventions by the trainers. Data provided by the video analyses helped to identify and quantify these reasons. But what caused the problems that eventually forced me to re-sample these groups? Can contributing factors be identified and are there any indications of how such problems might be avoided in the future?

The trainers (all males) of the three groups did not have much experience of teaching. This is very unlikely to be the main cause, however; some of the other trainers had also just started working with children yet did not encounter the same problems. It is also doubtful that the gender of the trainers was of any relevance; in

general, teaching ability is not related to gender. It is more promising to look at the trainers' behavior: two of the trainers tried to establish friendships with the children. Other indicators also make it clear that they did not take the training seriously enough: they were often late, meaning that training sessions could not start in time, and they did not insist that the children use the formal form of address ("Sie") when talking to them, a sign that they did not establish themselves as authorities. All of these circumstances and factors contribute to creating the learning "atmosphere," the environmental conditions for learning. The impression that these factors and the learning atmosphere in general was especially important for the groups in the present study being able to function effectively (see also Slavin, 1985) prompted me to take a closer look at some of the processes occurring during training. Video analyses provided a great deal of information about the working atmosphere and about the children's and trainers' behavior. In view of these findings, all groups in which it could not be ensured that the learning conditions were adequate and that the experimental condition could thus be implemented as intended were excluded from all further analyses in the present study.

In addition to the present study, diploma thesis has been written based on the video analyses of the training sessions (Haase, 2003, supervised by the author). The goal of Haase's thesis (2003) was to examine whether the learning environment had an impact on children's learning and performance during training. The behavior and personality traits of the trainers were of particular interest. In addition to the very detailed video analyses, therefore, all trainers filled out some personality questionnaires. They also were shown excerpts from the videos of their work with the children and asked to rate their own behavior during training and the behavior and learning success of the children. The scales used were mainly parallel to those used for the original video analyses, meaning that the objective judgments of video raters could be compared with the subjective impressions of the trainers themselves. It was shown that children in groups with a positive working atmosphere were more successful in applying the reading strategies during training, independent of their cognitive abilities and reading skills. Working atmosphere, measured during the first part of the strategy training program, explained a significant amount of the variance in children's performance at the end of training, independently from and additionally to their baseline level of performance at the beginning of the program. Trainer behavior proved to influence the working atmosphere in the group: a better working atmosphere and better discipline were observed in groups in which trainers monitored classroom rules effectively and

intervened consistently. These trainers were also the ones who reported preparing for the training program. Trainer proactivity correlated positively with behavior that promoted children's learning, such as establishing and monitoring classroom rules. Trainers who did not tend to resign in case of failure and trainers who tended to strive for perfection were also more effective in monitoring rules. These personality traits also proved to have a direct positive influence on the working atmosphere and discipline. Results indicated that it is necessary to establish a well-defined learning environment and to "set the stage" for working. Groups in which trainers established classroom rules at the beginning of the training program and implemented them consistently were better able to apply the reading strategies at the end of the program. It also had a positive effect when the trainers were the ones determining the course of the early sessions and the children had little to decide. It seems that children were able to learn more when trainers established a well-defined learning environment with clear rules, dominated at the beginning, and taught the children to learn independently and autonomously. The effects of trainer behavior and personality were largely independent of the experimental conditions.

4.5.3. Measures of metacognition

One of the major assumptions of this thesis was that the mechanism which makes the Reciprocal Teaching method so successful is metacognition – monitoring of one's own cognitive state and goal-directed regulation of cognition. Various measures tapping different aspects of metacognition were thus included in this study. Selecting a sensitive measure was not an easy decision (see Cavanaugh & Perlmutter, 1982, for a review of instruments in the field that is still relevant). In line with Borkowski et al.'s model of mature metamemory (1988), I decided to include several knowledge components. Some of them address knowledge that is closely related to the content of the strategy training program (i.e., knowledge about the *summarizing* and *clarifying* strategies that were trained), while others were not directly linked to the training program (*relational strategy knowledge*, *conditional strategy knowledge*, and *planning knowledge*).

However, all of these measures tap children's knowledge; i.e., the declarative component of metacognition. What about procedural metacognition; i.e., the quality of the actual application of strategies? It must be clearly stated that procedural metacognition was not assessed in the present study. It is very difficult to provide

measures of metacognition in action. In some respects, the human mind still resembles Skinner's "black box": only input and output can be directly observed – what happens in between needs to be inferred. There are two main ways to obtain information about the processes that occur in the black box. One is to ask people retrospectively what they were thinking while doing a task (like trying to comprehend a text), and if and how they monitored and/or regulated their cognitive processes and actions. The second approach brings us closer to the metacognitive processes actually occurring during comprehension: by asking people to think aloud during task-solving and analyzing their think-aloud protocols (Veenman & Beishuizen, 2004). Both of these possibilities for finding out more about what people actually think have one major shortcoming, however: people can only report what they consciously think. Yet this only constitutes part of their metacognitive activities, since it is assumed that much of metacognition occurs unconsciously and even automatically (see Brown, 1987; Flavell & Wohlwill,1969). According to Shiffrin and Schneider (1977), controlled strategy use is more likely to be detected than automatic processes. But controlled strategy application has the shortcoming that it consumes quite a lot of the limited processing resources available. Automated processing is not so capacity-demanding and thus is probably more effective – and desirable. These are just a few of the problems with observing "metacognition in action."

In this study, I got closest to children's actual strategy use in recording whether or not they underlined text in the original test booklets when required to summarize a text; i.e., in observing an external strategy that can help to identify the main ideas of a text and reduce the amount of textual information. Another indication of ongoing metacognitive activities may be children's attempts to find a title for their summary. This can be viewed as an indicator of the child trying to figure out a common theme that best characterizes the entire text and can thus serve as the title. Finding out more about the ongoing metacognitive activities that were affected by systematic strategy intervention would help a lot, not only to develop an understanding of the complicated processes of developing expertise in learning (mature metamemory and effective strategy use, Weinert, 1984, 1996), but potentially also to provide useful information on how to help students acquire metacognitive knowledge and develop effective routines to support their cognitive processing.

4.5.4. Sample size and power

There are a number of factors that need to be considered when designing a study, among them sample size, statistical power, and realizability. The larger the number of participants, the smaller the standard deviation and confidence intervals, and the greater the probability of finding statistically significant effects. By raising the number of participants or reducing the standard deviation, even small differences can prove to be "significant." These differences may not be of practical relevance, however. An important measure that provides information about the practical relevance of empirical (significant) results is the effect size (d). It is computed by weighing the difference of means between the experimental group and the control group by the standard deviation of the control group; thus, the mean difference is standardized and can be compared across experiments. Power analyses help to in determine the minimum sample size.

In the planning phase of this study, a power analysis (Erdfelder, Faul, & Buchner, 1996) was used to determine the sample size. The starting point for these considerations were the effect sizes usually obtained in experimental studies on Reciprocal Teaching: effect sizes of between .32 and 1.00 standard deviations, depending on the measure (mean effect size of .88; see Rosenshine & Meister, 1994). A significance level of alpha being .05 was adopted. A moderate significant effect with an effect size of .60 would be detected with a power of .70 and a sample size of 50 participants; it would take a minimum of 100 participants to obtain the same effect with a power of .90. After careful consideration, it was decided to draw a sample of 50 to 60 participants.

With the realized sample of 169 children, 78 of whom served as *control* children and 57 of whom participated in the various experimental training groups (17 *reciprocal* condition; 20 *monitor* condition; 18 *student* condition), effects of moderate to high size could be proved to be statistically significant. More specifically, effect sizes between .64 and 1.46 were observed for knowledge about the *summarizing* and *clarifying* reading strategies. The smallest, but significant effect (d = .34) was observed for the number of trained children relative to *control* children who produced a title for their summary. For some interesting results that just failed to reach significance in the predicted direction, small effect sizes were observed (greater ratio of main ideas to details in the summaries of trained than relative to *control* children, with p = .066 and d = .33; trained children acquiring more *relational strategy knowledge* relative to

control children from pre- to post-test, with p = .16 and d = .28). Evidently, the number of participants was not sufficient to detect differences with the adopted specifications for statistical tests in these cases.

4.6. Personal review of the effects of Reciprocal Teaching

In addition to the many results that have been reported throughout this thesis, much more was learned and observed during the study that is difficult to express quantitatively in terms of numbers or frequencies. At least, no attempt has been made to capture these aspects thus far. In my opinion, however, these aspects are especially characteristic of the Reciprocal Teaching method, and provide useful and necessary information for teachers and researchers intending to implement the method.

First of all, Reciprocal Teaching appeared to be fun for both the students and trainers. After only a short introduction and maybe a demonstration, the children understood the roles allocated and accepted the responsibilities. Of course, sometimes they still sought the feedback of the trainer by looking at him or her, but after a few sessions they acted as if the trainer was not present at all. This impression was confirmed by a few situations where the trainer had to leave the room for a short period (with the video camera still running), and the group kept on working as if nothing had happened. This, of course, was only observed in groups with a good working atmosphere and good discipline. It was especially interesting to observe that students who did not usually do well in school, and even students with learning disabilities, did a great job in adopting the role of the "teacher" and became engaged in the group and the learning process. For them, it was a rare opportunity in school to be really involved in the group in a completely different position than usual, feeling competent and even able to help other students when they forgot to do one of their tasks by adopting the "teacher" role.

Engagement and enjoyment played a major role in keeping the groups "on track," which was not easy because training took place in the children's spare time almost daily for a period of four weeks. This impression was verified by children arriving on time, helping to get started (arrange the desks and chairs, carry the material), and not complaining. Often they became so involved in the work that they forgot the time (i.e., experienced flow); trainers were not often asked "How long until we can go home?" All of this resulted in an excellent proportion of time on task – on average,

training sessions lasted about 55 minutes; with 60 minutes initially planned (i.e., longer than regular school lessons of just 45 minutes).

The unusual setting and mode of interaction common to all experimental conditions (a feature shared by many cooperative learning methods, see Slavin, 1991, 1996) – involving all participants including the teacher working around a small table – allowed a very special kind of cooperation to take place. The students were instructed not to raise their hands, told to speak freely, and had the chance to experience something new: acceptance of multiple ways of problem solving. When correcting, the trainer did not simply point out the student's mistakes and then demonstrate "how it should be done correctly," but worked with the student's answer and tried to show multiple ways of improving it. It was stated very clearly that there is no "one best way," but that there are many approaches and that each of them is correct. The trainers were also instructed to show the students that they also make mistakes and need to correct them. It was very important that the children lost their fear of saying something wrong and being laughed at. In the beginning, they often hesitated to answer for precisely this reason. However, it was soon made clear that it is not only acceptable to make mistakes, but that we need to make mistakes in order to know what should be improved. In many of the groups, one of the class rules set up was not to laugh at other students. This made it possible to establish a safe atmosphere where everyone could try to answer, simply doing his or her best: an accepting learning climate. One of the "outcomes" of these special working conditions was that the children learned to like each other. The groups consisted of children whose schedules (for sports and other afternoon activities) matched those of the rest of the group, and thus comprised children from different classes, some of whom were already friends, some of whom were not. Through changing the seating arrangements every day, all students had the chance to get to know one another. All of this contributed, in my opinion, to creating a good working climate and establishing new personal relationships. Some children mentioned this explicitly during our last training session, when we took time to talk about the program. It was a very special experience for both the children and the trainers.

Finally, I would like to give some more technical "advice." One of the recommendations I would like to make concerns the size and composition of the groups. The groups should consist of at least four and not more than five students. When one child was missing from a group of four, training proceeded faster, but there was less input into the interaction. When two children were missing, the interaction became

really boring, with one child acting as "teacher" and the other as his/her "student," no one interrupting or giving other ideas. With more than five children, we had the impression that the groups were too large to have everyone participate fully and be able to state their own ideas. The ability level of the group members is also worthy of consideration. Some research findings suggest that Reciprocal Teaching is a method suitable for students of almost all ability levels and that it appears to improve students' knowledge and skills no matter how the groups are set up. My experience, and that of some of the other trainers in this study, was that when a group consisted of mainly low-achieving students or students with learning disabilities (especially in reading) and only one high- or moderately achieving student, the quality and also quantity of the topics and discussions was greatly reduced and the training method alone was not sufficient to ensure successful management of the strategies and understanding of the text. This impression is consistent with other research findings (Bennett & Cass, 1988; Webb, 1985). I would argue strongly in favor of mixed-ability groups with enough high-achieving students who are able to carry the discourse, keep it up, introduce new ideas, and discover mistakes made by others. The last point in this section concerns the relationship between the trainer and students. Although a warm and friendly relationship is intended, with the trainer being a learning partner rather than an instructor, the relationship still needs to remain that of a teacher and his/her students. This means using the formal form of address ("Sie") for the trainer and adhering to class rules and the trainer's instructions. Empirical findings supportive of this notion have been provided by a detailed analysis of video data from the study (diploma thesis by Haase, 2003).

In sum, Reciprocal Teaching is a very special teaching method and cooperative learning approach because the interaction is manifold, fluent, and interesting. More rich and elaborated ideas were produced when children worked in groups of four or five. With fewer children, the procedure became more of a routine and some of the children seemed to get bored with the ongoing activities. When there are more competent children in the group who can serve as models, moreover, the chances are greater that children will get to know more than one "correct" approach by observing different students. There are many advantages to this method, but it takes competent, engaged teachers to implement it well and be responsive to students' needs and abilities.

4.7. Outlook

The major finding of this study is that HOW learning strategies are taught is the most important factor in producing benefits in metacognitive knowledge and allowing for transfer; for instance, to reading comprehension. The Reciprocal Teaching method seems to be a very effective way of promoting the acquisition of metacognitive knowledge and skills. It was also shown, however, that strategy training does not necessarily have to include the precise kind of interaction that characterizes this instructional approach. Instead, the tasks of monitoring, evaluating, and regulating other children's learning processes – i.e., tasks associated with the "teacher role" – are the ones that promote the acquisition of metacognitive knowledge and skills. Generally, any strategy training program that not only provides children with plentiful opportunities for practice, but also prompts them to engage in these kinds of metacognitive processes may help children to acquire metacognitive knowledge and skills. In an instructional setting such as peer tutoring, the children have plenty of learning opportunities and the chance to really practice their new skills and develop routines. Training programs other than Reciprocal Teaching have also been shown to produce good results in strategy application and to allow for transfer to reading comprehension (see Streblow, 2004 for an overview of training programs designed to improve reading comprehension and/or motivation). Two training approaches in which students are taught nearly the same strategies as Reciprocal Teaching and that incorporate peer tutoring activities are Peer-Assisted Learning Strategies (Fuchs, Fuchs, Mathes, & Simmons, 1997) and Collaborative Strategic Reading (Vaughn, Klingner, & Bryant, 2001). In PALS, children are taught reading strategies and then practice these in pairs of two; tutorial activities are modeling, guiding, correcting, and motivating. The program allocates rewards for cooperative behavior, but also includes competitive structures; it was originally developed for children in grades 2 to 6 and was extended downward to first-graders and kindergarten children, and upward to high school students. Collaborative Strategic Reading was developed for students in grades 4 through 8, and introduces interdependencies among group members and responsibilities by assigning students to one of three different roles ("leader," "clunk expert," "reporter").

There are more prerequisites for strategies to be internalized, generalized, and applied independently in transfer settings. It also seems important to provide children with a great deal of practice using strategies – a lot more than could be done in the

present study – for automatization to develop. The strategies should be applied in a variety of contexts to facilitate transfer. In addition, in order to be motivated to apply these powerful cognitive tools, children need to realize that the application of often these time- and resource-intensive strategies supports their learning processes and improves their learning outcomes . Traditional ways of providing children with feedback about their performance, namely grades, are not recommended here. Instead, teachers need to give students informative feedback about their learning gains (see Hattie, Biggs, & Purdie, 1996 for a meta-analysis). Receiving and giving corrective feedback is one of the central features of the Reciprocal Teaching method, and is necessary for the children to acquire metacognitive knowledge and skills. It places high demands on teachers' time and abilities, however. Time is one of the major reasons why these (cooperative) learning methods are not very often implemented in the curriculum. When considering the potential that mindful strategic learning has, however, it is clear that it is worth investing this extra time.

Other, non-technical factors also contribute to the success of cooperative learning methods like RT: the teacher's personality, knowledge, and skills; students' abilities and behavior; and the interaction of the two, as was also addressed with the data provided by this study. More specifically, video analyses provided a wealth of information about the working atmosphere and children's and trainers' behavior, resulting in a diploma thesis being written in addition to the present study (Haase, 2003, supervised by the author).

There is still a great deal to learn about metacognition, learning processes, and possible settings for learning. Cooperative learning settings such as Reciprocal Teaching are a good way of providing children with valuable learning experience and, if structured in ways that allow children to assume more responsibility for their learning processes, a very useful method which should be integrated in the curriculum. In the present study, it was shown that strategy instruction is most effective if instruction in and practice of the strategies is combined with instruction in and practice of monitoring and evaluative and regulative processes. The learning outcomes of strategy instruction are greatly enhanced when children practice these meta-strategies. These seem to be the most effective features of the Reciprocal Teaching method. The goal of full internalization, automatization, and successful transfer of reading strategies to other contexts was not achieved in the present study. Nevertheless, the findings of this study

contribute to a deeper understanding of the processes and components that make strategy learning possible and effective.

5. REFERENCES

Artelt, C. (2004). Zur Bedeutung von Lernstrategien beim Textverstehen. In J. Köster & W. Lütgert (Eds.), *Aufgabenkultur im Deutschunterricht - Ihre Bedeutung für die Entwicklung der Lesekompetenz*. Frankfurt: Peter Lang.

Artelt, C., Stanat, P., Schneider, W., & Schiefele, U. (2001). Lesekompetenz: Testkonzeption und Ergebnisse. In J. Baumert & E. Klieme & M. Neubrand & M. Prenzel & U. Schiefele & W. Schneider & P. Stanat & K.-J. Tillmann & M. Weiß (Eds.), *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (pp. 69-137). Opladen: Leske & Budrich.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 2, pp. 90-197). New York: Academic Press.

Baddeley, A. (1986). *Working memory*. Oxford: Oxford University Press.

Baddeley, A. (1992). Working memory. *Science, 255*, 556-559.

Bargh, J. A., & Schul, Y. (1980). On the cognitive benefits of teaching. *Journal of Educational Psychology, 72*(5), 593-604.

Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Stanat, P., Tillmann, K.-J., & Weiß, M. (2001). *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske & Budrich.

Bennett, N., & Cass, A. (1988). The effects of group composition on group interactive processes and pupil understanding. *British Educational Research Journal, 15*(1), 19-32.

Berliner, D. C. (1996). In pursuit of the expert pedagogue. *Educational Researcher, 15*(7), 5-13.

Bjorklund, D. F., & Harnishfeger, K. K. (1990). The resources construct in cognitive development: Diverse sources of evidence and a theory of inefficient inhibition. *Developmental Review, 10*, 48-71.

Bloom, B. S., Englehart, M. B., Furst, E. J., Hill, W. H., & Krathwohi, O. R. (1956). Taxonomy of educational objectives: The classification of educational goals., *Handbook 1: The cognitive domain*. New York: Longman.

Borkowski, J. G., Milstead, M., & Hale, C. (1988). Components of children's metamemory: implications for strategy generalization. In F. E. Weinert & M. Perlmutter (Eds.), *Memory development: Universal changes and individual differences* (pp. 73-100). Hillsdale, NJ: Lawrence Erlbaum Associates.

Borkowski, J. G., & Thorpe, P. K. (1994). Self-regulation and motivation: A life-span perspective on underachievement. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-regulation of learning and performance* (pp. 45-73). Hillsdale, NJ: Erlbaum.

Bradley, L., & Bryant, P. (1985). *Rhyme and reason in reading and spelling*. Ann Arbor: The University of Michigan Press.

Brady, P. L. (1990). Improving the reading comprehension of middle school students through reciprocal teaching and semantic mapping strategies. *Dissertation Abstracts International*.

Brown, A. (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In F. E. Weinert & R. H. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 65-116). Hillsdale, N.J.: Erlbaum.

Brown, A., & Palincsar, A. S. (1985). *Reciprocal teaching of comprehension strategies: A natural history of one programme for enhancing learning* (Technical 334). Urbana: University of Illinois: Center for the Study of Reading.

Brown, A. L., & DeLoache, J. S. (1978). Skills, plans, and self-regulation. In R. S. Siegler (Ed.), *Children's thinking: What develops?* (pp. 3-35). Hillsdale, NJ: Lawrence Erlbaum Associates.

Brown, A. L., & Smiley, S. S. (1978). The development of strategies for studying texts. *Child Development, 49*, 1076-1088.

Cariglia-Bull, T., & Pressley, M. (1990). Short-term memory differences between children predict imagery effects when sentences are read. *Journal of Experimental Child Psychology, 49*(3) Jun 1990), United Kingdom, www.

Case, R. (1972). Validation of a neo-Piagetian mental capacity construct. *Journal of Experimental Child Psychology, 14*, 287-302.

Case, R. (1978). Intellectual development from birth to adulthood: A neo-Piagetian interpretation. In R. S. Siegler (Ed.), *Children's thinking: What develops?* (pp. 37-72). Hillsdale, NJ: Lawrence Erlbaum Associates.

Case, R. (1985). *Intellectual development: Birth to adulthood*. Orlando, FL: Academic Press.

Cavanaugh, J. C., & Perlmutter, M. (1982). Metamemory: A critical examination. *Child Development, 53*, 11-28.

Chi, M. T. H. (1978). Knowledge structures and memory development. In R. S. Siegler (Ed.), *Children's thinking: What develops?* (pp. 73-96). Hillsdale, NJ: Lawrence Erlbaum Associates.

Chi, M. T. H. (1987). Representing knowledge and metaknowledge: Implications for interpreting metamemory research. In F. E. Weinert & R. H. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 239-266). Hillsdale, N.J.: Erlbaum.

Clark, R. E. (1990). When teaching kills learning: Studies of mathematantics. In H. Mandl & E. DeCorte & N. S. Bennett & H. F. Friedrich (Eds.), *Learning and instruction: European research in an international context* (pp. 1-22). Oxford, UK: Pergamon Press.

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the craft of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser*. Hillsdale, NJ: Erlbaum.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671-684.

Csikszentmihalyi, M. (1985). *Das Flow-Erlebnis*. Stuttgart: Klett-Cotta.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behavior, 19*(4), 450-466.

Dann, H. D., Diegritz, T., & Tosenbusch, H. S. (1999). *Gruppenunterricht im Schulalltag: Realität und Chancen*. Erlangen: Universitätsverbund.

Das mußt Du wissen. (1990). Weinheim: Zweiburgen Verlag.

Deci, E., & Ryan, R. M. (1985). Intrinsic motivation and self-determination in human behavior. New York: Plenum Press.

Deci, E. L., Vallerand, R. J., Pelletier, L. G., & Ryan, R. M. (1991). Motivation and education: The self-determination perspective. *Educational Psychologist, 26*, 325-346.

Dempster, F. N. (1985). Short-term memory development in childhood and adolescence. In C. J. Brainerd & M. Pressley (Eds.), *Basic processes in memory development : progress in cognitive development research* (pp. 209-248). New York: Springer.

Dossenbach, H. D. (1992). *Das Lexikon der Tiere*: Kinderbuchverlag Luzern.

Elley, W. B. (1994). The IEA study of reading literacy : achievement and instruction in 32 school systems. Oxford: Pergamon Pr.

Engle, R. W. (1996). Working memory and retrieval: an inhibition-resource approach. In I. T. Richardson (Ed.), *Working memory* (pp. 89-119). Oxford: Oxford University Press.

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, & Computers, 28*, 1-11.

Eroberer und Entdecker. (1990). Weinheim: Zweiburgen Verlag.

Eysenck, M. W., & Keane, M. (2000). *Cognitive psychology : a student's handbook* (4. ed.). Hove: Psychology Press.

Flavell, J. H. (1971). First discussant's comments: What is memory development the development of? *Human Development, 14*(4), 272-278.

Flavell, J. H. (1987). Speculations about the nature and development of metacognition. In F. E. Weinert & R. H. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 21-29). Hillsdale, N.J.: Erlbaum.

Flavell, J. H., & Wellman, H. M. (1977). Metamemory. In R. V. Kail & J. W. Hagen (Eds.), *Perspectives on the development of memory and cognition* (pp. 3-33). Hillsdale, NJ: Erlbaum.

Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.

Friedrich, H. F. (1995). Analyse und Förderung kognitiver Lernstrategien. *Empirische Pädagogik, 9*(2), 115-153.

Friedrich, H. F., & Mandl, H. (1992). Lern- und Denkstrategien - ein Problemaufriß. In H. Mandl & H. F. Friedrich (Eds.), *Lern- und Denkstrategien. Analyse und Intervention* (pp. 3-54). Göttingen: Hogrefe.

Fuchs, D., Fuchs, L. S., Mathes, P. G., & Simmons, D. C. (1997). Peer-assisted learning strategies: Making classrooms more responsive to diversity. *American Educational Research Journal, 34*(1), 174-206.

Garrod, S. C., & Sanford, J. A. (1994). How discourse representation affects language processing. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 675-698). San Diego, CA: Academic Press.

Gick, M. L., & Holyoak, K. J. (1980). Analytical problem solving. *Cognitive Psychology, 12*(306-355).

Gruehn, S. (2000). Unterricht und schulisches Lernen: Schüler als Quellen der Unterrichtsbeschreibung. Münster, New York, München, Berlin: Waxmann.

Haase, W. (2003). Training von Lesestrategien in Kleingruppen: Der Einfluß von Faktoren der Lernumwelt und von Persönlichkeitseigenschaften der Lehrkräfte. Unpublished Diplomarbeit, University of Potsdam, Potsdam, Germany.

Hasselhorn, M. (2001). Metakognition. In D. H. Rost (Ed.), *Handwörterbuch Pädagogische Psychologie* (2 ed., pp. 466-471). Weinheim: Beltz.

Hasselhorn, M., & Koerkel, J. (1986). Metacognitive versus traditional reading instructions: The mediating role of domain-specific knowledge on children's text-processing. *Human Learning: Journal of Practical Research & Applications, 5*(2), 75-90.

Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research, 66*(2), 99-136.

Heller, K. A., & Perleth, C. (2000). *Kognitiver Fähigkeits-Test (Rev.) für 5.-12. Klassen (KFT 5-12+ R)* (3. Aufl. ed.). Göttingen: Beltz-Testgesellschaft.

Jacobs, J. E., & Paris, S. G. (1987). Children's metacognition about reading: Issues in definition, measurement, and instruction. *Educational Psychologist, 22*(3-4), 255-278.

Jenkins, J. J. (1979). Four points to remember: A tetrahedral model and memory experiments. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 429-446). Hillsdale, NJ: Erlbaum.

Johnson-Laird, P. N. (1983). *Mental models*. Cambridge, MA: Harvard University Press.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*, 122-149.

Kästner, E. (1970). (Vol. 172): Ravensburger Taschenbuch Band.

Kintsch, W. (1982). Aspects of text comprehension. *Language and comprehension*, 301-312.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York, NY, US: Cambridge University Press.

Kintsch, W., & Kozminsky, E. (1977). Summarizing stories after reading and listening. *Journal of Educational Psychology, 74*, 828-834.

Kirby, J. R. (1988). Style, strategy, and skill in reading. In R. R. Schmeck (Ed.), *Learning strategies and learning styles* (pp. 229-274). New York: Plenum Press.

Klauer, K. J. (1996). Denktraining oder Lesetraining? Über die Auswirkungen eines Trainings zum induktiven Denken sowie eines Lesetrainings auf Leseverständnis und induktives Denken. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 28*(1), 67-89.

Klieme, E., Neubrand, M., & Lüdtke, O. (2001). Mathematische Grundbildung: Testkonzeption und Ergebnisse. In J. Baumert & E. Klieme & M. Neubrand & M. Prenzel & U. Schiefele & W. Schneider & P. Stanat & K.-J. Tillmann & M. Weiß

(Eds.), *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (pp. 139-190). Opladen: Leske & Budrich.

Körkel, J. (1987). Die Entwicklung von Gedächtnis- und Metagedächtnisleistungen in Abhängigkeit von bereichsspezifischen Vorkenntnissen. Frankfurt/Main: Lang.

Krapp, A. (1992). Konzepte und Forschungsansätze zur Analyse des Zusammenhangs von Interesse, Lernen und Leistung. In A. Krapp & M. Prenzel (Eds.), *Interesse, Lernen, Leistung. Neuere Ansätze einer pädagogisch-psychologischen Interessenforschung* (pp. 9-52). Münster: Aschendorff.

Krapp, A. (1998). Entwicklung und Förderung von Interessen im Unterricht. *Psychologie und Erziehung im Unterricht, 44*, 185-201.

Krapp, A. (1999). Intrinsische Lernmotivation und Interesse. Forschungsansätze und konzeptuelle Überlegungen. *Zeitschrift für Pädagogik, 45*(3), 387-406.

Kreutzer, M. A., Leonard, S. C., & Flavell, J. H. (1975). An interview study of children's knowledge about memory. *Monographs of the Society for Research in Child Development, 40*(1), 1-57.

Kurtz, B., & Borkowski, J. G. (1984). Children's metacognition: Exploring relations among knowledge, process, and motivational variables. *Journal of Experimental Child Psychology, 37*, 335-354.

Kurtz, B., Reid, M. K., Borkowski, J. G., & Cavanaugh, J. C. (1982). On the reliability and validity of children's metamemory. *Bulletin of the Psychonomic Society, 19*, 137-140.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95*(4), 492-527.

Lompscher, J. (2001). Lehrstrategien. In D. H. Rost (Ed.), *Handwörterbuch Pädagogische Psychologie* (2 ed., pp. 394-401). Weinheim: Beltz.

Lonberger, R. B. (1989). The effects of training in self-generated learning strategy on the prose processing abilities of fourth and sixth graders. *Dissertation Abstracts International, 49*(12-A, Pt 1), 3674.

Magazin, Nr. 51. (1999, 23.12.1999). *Süddeutsche Zeitung*, pp. 16-19.

Mähler, C., & Hasselhorn, M. (2000). Lern- und Gedächtnistraining bei Kindern. In K. J. Klauer (Ed.), *Handbuch Kognitives Training* (pp. 407-429). Göttingen: Hogrefe.

Mandler, J. M. (1978). A code in the node: The use of a story scheme in retrieval. *Discourse Processes, 1*, 14-35.

Mandler, J. M. (1984). *Stories, scripts and scenes*. Hillsdale, NJ: Erlbaum.

Marton, F., & Saljö, R. (1984). Approaches to learning. In F. Marton & D. J. Hounsell & N. J. Entwistle (Eds.), *The experience of learning* (pp. 36-55). Edinburgh: Scottish Academic Press.

Marx, H. (1998). *Knuspels Leseaufgaben*. Göttingen: Hogrefe.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88*, 375-407.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review, 63*, 81-97.

Moore, P. J. (1988). Reciprocal teaching and reading comprehension: A review. *Journal of Research in Reading, 11*(1), 3-14.

Nicholls, J. G. (1978). The development of the concepts of effort and ability, perception of academic attainment, and the understanding that difficult tasks require more ability. *Child Development, 49*, 800-814.

OECD. (2000). Literacy in the information age : final report of the international adult literacy survey. Paris: Oecd.

O'Sullivan, J. T., & Pressley, M. (1984). Completeness of instruction and strategy transfer. *Journal of Experimental Child Psychology, 38*(2), 275-288.

Palincsar, A., David, Y., & Brown, A. (1989). Using Reciprocal Teaching in the Classroom: A Guide for Teachers.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition & Instruction, 1*(2), 117-175.

Palincsar, A. S., Brown, A. L., & Martin, S. M. (1987). Peer interaction in reading comprehension instruction. *Educational Psychologist, 22*(3-4), 231-253.

Paris, S. G., Cross, D. R., & Lipson, M. Y. (1984). Informed Strategies for Learning: A program to improve children's reading awareness and comprehension. *Journal of Educational Psychology, 76*(6), 1239-1252.

Paris, S. G., & Jacobs, J. E. (1984). The benefits of informed instruction for children's reading awareness and comprehension skills. *Child Development, 55*(6), 2083-2093.

Pearson, P. D., Hansen, J., & Gordon, C. (1979). The effect of background knowledge on young children's comprehension of explicit and implicit information. *Journal of Reading Behavior, 11*, 201-209.

Pressley, M., Borkowski, J. G., & O'Sullivan, J. T. (1984). Memory strategy instruction is made of this: Metamemory and durable strategy use. *Educational Psychologist, 19*(2), 94-107.

Pressley, M., Borkowski, J. G., & Schneider, W. (1989). Good information processing: What it is and how education can promote it. *International Journal of Educational Research, 13*, 857-867.

Pressley, M., Levin, J. R., & Ghatala, E. S. (1988). Strategy-comparison opportunities promote long-term strategy use. *Contemporary Educational Psychology, 13*, 157-168.

Pressley, M., Ross, K. A., Levin, J. R., & Ghatala, E. S. (1984). The role of strategy utility knowledge in children's strategy decision making. *Journal of Experimental Child Psychology, 38*, 491-504.

Pressley, M., Snyder, B. L., & Cariglia-Bull, T. (1987). How can good strategy use be taught to children? Evaluation of six alternative approaches. In S. M. Cormier & J. D. Hagman (Eds.), *Transfer of learning. Contemporary research and applications* (pp. 81-120). San Diego: Academic Press.

Renkl, A., Mandl, H., & Gruber, H. (1996). Inert knowledge: Analyses and remedies. *Educational Psychologist, 31*(2), 115-121.

Rheinberg, F. (1986). Lernmotivation. In W. Sarges & R. Fricke (Eds.), *Psychologie für die Erwachsenenbildung* (pp. 360-365). Göttingen: Hogrefe.

Rich Zelmanowicz, R. (1989). The effects of training adult poor readers tu use text comprehension strategies. *Dissertation Abstracts International*.

Richter, T., & Christmann, U. (2002). Lesekompetenz: Prozessebenen und interindividuelle Unterschiede. In N. Groeben & B. Hurrelmann (Eds.), *Lesekompetenz: Bedingungen, Dimensionen, Funktionen* (pp. 25-58). Weinheim: Inventa Verlag.

Rogoff, B., & Gardner, W. (1984). Adult guidance of cognitive development. In B. Rogoff & J. Lave (Eds.), *Everyday cognition*. Cambridge, MA: Harvard University Press.

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

Salomon, G., & Globerson, T. (1987). *Skill is not enough: The role of mindfulness in learning and transfer* (Technical Report 11). Tel Aviv, Israel: Tel Aviv University, School of Education.

Scardamalia, M., & Bereiter, C. (1986). Research on written composition. In Wittrock (Ed.), *Handbook of research on teaching* (3 rd ed., pp. 778-803). New York: Macmillan.

Schiefele, H. (1986). Interesse - Neue Antworten auf ein altes Problem. *Zeitschrift für Pädagogik, 32*, 153-162.

Schiefele, U. (1996). *Motivation und Lernen mit Texten*. Göttingen: Hogrefe.

Schlagmüller, M., Visé, M., & Schneider, W. (2001). Zur Erfassung des Gedächtniswissens bei Grundschulkindern: Konstruktionsprinzipien und empirische Bewährung der Würzburger Testbatterie zum deklarativen Metagedächtnis. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 33*(2), 91-102.

Schneider, W. (2004). Frühe Entwicklung von Lesekompetenz: Zur Relevanz vorschulischer Sprachkompetenzen. In C. Artelt & U. Schiefele & W. Schneider & P. Stanat (Eds.), *Entwicklung, Bedingungen und Förderung der Lesekompetenz: Vertiefende Analysen der PISA-2000-Daten*. Opladen: Leske & Budrich.

Schneider, W., & Pressley, M. (1997). *Memory development between two and twenty* (2nd ed.). Mahwah, NJ, US: Lawrence Erlbaum Associates, Inc., Publishers.

Schneider, W., & Schlagmüller, M. (2002). *Metakognitives Wissen über Textverarbeitung*: Universität Würzburg: Lehrstuhl für Psychologie IV.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84*(1), 1-66.

Schneider, W., & Weinert, F. E. (1995). Memory development during early and middle childhood: Findings from the Munich Longitudinal Study (LOGIC). In F. E. Weinert & W. Schneider (Eds.), *Memory performance and competencies: Issues in growth and development* (pp. 263-279). Mahwah, NJ: Lawrence Erlbaum Associates.

Schnotz, W. (1994). *Aufbau von Wissensstrukturen*. Weinheim: Psychologie Verlags Union.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review, 84*(2), 127-190.

Slavin, R. E. (1985). An introduction to cooperative learning research. In R. E. Slavin & S. Shaven & S. Kagan & R. Hertz-Lazarowitz & C. Webb & R. Schmuck (Eds.), *Learning to cooperate, cooperating to learn* (pp. 5-15). New York: Plenum Press.

Slavin, R. E. (1991). Synthesis of research on cooperative learning. *Educational Leadership, 48*, 71-82.

Slavin, R. E. (1994). *Educational Psychology: Theory and Practice*. Boston: Allyn and Bacon.

Slavin, R. E. (1996). Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology, 21*, 43-69.

Stiftung Lesen. (2001). Leseverhalten in Deutschland im neuen Jahrtausend : eine Studie der Stiftung Lesen. Hamburg: Spiegel-Verlag Augstein.

Streblow, L. (2004). Förderung von Lesekompetenz. In U. Schiefele & C. Artelt & W. Schneider & P. Stanat (Eds.), *Entwicklung, Bedingungen und Förderung der Lesekompetenz: Vertiefende Analysen der PISA-2000-Daten*. Opladen: Leske & Budrich.

Thomas, J. W., Bol, L., Warkentin, R. W., Wilson, M., Strage, A., & Rohwer, W. D., Jr. (1993). Interrelationships among student's study activities, self-concept of academic ability, and achievement as a function of characteristics of high-school biology courses. *Applied Cognitive Psychology, 7*, 499-532.

Todt, E. (1978). *Das Interesse*. Bern: Huber.

Treff-Schülerbuch 1993. (1992). Seelze: Velber Verlag.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York, NY: Academic Press.

Vanderbilt, C. a. T. G. a. (1990). Anchored instruction and its relationship to situated cognition. *Educational Researcher, 19*, 2-10.

Vanderbilt, C. a. T. G. a. (1992). The Jasper Series as an example of anchored instruction: Theory, program description, and assessment data. *Educational Psychologist, 27*, 291-315.

Vaughn, S., Klingner, J. K., & Bryant, D. P. (2001). Collaborative strategic reading as a means to enhance peer-mediated instruction for reading comprehension and content-area learning. *Rase: Remedial & Special Education, 22*(2).

Veenman, M. V. J., & Beishuizen, J. J. (2004). Intellectual and metacognitive skills of novices while studying texts under conditions of text difficulty and time constraint. *Learning and Instruction, 14*, 621-640.

Webb, C. (1985). Cooperative learning in mathematics and science. In R. E. Slavin & S. Shaven & S. Kagan & R. Hertz-Lazarowitz & C. Webb & R. Schmuck (Eds.), *Learning to cooperate, cooperating to learn* (pp. 173-176). New York: Plenum Press.

Webb, N. M. (1989). Peer interaction and learning in small groups. *International Journal of Educational Research, 13*, 21-39.

Wechsler, D. (1983). HAWIK-R : Hamburg-Wechsler-Intelligenztest für Kinder; Revision 1983 ; Handbuch und Testanweisung. Bern ; Stuttgart ; Wien: Huber.

Weinert, F. E. (1984). Metakognition und Motivation als Determinanten der Lerneffektivität: Einführung und Überblick. In F. E. Weinert & R. H. Kluwe (Eds.), *Metakognition, Motivation und Lernen* (pp. 9-21). Stuttgart: Kohlhammer.

Weinert, F. E. (1996). Für und Wider die "neuen Lerntheorien" als Grundlagen pädagogisch-psychologischer Forschung. *Zeitschrift für Pädagogische Psychologie, 10*, 1-12.

Weinert, F. E., & Schneider, W. (1987). *The Munich Longitudinal Study on the Genesis of Individual Competencies (LOGIC)* (Technical Report 2. Documentation of assessment procedures used in waves one to three). Munich: Max Planck Institute for Psychological Research.

Weinert, F. E., & Schneider, W. (1999). *Individual development from 3 to 12: Findings from the Munich Longitudinal Study*. Cambridge: Cambridge University Press.

Weinstein, C. E., & Mayer, R. E. (1986). The teaching of learning strategies. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 815-327). New York: Macmillan.

Wellman, H. M. (1983). Metamemory revisited. In M. T. H. Chi (Ed.), *Trends in memory development research*. Basel, Switzerland: Karger.

Wellman, H. M. (1985). A child's theory of mind: The development of conceptions of cognition. In S. R. Yussen (Ed.), *The growth of reflection in children* (pp. 169-206). New York: Academic Press.

Wellman, H. M., Collins, J., & Glieberman, J. (1981). Understanding the combination of memory variables: Developing conceptions of memory limitations. *Child Development, 52*, 1313-1317.

Wertsch, J. V. (1978). Adult-child interaction and the roots of metacognition. *Quarterly Newsletter of the Institute for Comparative Human Development, 2*(1), 15-18.

Wertsch, J. V. (1985). *Vygotsky and the social formation of mind*. Cambridge, Mass. [u.a.]: Harvard University Press.

Wertsch, J. V., & Hickmann, M. (1987). Problem solving in interaction: Amicrogenetic analysis. In M. Hickmann (Ed.), *Social and functional approaches to language and thought* (pp. 251-266).

Widmann, G. (1994). *Diktate 5/6*. München: Adolf Hauschka Verlag.

Wild, K.-P. (2000). Die Bedeutung betrieblicher Lernumgebungen für die langfristige Entwicklung intrinsischer und extrinsischer motivationaler Lernorientierungen. In U. Schiefele & K. P. Wild (Eds.), *Interesse und Lernmotivation: Untersuchungen zu Entwicklung, Förderung und Wirkung* (pp. 73-93). Münster: Waxmann.

Wild, K.-P., & Schiefele, U. (1992). Aufmerksamkeit als Mediator des Einflusses von Interesse auf die Lernleistung. Neubiberg: Gelbe Reihe.

Willson, V. L., & Putnam, R. R. (1982). A meta-analysis of pretest sensitization effects in experimental design. *American Educational Research Journal, 19*(2), 249-258.

Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology*, *82*(1), 51-59.

Zimmermann, B. J. (1989). Models of self-regulated learning and academic achievement. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement. Theory, research and practice* (pp. 1-25). New York: Springer.

A. Material

A 1. Test material for all children

A 1.1. Text and questions for assessing reading comprehension at pre-test

Die verschwundenen Mandan-Indianer

Anfang des 19. Jahrhunderts lebten etwa 1500 Mandan-Indianer in festen Dörfern am oberen Missouri. 1837 kam ein Dampfschiff den Fluss hinauf und weiße Siedler besuchten die Dörfer. Kurze Zeit später lebten nur noch 31 Indianer. Was war passiert?

Es hatte keinen Kampf gegeben. Es waren nicht die Waffen der Feinde, die diese Indianer töteten, sondern eine ansteckende Krankheit, gegen die die Indianer keinerlei Abwehrkräfte hatten. Zwei der Fremden, die dieses Dorf besucht hatten, waren mit Pocken infiziert. Die Mandan-Indianer waren die ersten von vielen indianischen Stämmen, die unter der großen Pockenepidemie zu leiden hatten, die die westlichen Ebenen zwischen 1837 und 1841 heimsuchte.

Die Mandan-Indianer sind deshalb von großem Interesse, da sie zu den fortschrittlichsten Prärie-Indianern gehörten. Sie jagten Büffel, betrieben aber auch Ackerbau, bauten Getreide, Bohnen, Kürbisse und Sonnenblumen an. Sie bauten runde Holzhäuser von 12 bis 18 Metern im Durchmesser; viel größer und stabiler als die meisten anderen indianischen Unterkünfte.

Mehrere Familien teilten sich ein Haus, aber für den Privatbereich hatte jede Familie einen durch einen Vorhang abgetrennten Raum in der Nähe der Außenwand. Die Mandan-Indianer hatten eigene Legenden und Zeremonien, und sie stellten Tonwaren und ungewöhnlich hübsche Körbe her. Dieser Lebensstil verschwand 1837 mit den Mandan-Indianern.

Die Pocken breiteten sich in den benachbarten Dörfern der Arikari- und Hidatsa-Indianer aus, wo rund die Hälfte der 4000 Einwohner starb. Dann verbreitete sich die Krankheit auch unter den als Nomaden lebenden Prärie-Indianern, die weiter westlich lebten. Sie traf die Blackfeet, die Crow, die Sioux, die Pawnee, die Osage, die Kiowa und die Comanchen.

Wieviele Indianer starben an den Pocken? Niemand wird es je wissen, da es keine Aufzeichnungen darüber gibt. Aber viele Stämme verloren die Hälfte ihrer Leute. Der Zufall ergab, daß die im Westen der USA lebenden Indianer zur Hälfte vernichtet waren, als die Soldaten sich auf den Weg machten, um sie zu besiegen.

Fragen zum Text „Die verschwundenen Mandan-Indianer"

1. Warum sind Anfang des 19. Jahrhunderts sehr viele Mandan-Indianer gestorben?

2. Kennst du noch andere Indianerstämme? Nenne ihre Namen. Was passierte mit diesen anderen Stämmen?

3. Wie lebten die Mandan-Indianer? Nenne einige Beispiele für Ihre Lebensweise!

4. Nenne mindestens zwei Dinge, in denen die Mandan-Indianer fortschrittlicher waren als andere indianische Stämme.

A 1.2. Text and questions for assessing reading comprehension at post-test.

Eine braune Welle von Ameisen

Joe lebte und arbeitete auf einer kleinen Farm im Lares-Tal von Peru. Dieses üppige, grüne Land ist bekannt für seine Tee- und Kakaoproduktion. Es ist ein reiches und schönes Tal. Joe erzielte reichliche Erträge von den Kakaopflanzen, die er auf seinem kleinen Stück Land anbaute. Zur Erntezeit pflückte er glücklich seine Kakaobohnen, brachte die Ernte zum Markt und konnte seine Familie mit dem Erlös gut versorgen.

Eines Morgens im Jahr 1969 erwachte Joe von einem fremdartigen knisternden Geräusch, so als ob sich ein Strohfeuer in seinen Feldern ausbreiten würde. Joe sprang von seinem Strohlager auf und rannte hinaus. Er sah keinen Rauch, aber das merkwürdige Geräusch schien näher zu kommen.

Er lief auf den kleinen Hügel hinter seiner Hütte und schaute im frühen Morgennebel in die Richtung, aus der das Geräusch zu kommen schien. Joe war geschockt, denn was er sah, war schlimmer als jedes Feuer hätte sein können. Sich langsam in Richtung der Kakao-Plantage bewegend, sah er Wellen von Millionen und aber Millionen riesiger brauner Ameisen. Die Ameisen fegten ins Tal nieder und erzeugten dabei Geräusche wie elektrische Rasenmäher.

Die Ameisen verwüsteten die grüne Vegetation des ganzen Tals – die kostbaren Tee- und Kakaopflanzen, die Früchte und das Gemüse, sogar das grüne Gras des Rasens – alles war verschwunden. Die Ameisen hinterließen hellbraunes, unfruchtbares Ödland.

Die Ameisen, die das Lares-Tal im Süden Perus im Jahr 1969 verwüsteten, waren bekannt unter dem Namen „Coqui", einer zerstörerischen Art von Blattschneideameisen. Diese Art von Ameisen hatte den Bauern in diesem Gebiet schon mehrere Male zuvor Probleme bereitet. Jeder Bauer versuchte, sich auf diese Insekten vorzubereiten, aber dieses Mal konnte keiner mit den Millionen von Ameisen fertig werden. Die bloße Zahl der Ameisen dieser Invasion machte es unmöglich, sie unter Kontrolle zu halten. Es waren einfach zu viele.

Noch schlimmer war, daß die Ameisen eine Vorliebe für die Tee- und Kakaoblätter hatten, die der Stolz dieses ganzen Tals waren. Eine ganze Tee- oder Kakaopflanze konnte von diesen hungrigen Insekten innerhalb von Sekunden vernichtet werden. Schließlich, nachdem sie die Ernte, die Tausende von Dollars wert war, vernichtet hatten, verließen die Ameisen die Farmen und steuerten in Richtung der offenen Felder. Schließlich wurden die Ameisen durch ihre eigene Überbevölkerung vernichtet. Es waren einfach zu viele, so daß es nicht genug Nahrung für sie gab. Das Land konnte diese große Anzahl nicht verkraften.

Obwohl viele Bauern durch diese Invasion der braunen Ameisen ruiniert waren, hätte die Situation noch schlimmer sein können. In einigen Teilen Südamerikas gibt es eine Art von Ameisen, die sowohl Tierfleisch als auch Pflanzen frißt. Diese Ameisen werden Armeeameisen genannt, und sie sind dafür bekannt, daß sie nicht nur ganze Wälder fressen, sondern auch die Tiere und Menschen, die ihnen in den Weg geraten.

Fragen zum Text „Die braune Welle von Ameisen"

1. Wo lebte Joe und womit verdiente er sein Geld?

2. Was ereignete sich 1969 an dem Ort, wo Joe lebte?

3. Warum war dieses Ereignis so schlimm für die Bevölkerung?

4. Hätte etwas noch Schlimmeres passieren können? Begründe deine Antwort?

5. Wer oder was sind „Coqui"?

A 2. Training Material

A 2.1. Worksheet Summarizing

Arbeitsblatt zum **Zusammenfassen** 1

Regel 1: Finde heraus, welches der Hauptaussagesatz ist.

Unterstreiche den Hauptaussagesatz, wenn Du ihn gefunden hast!

1 Computer sind sehr nützliche Maschinen, da sie sehr schnell arbeiten. Ein Computer kann einem Geschäftsführer sagen, wie viele Pizzen oder Kartons von Coca-Cola in den Regalen sind und dies in der Zeit, die man braucht um ein paar Knöpfe zu drücken. In weniger als 15 Sekunden verbinden Telefon-Computer Anrufer, die Tausende von Kilometern voneinander entfernt sind. Diese schnellen Maschinen können 2 Millionen Multiplikations-Aufgaben in einer Sekunde erledigen.

2 Im Herbst beginnen wilde Tiere Futtervorräte anzulegen und ihnen wächst ein langer Fellmantel. Viele Vögel fliegen in den Süden. Die Blätter verfärben sich und bekommen wunderschöne Farben, sterben dann und fallen auf den Boden. Die Tage werden kürzer und es wird kälter. An vielen Dingen kann man erkennen, dass der Winter kommt.

Regel 2: Erfinde einen Hauptaussagesatz, wenn es keinen gibt.

3 Wenn ein Vulkan ausbricht, werden geschmolzenes Gestein, Dampf und Asche durch die Öffnung des Berggipfels gezwungen. Das Gebiet rund um den Vulkan ist mit Asche und brodelnder Flüssigkeit, Lava genannt, umgeben. Bäume und Gebäude, die sich auf ihrem Weg befinden, werden zerstört. Die Tier- und Pflanzenwelt und auch Menschen werden getötet.

Kreuze an, was der Hauptaussagesatz für diesen Abschnitt sein könnte!

A. Die Insel Hawai entstand durch Vulkanausbrüche.
B. Städte, die sich in der Nähe von ausbrechenden Vulkanen befinden, können mit Asche bedeckt werden.
C. Wenn ein Vulkan ausbricht, zerstört er sowohl das Land als auch die Menschen, die sich um ihn herum befinden.

Arbeitsblatt zum **Zusammenfassen** 2

4 Inseln wie Hawai sind vollständig durch Vulkanausbrüche entstanden. In einigen Teilen der Erde wird der Dampf aus aktiven Vulkanen benutzt, um Kraftwerke für Fabriken oder Einfamilienhäuser zu betreiben. Wenn Lava in den Boden einzieht, hinterlässt sie Mineralien, die das Erdreich anreichern und fruchtbar machen. Die abgekühlte Lava wird auch als Belag im Straßenbau verwendet.

A. Viele Vulkane sind zu Touristenattraktionen geworden.
B. Lava und Dampf aus Vulkanen können für die Menschen sehr hilfreich sein.
C. Lava gibt der Erde Mineralien.

5 Koffein und Zucker in colahaltigen Getränken können deine Zähne regelrecht auffressen. Eine meiner Freundinnen war einmal „colasüchtig“. Ich schätze sie trank ungefähr 6 oder 7 Colas am Tag, von jüngster Kindheit an. Im Alter von 22 entdeckte sie, dass sie ihren Zahnschmelz auf ihren Zähnen verlor. Der Zahnarzt teilte ihr mit, dass der Verlust an Zahnschmelz die Folge ihres Colakonsums war.

Thema: ______________________

Hauptaussagesatz: ______________________

____

____

kurze Zusammenfassung des Inhalts: ______________________

____

____

____

____

6 Die riesigen Wolkenkratzer verdanken ihr Entstehen einem lustigen Vorfall: Die Frau des amerikanischen Architekten M. Jenney ließ aus Versehen ein dickes Buch auf einen Vogelkäfig fallen. Trotz der Wucht des Aufpralls verformte sich dieser nicht. Ihr Mann war erstaunt, dass sich die scheinbar zerbrechliche Drahtkonstruktion als so stabil erwies. Dies brachte ihn auf den Gedanken, Hochhäuser zu entwerfen, deren Metallkonstruktion der des Käfigs ähnelte.

Thema: _____

Hauptaussagesatz: _____

_____

_____

Regel 3: Laß unwichtige Informationen weg!

7 Amelia hat 3 Jeans. Eine ist marine-blau. Eine ist hell-blau (zu ihren Augen passend). Die dritte ist grün. Sie mag sie alle.

Streiche alle unwichtigen Informationen durch. Schreibe dann den Hauptaussagesatz auf die Linie.

_____

_____

8 England ist bekannt für sein köstliches Frühstück. Es handelt sich um eine herzhafte Mahlzeit. Man sollte eine ganze Stunde einplanen, um es zu sich zu nehmen. Es besteht aus Saft, Cerealien (Getreideflocken), Milch, Schinken mit Ei, Toast, Marmelade und Tee. Jeder einzelne Geschmack ist so speziell, dass es schwierig ist irgendetwas auszulassen.

Thema: _____

Hauptaussagesatz: _____

_____

_____

kurze Zusammenfassung des Inhalts: _____

_____

_____

_____

_____

Regel 4: Gib Abschnitten oder Aufzählungen einen Titel!

9 Nachdem du das Kuchenrezept gelesen hast, such dir die Zutaten zusammen. Zuerst rühre die Butter und den Zucker schaumig. Dann füge die Eier hinzu. Dann siebe das Mehl. Dann rühre alle trockenen Zutaten hinein. Zum Schluss füge die Flüssigkeit hinzu. Die Flüssigkeit kann Wasser sein. Rühre die Mischung gut durch. Fülle die Mischung in eine Backform. Backe alles 1 Stunde lang bei 200°C.

Dieser Abschnitt handelt von: _____

_____

_____

Titel: _____

10 Es macht immer wieder Spaß zuzusehen, wie eine Pizza gemacht wird. Der Pizzabäcker greift sich zunächst einen Teigklumpen und klopft ihn zu einem flachen Kuchen. Dann dreht er ihn auf seiner geschlossenen Faust und wirbelt ihn in der Luft umher, bis er zu einem flachen Pfannkuchen wird. Danach wird er vorsichtig auf einem Ofenschieber platziert und mit Käse, Fleisch und Tomatensoße belegt. Zum Schluss schiebt der Pizzabäcker den Schieber in einen speziellen heißen Ofen. In fünf Minuten wird die Pizza sehr heiß, wirft Blasen und bekommt eine braune Kruste.

Dieser Abschnitt handelt von: _____

_____

_____

Titel: _____

11 Wenn Julie in den Zoo geht, schaut sie sich am liebsten die Wellensittiche, Kanarienvögel, Papageien und Pfaue an.

Titel der Aufzählung: _____

_____

12 Ernährungswissenschaftler möchten herausfinden, wie viel Mineralien und Vitamine sich in Pizza, Hot Dogs, Hamburgern und Tacos befinden.

Titel der Aufzählung: _____

_____

13 Orangen, Zitronen, Mandarinen und Bananen werden in warmen Ländern angebaut.

Titel der Aufzählung: _____

_____

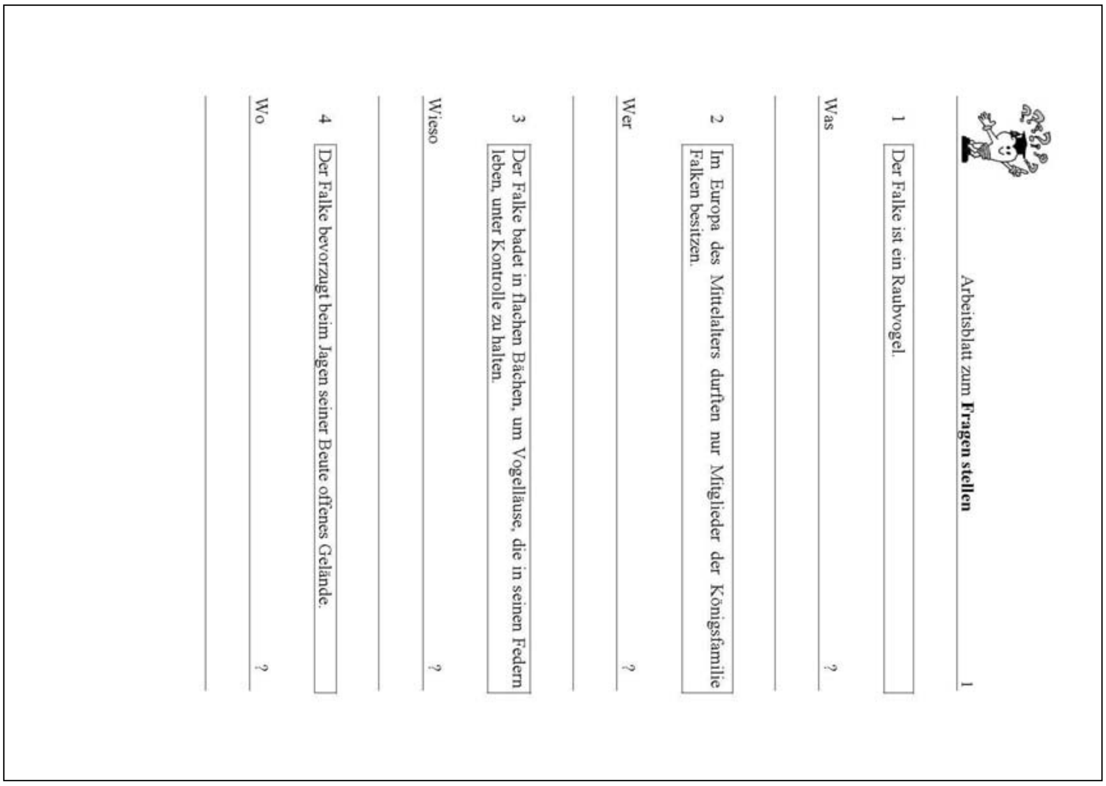
Regel 5: Laß überflüssige Informationen weg!

Unterstreiche die Textstellen, die man weglassen kann!

14 Der Nordpol gehört zu den kältesten Regionen der Erde. Die Temperatur fällt dort oft unter 0°C. Wettervorhersager berichten oft von Minustemperaturen dort. Die Winde sind sehr hart am Nordpol, haben zum Teil eine Stärke von ca. 100 km pro Stunde. Es ist recht windig am Nordpol.

15 In Afrika werden viele Sprachen gesprochen. In Westafrika werden 126 Hauptsprachen gesprochen. Jeder Stamm spricht eine andere Sprache. Arabisch ist die Hauptsprache in Nordafrika, während in Ostafrika hauptsächlich Swahili gesprochen wird. Afrika ist ein Kontinent mit vielen Sprachen. Nord- und Ostafrikaner sprechen verschiedene Sprachen.

A 2.2. Worksheet Questioning



Arbeitsblatt zum **Fragen stellen** 1

1 Der Falke ist ein Raubvogel.

Was ______ ?

2 Im Europa des Mittelalters durften nur Mitglieder der Königsfamilie Falken besitzen.

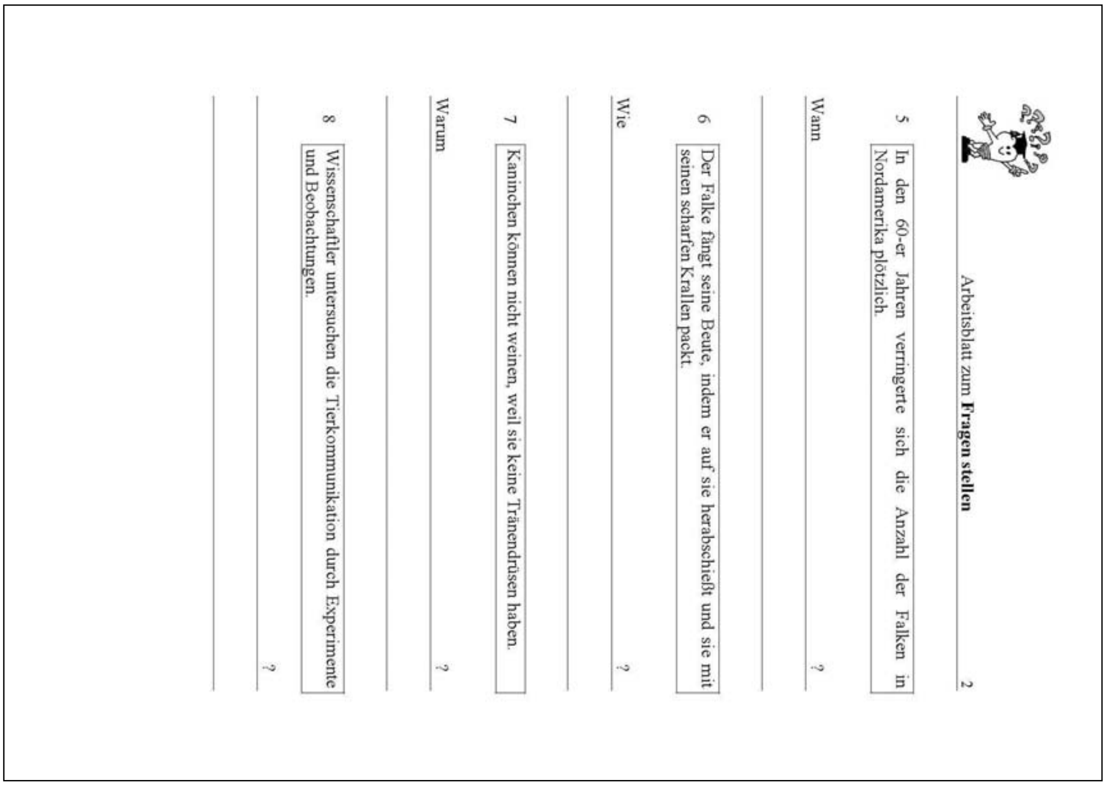
Wer ______ ?

3 Der Falke badet in flachen Bächen, um Vogelläuse, die in seinen Federn leben, unter Kontrolle zu halten.

Wieso ______ ?

4 Der Falke bevorzugt beim Jagen seiner Beute offenes Gelände.

Wo ______ ?



Arbeitsblatt zum **Fragen stellen** 2

5 In den 60-er Jahren verringerte sich die Anzahl der Falken in Nordamerika plötzlich.

Wann ______ ?

6 Der Falke fängt seine Beute, indem er auf sie herabschießt und sie mit seinen scharfen Krallen packt.

Wie ______ ?

7 Kaninchen können nicht weinen, weil sie keine Tränendrüsen haben.

Warum ______ ?

8 Wissenschaftler untersuchen die Tierkommunikation durch Experimente und Beobachtungen.

______ ?

9 Da Schlangen völlig taub sind, ist es die Bewegung des Schlangenbeschwörers, die die Schlange verzaubert und nicht die Musik, die er spielt.

_____ ?

_____

10 Einige Ameisen sondern einen speziellen Alarm-Duftstoff ab, der andere Ameisen, die in der Nähe sind, vor Gefahr warnt.

_____ ?

_____

11 Die Geräusche, die Fledermäuse, Nachtfalter und Wale machen, sind so hoch, daß sie vom Menschen nicht mehr gehört werden können.

_____ ?

_____

12 Todesfälle durch Schlangenbisse konnten in den letzten Jahren durch Impfstoffe aus Ameisengift, die gegen Schlangengifte wirken, reduziert werden.

___ 1. Warum beißen Schlangen Menschen?
___ 2. In welchen Ländern sterben einige Menschen an Schlangenbissen?
___ 3. Warum sterben heutzutage weniger Menschen an Schlangenbissen?

13 Im Gegensatz zu dem, was einige Leute glauben, stechen Schlangen nicht mit ihren Zungen. Sie brauchen ihre Zungen, um ihren Geruchssinn zu verstärken. Die Schlangen nehmen kleine Substanzpartikel aus der Luft auf und geben sie in zwei kleine Löcher am Ende ihrer Nasenlöcher, so können sie besser riechen.

___ 1. Wie viele Löcher hat eine Schlange am Ende ihrer Nasenlöcher?
___ 2. Wozu braucht eine Schlange ihre Zunge?
___ 3. Warum benutzen die Menschen den Spruch: "Er spricht mit gespaltener Zunge?

14 Es gibt viele Arten von Schlangen und sie kommen in unterschiedlichen Längen vor. Die kleinste Schlange ist gerade einmal so groß wie ein Wurm. Die längste Schlange, die man kennt, erreicht ca. 9 Meter Länge, das ist ungefähr so groß wie zwei Autos.

___ 1. Wie lang werden Schlangen?
___ 2. Wie viele Autos würden in 9 Meter hineinpassen?
___ 3. Wo kannst du die längsten Schlangen finden.

15 Schlangen sind sehr beweglich, weil ihr Körper wie ein Gummischlauch mit vielen Knochen ist. Tatsächlich kann die Wirbelsäule einer Schlange bis zu 300 Wirbel haben, fast zehnmal so viel wie die von einem Menschen. Wegen all dieser Knochen kann die Schlange ihren Körper in fast jede Richtung drehen.

___ 1. Warum kann eine Schlange ihren Körper so gut bewegen?
___ 2. Brauchen Schlangen gelegentlich Massagen so wie wir Menschen?
___ 3. Wie viele Rückenwirbel haben Schlangen?

16 Während sehr kleine Schlangen sehr kleine Insekten oder Würmer fressen, können große Schlangen sogar kleine Rehe, Leoparden und Ziegen fressen. Alle Schlangen, die Größe spielt dabei keine Rolle, fressen lebende Tiere oder Tier-Eier. Tatsache ist auch, daß einige Schlangen sich gegenseitig verschlingen.

___ 1. Welche Schlange frißt ihre Nachbarn?
___ 2. Was fressen Schlangen?
___ 3. Wie unterscheidet sich die Ernährungsweise kleiner Schlangen von der Ernährungsweise großer Schlangen?

17 Für Menschen, die in Wüsten leben, sind Kamele schon seit Tausenden von Jahren sehr hilfreich. Sie tragen die Menschen, wie auch deren Gepäck, auf ihren sonderbar geformten Rücken. Sie können Wanderungen durch Wüsten und Gebirge durchhalten, die zwei Monate dauern.

_____ ?

18 Wissenschaftler führten gründliche Untersuchungen über das Kamel durch, um herauszufinden, warum es dort überleben kann, wo andere Tiere sterben müssen. Sie fanden heraus, daß das Kamel erstaunlich gut geschaffen ist für das Leben in heißen, trockenen und sandigen Teilen der Erde. Es hat viele Körperteile, wie seine Füße, Beine, Augenwimpern und Nasenlöcher, die speziell an dieses Leben angepaßt sind.

_____ ?

19 Die Macateco-Indianer in Mexiko können sich miteinander unterhalten, ohne ein einziges Wort zu sprechen. Sie verständigen sich durch Pfeiflaute, die die Tonhöhe und den Rhythmus der gesprochenen Sprache nachahmen.

_____ ?

20 Viele Wörter, die wir benutzen, stammen aus der englischen Sprache, beispielsweise Computer. Aber auch Wörter wie Partner, Flirt, Sport und Training gehören dazu.

_____ ?

A 2.3. Worksheet Clarifying

Arbeitsblatt zum **Unklarheiten beseitigen** 2

4 Das Bushbaby ist ein kleines Tier, das zur Familie der Lemuren gehört. Lemuren sind eine Affenart. ***Es*** ist ungefähr so groß wie eine junge Katze und hat ein Gesicht mit einem seltsam menschlichen Ausdruck, sehr große runde Augen und schmalen spitzen Ohren.

D. ***Es*** bezieht sich auf: ______________________

____

5 Die Inkas waren berühmt für ihre baulichen Fähigkeiten. Beispiele ihrer Werke sind die antike Stadt von Machu Pichu und die älteste Brücke der Welt, die über den Fluss San Luis geht.

E. ***ihrer*** bezieht sich auf: ______________________

____

6 Menschen, die zweisprachig sind – das heißt, Menschen, die mehr als eine Sprache sprechen – werden mit zunehmender Nachfrage gesucht. Ihre Fähigkeiten werden bei Fluglinien, Schulen und Regierungen benötigt.

F. ***zweisprachig*** bedeutet: ______________________

____

7 Während des Sommers kommen die Vögel in die Mauser, oder verlieren ihre Federn.

G. ***Mauser*** bedeutet: ______________________

____

Arbeitsblatt zum **Unklarheiten beseitigen** 1

1 Es gibt Tiere, Lemminge genannt, die einige sehr unübliche Verhaltensweisen zeigen. Alle paar Jahre wird beobachtet, wie sie sich ins Meer stürzen, was wie Selbstmord aussieht.

A. ***Sie*** bezieht sich auf: ______________________

____

2 Ein Korallenriff ist eigentlich eine Sammlung von vielen Seetieren, die zusammen leben. Eins der wichtigsten Tiere, die in einem Korallenriff vorkommen, sind Schwämme. Sie heften sich an das Riff und versorgen es mit Nahrung.

B. ***es*** bezieht sich auf: ______________________

____

3 Wann immer ***sie*** sich selbst verletzte, lernte das Schimpansenweibchen Washoe das Zeichen für „verletzt“ oder „Schmerz“ zu machen. Zu einem späteren Zeitpunkt hat ***sie*** immer dann, wenn ***sie*** Menschen mit roten Flecken auf ihren Körpern sah, das Zeichen für „verletzt“ gemacht.

C. ***Sie*** bezieht sich auf: ______________________

____

8 Die Stadtbewohner dachten, dass der Berg mystische (magische) Kräfte besitzt.

H. *mystisch* bedeutet: _____

_____

9 KEIN AUFENTHALTSRAUM. NICHT HERUMLUNGERN. DIES IST KEINE WARTEHALLE.

I. *herumlungern* bedeutet: _____

_____

10 Der Redner schenkte dem Zwischenrufer, der immer wieder rüde Kommentare aus dem hinteren Teil des Raumes dazwischen rief, keinerlei Beachtung.

J. Ein *Zwischenrufer* ist: :_____

_____

11 Wie schneidet man Glas, das so zerbrechlich und doch so hart ist? Der Glaser benutzt dazu einen Glasschneider aus Wolfram, einem sehr harten Metall.

K. *Wolfram* ist: _____

_____

A 2.4. Worksheet for Predicting

Arbeitsblatt zum **Vorhersagen** 1

1 In die Sterne schauen

Ich könnte mir denken, daß diese Geschichte von ... handelt:

Was weißt Du bereits über dieses Thema?

2 Ich bestieg ganz allein den Mount Everest

Ich könnte mir denken, daß diese Geschichte von ... handelt:

Was weißt Du bereits über dieses Thema?

Arbeitsblatt zum **Vorhersagen** 2

3 Manchmal hören wir von Bränden, die großflächigen Schaden anrichten. Viele davon werden von unachtsamen Rauchern verursacht oder durch Kurzschlüsse. Kennst du die Legende, die von einer Kuh erzählt, die für das schrecklichste Feuer in Chicago verantwortlich gewesen sein soll?

Ich könnte mir denken, daß diese Geschichte von ... handelt:

Wie könnte der Titel des Artikels lauten?

4 a 30 Millionen Haushalte in Deutschland besitzen mindestens einen. In 87% der Haushalte war er täglich in Betrieb, und zwar durchschnittlich 208 Minuten. Das sind mehr als 3 Stunden! So verwundert es auch nicht, wenn immer weniger Menschen ins Kino oder zu Konzerten gehen, kaum noch Sport treiben und keine Zeit für Freunde und Bekannte haben.

Wie könnte der Artikel weitergehen?

Wie könnte der Titel des Artikels lauten?

4 b | Eindeutig hat auch das Lesen unter dem Fernsehkonsum gelitten. Kaum einer liest noch Bücher, die als Verfilmung im Fernsehen laufen ...

Wurde deine Vermutung bestätigt, wie der Artikel weitergeht?

_____

_____

5 | Immer wieder kann man in der Geschichte von großen, oftmals zufälligen Entdeckungen lesen. Eine der größten Entdeckungen des letzten Jahrtausends liegt nun mehr als 500 Jahre zurück. Damals glaubte ein bekannter italienischer Seefahrer, als er 1492 mit seinen Schiffen den Strand einer großen Insel erreicht hatte, endlich an seinem Ziel angekommen zu sein und das lang gesuchte Land zu betreten. Nachdem er den Strand hinter sich gelassen hatte, entdeckte er bald, daß er und seine Männer nicht die einzigen Menschen auf dieser Insel waren. Freundlich und dennoch mißtrauisch wurden der Seefahrer und seine Mannschaft von den Einheimischen begrüßt. Doch welches Land hatte er betreten?

Ich könnte mir denken, daß diese Geschichte von ... handelt:

_____

_____

Wie könnte der Titel der Geschichte lauten?

_____

_____

Wie könnte die Geschichte weitergehen?

_____

_____

A 2.5. Bookmark

Zusammenfassen

- kurze Fassung
- enthält Wichtigstes

Fragen stellen

- beginnen meist mit „W“
- haben ein Fragezeichen

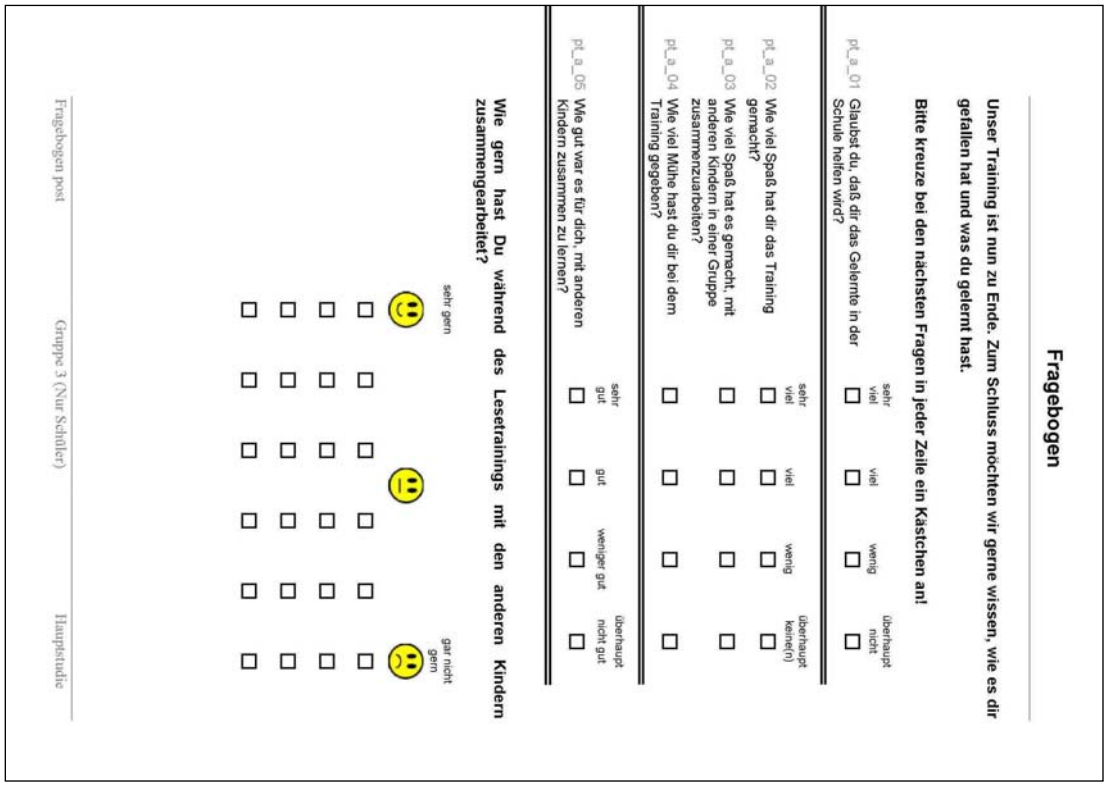
Vorhersagen

- Was kommt als nächstes?
- Was weiß ich schon über das Thema?

Unklarheiten beseitigen

- noch einmal lesen
- weiterlesen
- Hinweise suchen

A 2.6. Post-training questionnaire



Fragebogen

Unser Training ist nun zu Ende. Zum Schluss möchten wir gerne wissen, wie es dir gefallen hat und was du gelernt hast.

Bitte Kreuze bei den nächsten Fragen in jeder Zeile ein Kästchen an!

		sehr viel	viel	wenig	überhaupt nicht
pt_a_01	Glaubst du, daß dir das Gelernte in der Schule helfen wird?	☐	☐	☐	☐

		sehr viel	viel	wenig	überhaupt keine(n)
pt_a_02	Wie viel Spaß hat dir das Training gemacht?	☐	☐	☐	☐
pt_a_03	Wie viel Spaß hat es gemacht, mit anderen Kindern in einer Gruppe zusammenzuarbeiten?	☐	☐	☐	☐
pt_a_04	Wie viel Mühe hast du dir bei dem Training gegeben?	☐	☐	☐	☐

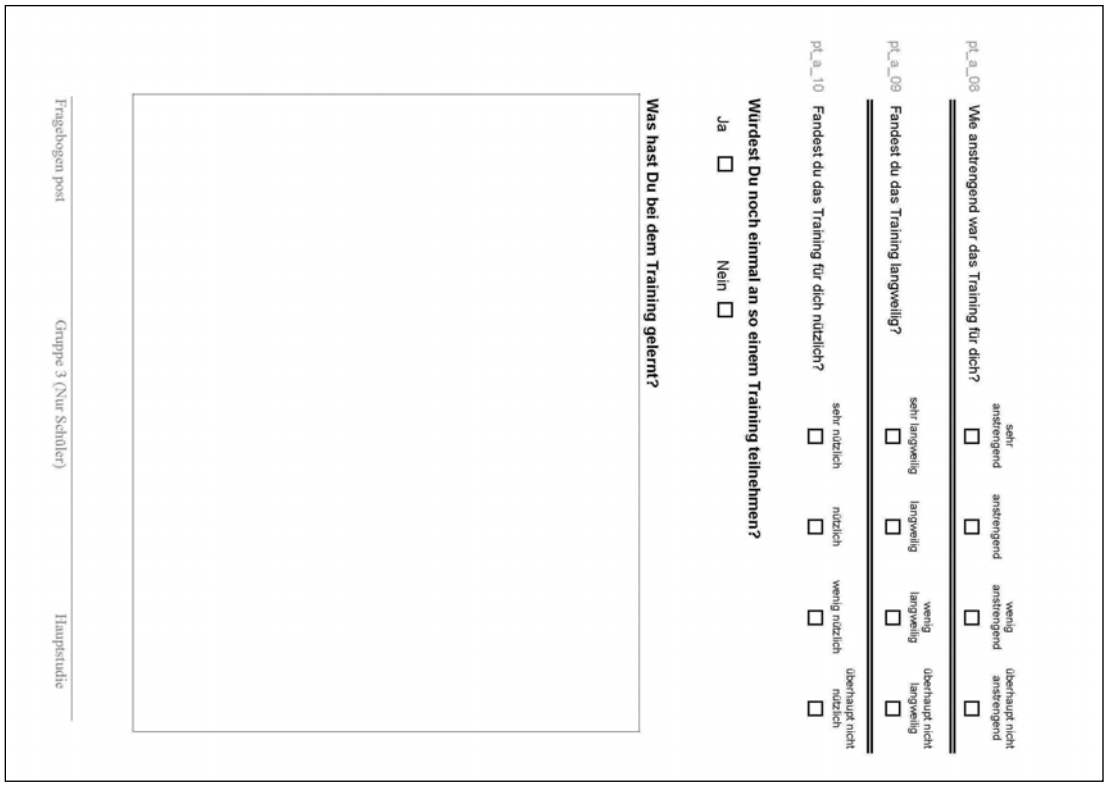
		sehr gut	gut	weniger gut	überhaupt nicht gut
pt_a_05	Wie gut war es für dich, mit anderen Kindern zusammen zu lernen?	☐	☐	☐	☐

Wie gern hast Du während des Lesetrainings mit den anderen Kindern zusammengearbeitet?

sehr gern — gar nicht gern

☐ ☐ ☐ ☐ ☐

Fragebogen post | Gruppe 3 (Nur Schüler) | Hauptstudie



		sehr anstrengend	anstrengend	wenig anstrengend	überhaupt nicht anstrengend
pt_a_08	Wie anstrengend war das Training für dich?	☐	☐	☐	☐

		sehr langweilig	langweilig	wenig langweilig	überhaupt nicht langweilig
pt_a_09	Fandest du das Training langweilig?	☐	☐	☐	☐

		sehr nützlich	nützlich	wenig nützlich	überhaupt nicht nützlich
pt_a_10	Fandest du das Training für dich nützlich?	☐	☐	☐	☐

Würdest Du noch einmal an so einem Training teilnehmen?

Ja ☐ Nein ☐

Was hast Du bei dem Training gelernt?

Fragebogen post | Gruppe 3 (Nur Schüler) | Hauptstudie

Vielleicht gibt es etwas, das dir besonders viel Spaß gemacht hat, oder etwas, das dir nicht gefallen hat. Schreibe es bitte in die Kästchen.

☺ Was ich besonders gut fand:

☹ Was mir nicht gefallen hat:

A 2.7. Text: Polar Bears

Eisbären

Der Eisbär ist eines der größten und stärksten fleischfressenden Tiere. Er kommt nur in der Arktis vor. Wenn er ausgewachsen ist wiegt er bis zu 1600 Pfund. Mit seinem langen, schweren Körper und schmalen Kopf wirkt er schwerfällig; er kann sich aber sehr schnell bewegen. Der Eisbär kann sich im Wasser leicht bewegen. Die meisten Tiere müssen heftig paddeln, um sich über Wasser zu halten, aber der Eisbär kann lange Zeit ruhig im Wasser treiben. Luftkammern in seinem Fell sowie eine ölige Haut und eine dicke Fettschicht ermöglichen es ihm, sich über Wasser zu halten.

Der Bär hat viele Körperteile, die ihn schützen. Er hat eine spezielle „Sonnenbrille" – ein drittes Augenlid, das seine Augen vor dem grellen Licht von Eis und Schnee schützt. Seine Zehen sind teilweise zusammengewachsen; dadurch kann er schwimmen. Haare auf der Unterseite der Füße ermöglichen es ihm, auf dem Eis zu laufen ohne auszurutschen. Seine Tatzen sind so scharf wie die einer Katze. Ein Eisbär kann Geräusche machen, um zu kommunizieren. Er faucht wie eine Katze, wenn er sich ärgert. Wenn er schwer verletzt ist, brüllt er, aber normalerweise ist er so leise wie der Schnee um ihn herum.

Neben seiner großen Kraft besitzt der Eisbär einen ausgeprägten Geruchssinn. Er kann Robbenfett auf 20 Meilen Entfernung riechen. Robben sind die bevorzugte Nahrung des Eisbären. Manchmal verbrennen Jäger Robbenfett, um einen in der Nähe befindlichen Bären anzulocken und ihn zu erlegen. Die schwarze Nase des Eisbären ist ein Problem. Sie ist das einzige Körperteil, das man gegen den weißen Schnee und das Eis sehen kann. Um sie zu verstecken, deckt der Eisbär sie manchmal beim Jagen mit seiner Tatze ab.

Der Eisbär ist ein geschickter Jäger. Im Herbst wandert er mit der Sonne nach Süden. Im Frühling macht er sich nach Norden auf, und hält sich auf dem Eis entlang der Küste auf. Dort kann er Robben finden. Auf seinen Reisen folgt der Bär der Bewegung von winzigen Seetieren, die Krill genannt werden. Der Bär ist immer unterwegs und kann bis zu 75 Meilen in einer Woche zurücklegen, um diesem tierischen Plankton zu folgen. Krill hält sich am Rande von schmelzenden Eisbergen auf, weil es nur in Wasser leben kann, das nicht zu salzig ist. Fische ernähren sich von Krill. Robben fressen die Fische und Bären fressen die Robben.

Der Eisbär hat eine besondere Methode, die Robben zu fangen. Er wartet auf die Robben, die mit ihren Köpfen durch Luftlöcher im Eis hervorlugen. Die meisten Arten der arktischen Robben bleiben ca. neun Minuten unter Wasser, dann tauchen sie 45 Sekunden lang auf, um zu atmen. Ein wartender Bär kann eine Robbe mit einem Hieb mit seiner riesigen Tatze aus dem Wasser schubsen und sie mit einem einzigen Biß töten. Der Eisbär ist ein kluger Jäger. Wenn die Sonne scheint, taucht manchmal eine Robbe auf, um sich auf das Eis neben die Luftlöcher zu legen. Sie hebt ihren Kopf sobald sie Gefahr spürt – aber der Eisbär gleitet langsam auf sie zu, dabei bewegt er sich nur dann, wenn der Kopf der Robbe unten ist. Schließlich ist er nahe genug, um zu springen und sich die Robbe zu packen.

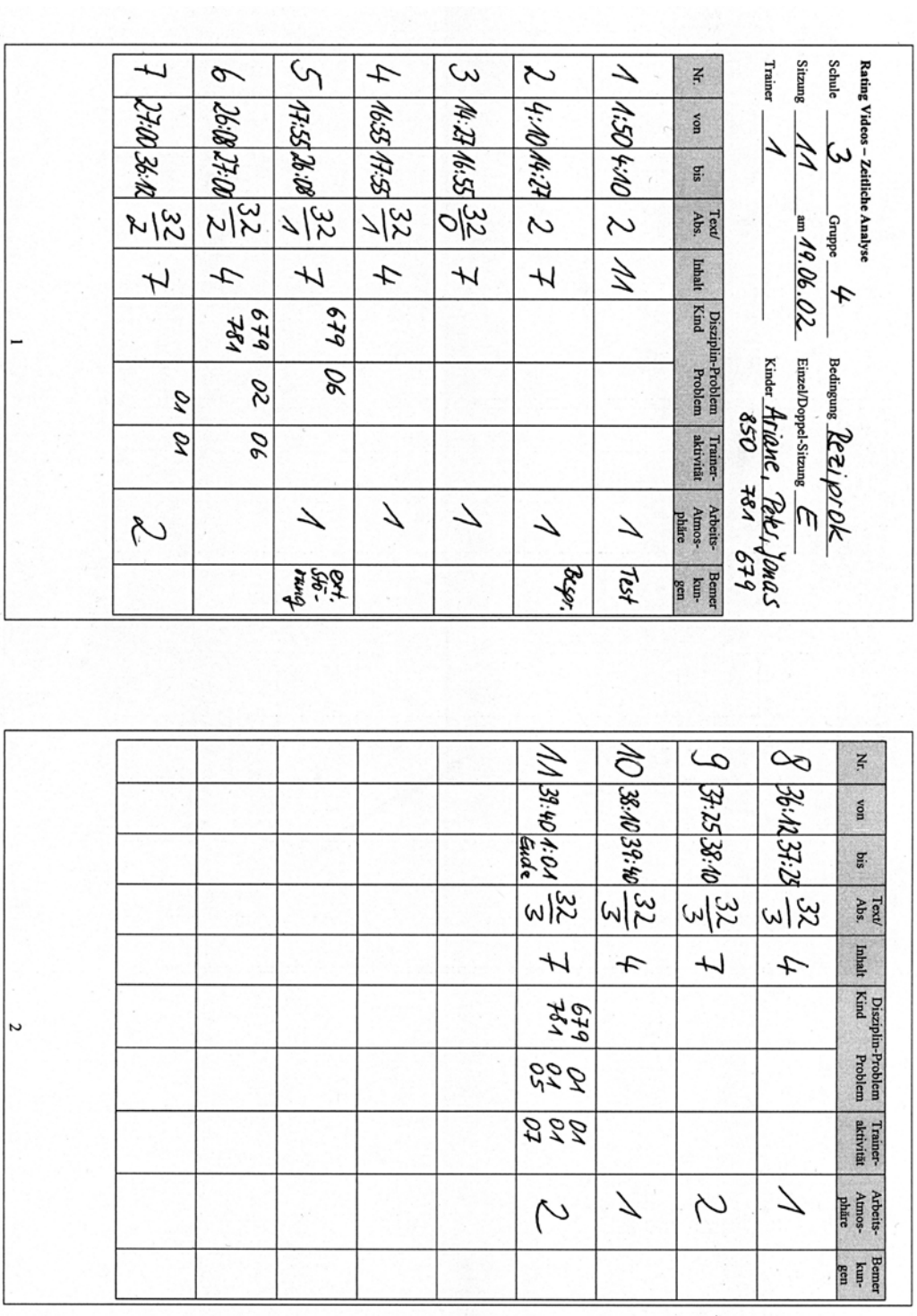
Neben dem Menschen sind die einzigen Tiere, die der Eisbär fürchten muß, der Killerwal und das Walroß. Killerwale fangen Bären unter Wasser. Das Walroß, das drei mal so groß ist wie der Eisbär, kann schneller schwimmen als der Bär und ihn ertränken. Aber das Walroß ist dumm und der Bär ist schlau. Einmal haben Eskimos einen Eisbären beobachtet, der sich an ein schlafendes Walroß heranschlich und seinen Kopf mit einem Eisblock zerschmetterte!

Die Eisbärin ist eine gute Mutter. Im Alter von drei oder vier Jahren paart sie sich im Frühling. Ihre Jungen werden dann im nächsten Januar geboren. Ihre Unterkunft ist eine Eis- oder Schneehöhle. Die Eisbärin bleibt mit ihren Jungen bis April in ihrer Höhle. Sie bleibt dort, obwohl Eisbären keinen Winterschlaf halten. Sie füttert ihre Jungen den Winter hindurch, während die sich an ihrem warmen Pelz kuscheln. Sie behält die Jungen zwei Jahre lang bei sich und lehrt sie das Jagen.

Obwohl die Lebenserwartung eines Eisbären etwa 25 Jahre beträgt, wird er heutzutage wahrscheinlich nicht mehr so alt. Die Menschheit rottet ihn aus. Zoos bezahlen etwa 500 Dollar für ein Eisbärenjunges. Jäger bekommen etwa 200 Dollar für das Fell. Wissenschaftler schätzen, daß es heutzutage nur noch etwa 25 000 Eisbären auf der Welt gibt. Eisbären sind gefährdet. Einige Länder billigen das Erlegen der Eisbären. Andere Länder tun das nicht mehr. Aber jedes Jahr reisen Eisbären über weite Strecken und überqueren dabei Landesgrenzen. Sie sind so lange gefährdet, wie die Jagd auf Eisbären von einigen Ländern geduldet wird.

B. Video rating

B 1. Session analysis (school 3, group 4, session 11)



Rating Videos – Zeitliche Analyse

Schule 3 Gruppe 4 Bedingung Reziprok

Sitzung 11 am 19.06.02 Einzel/Doppel-Sitzung E

Trainer 1 Kinder Ariane, Peter, Jonas
850 781 679

Nr.	von	bis	Text/ Abs.	Inhalt	Disziplin-Problem Kind	Problem	Trainer-aktivität	Arbeits-Atmos-phäre	Bemer-kun-gen
1	1:50	4:10	2	11			1	Test	
2	4:10	14:27	2	7			1	Bsp.	
3	14:27	16:55	32/0	7			1		
4	16:55	17:55	32/1	4			1		
5	17:55	26:08	32/1	7	679	06	1	ext. Stö-rung	
6	26:08	27:00	32/2	4	679 781	02	06		
7	27:00	36:10	32/2	7		01	01	2	

1

Nr.	von	bis	Text/ Abs.	Inhalt	Disziplin-Problem Kind	Problem	Trainer-aktivität	Arbeits-Atmos-phäre	Bemer-kun-gen
8	36:12	37:25	32/3	4				1	
9	37:25	38:10	32/3	7				2	
10	38:10	39:40	32/3	4				1	
11	39:40	1:01 Ende	32/3	7	679 781	01 01 05	01 01 07	2	

2

1. Anteil Redezeit: Trainer 10 % (zscr) + Kinder 90 % (zskr) = 100 %

2. globale Lärmpegeleinschätzung

Codes	Video – Frequenz	sehr leise 1	2	3	4	5	sehr laut 6
zsglp	Abschnitte	X					
zsglt	Textarbeit (Tafelarbeit, Unterstreichen, kurze schriftl. Zusammenfassung)						
zsgld	Test	X					
zsglo	Sonstiges (z.B. Organisatorisches, Begrüßung)						

3. globale Disziplin

Codes	Video - Frequenz	sehr diszipliniert 1	2	3	4	5	Absolutes Chaos 6
Zsgdp	Abschnitte		X				
Zsgdt	Textarbeit (Tafelarbeit, Unterstreichen, kurze schriftl. Zusammenfassung)						
Zsgdd	Test	X					
Zsgdo	Sonstiges (z.B. Organisatorisches, Begrüßung)						

4. Angemessenheit des Eingreifens des Trainers

Codes	Video - Frequenz	völlig angemessen 1	2	3	4	5	Völlig unangemessen 6
zsckp	Abschnitte	X					
zsckt	Textarbeit (Tafelarbeit, Unterstreichen, kurze schriftl. Zusammenfassung)						
zsckd	Test	X					
zscko	Sonstiges (z.B. Organisatorisches, Begrüßung)						

5. Arbeitsatmosphäre

Codes	Video – Frequenz	sehr gut 1	2	3	4	5	nahezu unmöglich 6
zsgzp	Abschnitte	X					
zsgzt	Textarbeit (Tafelarbeit, Unterstreichen, kurze schriftl. Zusammenfassung)						
zsgzd	Test	X					
zsgzo	Sonstiges (z.B. Organisatorisches, Begrüßung)						

6. Lärmeinschätzung differentiell – ganze Sitzung

Codes	Item	Abschnitte: trifft zu	trifft eher zu	trifft eher nicht zu	trifft nicht zu
zsglp01	Es ist ruhig	X	□	□	□
zsglp02	Man kann so gut wie jedes Wort verstehen.	X	□	□	□
zsglp03	Man muss teilweise schreien, um sich Gehör zu verschaffen.	□	□	□	X
zsglp04	Man kann Nebengeräusche gut hören (z.B. andere Kinder auf dem Flur oder draußen).	X	□	□	□
zsglp05	Die Kinder reden durcheinander.	□	□	□	X
zsglp06	Die Äußerungen einzelner Personen sind nicht mehr zu identifizieren.	□	□	□	X
zsglp07	Die Kinder müssen zur Ruhe aufgefordert werden.	□	□	□	X
zsglp08	Die Kinder haben Mühe, sich zu Wort zu melden.	□	□	□	X
zsglp09	Es ist so laut, dass man kaum arbeiten kann.	□	□	□	X

Codes	Item	Textarbeit: trifft zu	trifft eher zu	trifft eher nicht zu	trifft nicht zu
zsglt01	Es ist ruhig	□	□	□	□
zsglt02	Man kann so gut wie jedes Wort verstehen.	□	□	□	□
zsglt03	Man muss teilweise schreien, um sich Gehör zu verschaffen.	□	□	□	□
zsglt04	Man kann Nebengeräusche gut hören (z.B. andere Kinder auf dem Flur oder draußen).	□	□	□	□
zsglt05	Die Kinder reden durcheinander.	□	□	□	□
zsglt06	Die Äußerungen einzelner Personen sind nicht mehr zu identifizieren.	□	□	□	□
zsglt07	Die Kinder müssen zur Ruhe aufgefordert werden.	□	□	□	□
zsglt08	Die Kinder haben Mühe, sich zu Wort zu melden.	□	□	□	□
zsglt09	Es ist so laut, dass man kaum arbeiten kann.	□	□	□	□

6. Lärmeinschätzung differentiell – ganze Sitzung

Codes	Item	Test			
		trifft zu	trifft eher zu	trifft eher nicht zu	trifft nicht zu
zsgld01	Es ist ruhig	X	☐	☐	☐
zsgld02	Man kann so gut wie jedes Wort verstehen.	X	☐	☐	☐
zsgld03	Man muss teilweise schreien, um sich Gehör zu verschaffen.	☐	☐	☐	X
zsgld04	Man kann Nebengeräusche gut hören (z.B. andere Kinder auf dem Flur oder draußen).	X	☐	☐	☐
zsgld05	Die Kinder reden durcheinander.	☐	☐	☐	X
zsgld06	Die Äußerungen einzelner Personen sind nicht mehr zu identifizieren.	☐	☐	☐	X
zsgld07	Die Kinder müssen zur Ruhe aufgefordert werden.	☐	☐	☐	X
zsgld08	Die Kinder haben Mühe, sich zu Wort zu melden.	☐	☐	☐	X
zsgld09	Es ist so laut, dass man kaum arbeiten kann.	☐	☐	☐	X

Codes	Item	Sonstiges			
		trifft zu	trifft eher zu	trifft eher nicht zu	trifft nicht zu
zsglo01	Es ist ruhig	☐	☐	☐	☐
zsglo02	Man kann so gut wie jedes Wort verstehen.	☐	☐	☐	☐
zsglo03	Man muss teilweise schreien, um sich Gehör zu verschaffen.	☐	☐	☐	☐
zsglo04	Man kann Nebengeräusche gut hören (z.B. andere Kinder auf dem Flur oder draußen).	☐	☐	☐	☐
zsglo05	Die Kinder reden durcheinander.	☐	☐	☐	☐
zsglo06	Die Äußerungen einzelner Personen sind nicht mehr zu identifizieren.	☐	☐	☐	☐
zsglo07	Die Kinder müssen zur Ruhe aufgefordert werden.	☐	☐	☐	☐
zsglo08	Die Kinder haben Mühe, sich zu Wort zu melden.	☐	☐	☐	☐
zsglo09	Es ist so laut, dass man kaum arbeiten kann.	☐	☐	☐	☐

7. Einschätzung der Arbeitsweise der Gruppe

(Rollenübernahme – Identifikation der experimentellen Bedingung)

Code	Item	trifft zu	trifft eher zu	trifft eher nicht zu	Trifft nicht zu	Kam nicht vor
	Verhältnis Trainer – Kinder					
zsgy01	Der Trainer bestimmt, welches Kind wann was macht.	☐	☐	☐	X	
zsgy02	Der Trainer greift ein, wenn die Kinder ihre „Rolle" nicht erfüllen.	☐	X	☐	☐	☐
zsgy03	Der Trainer greift ein, wenn die Kinder die Strategien nicht richtig ausführen.	☐	X	☐	☐	☐
zsgy04	Der Trainer hält sich zurück.	☐	X	☐	☐	
zsgy05	Der Trainer weist darauf hin, wenn die Disziplin schlecht ist.	☐	☐	☐	☐	X
zsgy06	Der Trainer hat „die Fäden in der Hand".	☐	☐	X	☐	
zsgy07	Die Kinder werden vom Trainer ermutigt, sich gegenseitig zu helfen.	☐	X	☐	☐	
zsgy08	Eines der Kinder sagt, wann es zum nächsten Textabschnitt geht.	X	☐	☐	☐	
zsgy09	Es wird ein Kind bestimmt, dass die Leistung des anderen bewertet.	X	☐	☐	☐	
zsgy10	Ein Kind legt fest, wer eine bestimmte Strategie ausführt.	X	☐	☐	☐	
zsgy11	Die Kinder bestimmen, wer vorliest.	X	☐	☐	☐	
zsgy12	Wenn Unruhe herrscht, bemerken die Kinder das von selbst und fordern sich gegenseitig auf, leiser zu sein.	☐	☐	☐	☐	X
zsgy13	Die Kinder lassen sich gegenseitig ausreden.	X	☐	☐	☐	☐
zsgy14	Die Kinder fragen vorher den Trainer, wenn sie aufstehen möchten.	☐	☐	☐	☐	X

	Item	trifft zu	trifft eher zu	trifft eher nicht zu	Trifft nicht zu	Kam nicht vor
	Rückmeldung					☐
zsgu01	Die Schüler geben sich gegenseitig Rückmeldung über die erbrachte Leistung.	X	☐	☐	☐	

	Item	trifft zu	trifft eher zu	trifft eher nicht zu	Trifft nicht zu	Kam nicht vor
zsgu02	Der Trainer gibt nach der Ausführung einer Strategie eine Rückmeldung über die Richtigkeit des Inhalts der Aussage.	☐	☐	☐	☒	
zsgu03	Die Kinder fordern eine Rückmeldung ein, wenn keine gegeben wird (einschließlich Verbesserung).	☐	☐	☐	☐	☒
zsgu04	Bei der Rückmeldung wird nur gesagt, ob es gut oder schlecht war.	☐	☐	☐	☒	

8. Einschätzung der Arbeitsweise der Gruppe (Trainer)

Code	Der Trainer	trifft zu	trifft eher zu	trifft eher nicht zu	Trifft nicht zu	Kam nicht vor
	Verhalten des Trainers					
zscx01	Reagiert unangemessen heftig.	☐	☐	☐	☒	
zscx02	Interveniert bei Störungen so, dass der Trainingsablauf nicht unterbrochen wird.	☐	☐	☒	☐	
zscx03	Interveniert bei Störungen mit Tadel.	☐	☐	☐	☒	
zscx04	Ignoriert Störungen.	☐	☒	☐	☐	
zscx05	Lobt mehr als er tadelt.	☒	☐	☐	☐	
zscx06	Ermahnt ein Kind.	☐	☐	☐	☒	
zscx07	Wird ärgerlich.	☐	☐	☐	☒	
zscx08	Ist sehr streng.	☐	☐	☐	☒	
zscx09	Spricht mit den Kindern über persönliche Belange.	☐	☐	☐	☒	
zscx10	Ist inkonsistent bei Regelüberwachung.	☐	☐	☐	☒	
zscw	Hinweise auf (organisatorische) Regeln				☒	
zscw01	Bespricht explizit Regeln.	☐	☐	☐	☐	
zscw02	Bespricht Folgen von nicht eingehaltenen Regeln.	☐	☐	☐	☐	
zscw03	Verweist auf aufgestellte Regeln.	☐	☐	☐	☐	
zscw04	Geht über deutliche Regelverstöße hinweg.	☐	☐	☐	☐	☐

	Instruktion					
zstc01	Fragt Wissen über Strategien ab.	☐	☐	☐	☒	
zstc02	Erklärt die Strategien noch einmal.	☐	☐	☐	☒	
zstc03	Arbeitet mit Veranschaulichungshilfen (Tafelbild, Flip...).	☐	☐	☐	☒	
zstc04	Demonstriert laut. *(gibt Beispiele)*	☐	☐	☐	☒	

9. Einschätzung der Arbeitsweise in der Gruppe

Code	Item	trifft zu	trifft eher zu	trifft eher nicht zu	trifft nicht zu
zsgz01	Den Ablauf der Sitzung gestalten größtenteils die Kinder.	☒	☐	☐	☐
zsgz02	Haben die Kinder untereinander häufig Blickkontakt.	☒	☐	☐	☐
zsgz03	Haben die Kinder mit dem Trainer häufig Blickkontakt.	☒	☐	☐	☐
zsgz04	Es sind klare Regeln erkennbar.	☐	☐	☒	☐

10. Einschätzung der Rollenübernahme der Lehrer-Kinder (experimentelle Bedingung – RT)

Code	Item	trifft zu	trifft eher zu	trifft eher nicht zu	trifft nicht zu	kam nicht vor
zskj01	Der „Lehrer" macht Vorschläge, wie der Schüler seine Antwort verbessern kann.	☐	☐	☒	☐	
zskj02	Der „Lehrer" bestimmt, welches andere Kind eine Strategie ausführen soll.	☒	☐	☐	☐	
zskj03	Wenn das „Schüler"-Kind nicht weiter weiß, hilft der „Lehrer".	☐	☒	☐	☐	☐
zskj04	Der „Schüler" bekommt eine Rückmeldung.	☒	☐	☐	☐	
zskj05	Die „Lehrer"-Kinder sprechen laut und deutlich.	☒	☐	☐	☐	
zskj06	Die Anwendung der Strategie wird vom „Lehrer" bewertet.	☐	☒	☐	☐	
zskj07	Der „Lehrer" fordert den „Schüler" zur Verbesserung auf.	☐	☒	☐	☐	☐

11. Einschätzung der Rollenübernahme der Monitor-Kinder (experimentelle Bedingung – Monitor)

Code	Item	trifft zu	trifft eher zu	trifft eher nicht zu	trifft nicht zu	Kam nicht vor
zskm01	Die „Monitor"-Kinder schlagen vor, wie der andere seine Antwort verbessern kann.	☐	☐	☐	☐	
zskm02	Wenn das „Schüler"-Kind nicht weiter weiß, hilft der „Monitor".	☐	☐	☐	☐	☐
zskm03	Die „Schüler" bekommen eine inhaltliche Rückmeldung.	☐	☐	☐	☐	
zskm04	Die „Monitor"-Kinder sprechen laut und deutlich.	☐	☐	☐	☐	
zskm05	Die Anwendung der Strategie und nicht der Schüler werden von den „Monitor"-Kindern bewertet.	☐	☐	☐	☐	
zskm06	Das „Monitor"-Kind leitet den „Schüler" bei der Verbesserung an.	☐	☐	☐	☐	☐

12. Einschätzung der Bedingungserfüllung durch den Trainer

Bezogen auf die Abschnitte !

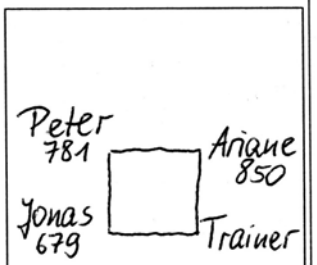
Die Sitzung entspricht hinsichtlich der Erfüllung der experimentellen Bedingung der Note...

Code	Bedingung	Note 1	2	3	4	5	6
zscs01	Reziprok	☒					
zscs02	Monitor						☒
zscs03	Schüler						☒

B 2. Paragraph analysis.

Datum: 04.09.02 Rater: H Rating Nr: 35

Schule	3
Bedingung	1
Sitzungsnummer	11
Text isgt	32
Abschnitt isgp	1
Videozeit isgpb isgpe isgpf	16:55- 26:08
Trainer	1
Kinder isk	1.
	2. Ariane
	3.
	4. Peter
	5.
	6. Jonas



Kind Nr. isk	Rolle ivko	Strategie ivkf	Freiwilligkeit ivkx	Ausführung ivkb	Hilfe ivkh0 ivkh1 ivkh2 ivkh3	Erfolg ivke (für die Strategien 1-4)
(bzw. T... Trainer)	1 Lehrerkind 2 Monitorkind	1 Zus.fassen 2 Fragen 3 Unklarheiten 3.1) aufwerfen 3.2) beseitigen 4 Vorhersagen 5 Vorlesen	0 nein 1 ja	0 nein 1 ja	0 keine Hilfe 1 Peers 2 Trainer 3 Lehrer-/Monitorkind	1 sehr gut 2 gut 3 befriedigend 4 ausreichend 5 mangelhaft 6 ungenügend
2	1					
4		5	0	1		
6		31	1	1		
2		32	1	1		
6		32	1	1		
4		31	1	1	2	

1

INHALTSANALYSE

Kind Nr. isk	Rolle ivko	Strategie ivkf	Freiwilligkeit ivkx	Ausführung ivkb	Hilfe ivkh0 ivkh1 ivkh2 ivkh3	Erfolg ivke (für die Strategien 1-4)
(bzw. T... Trainer)	1 Lehrerkind 2 Monitorkind	1 Zus.fassen 2 Fragen 3 Unklarheiten 3.1) aufwerfen 3.2) beseitigen 4 Vorhersagen 5 Vorlesen	0 nein 1 ja	0 nein 1 ja	0 keine Hilfe 1 Peers 2 Trainer 3 Lehrer-/Monitorkind	1 sehr gut 2 gut 3 befriedigend 4 ausreichend 5 mangelhaft 6 ungenügend
6		32	1	1	1	2
6		1	0	1		
4		2	0	1	1	3
6		2	1	1		
6		4	0	1		1

2

INHALTSANALYSE

1. Anteil Redezeit: Trainer 10 % (iscr) + Kinder 90 % (iskr) = 100 %

2. Einschätzung der Implementation der Strategien

Code	Item	trifft zu	trifft eher zu	trifft eher nicht zu	Trifft nicht zu	kam nicht vor
isgsn	**Zusammenfassung**					☒
isgsn01	Wenn eine Zusammenfassung verbesserungswürdig ist, weist jemand darauf hin.	☐	☐	☐	☐	☐
isgsn02	Wenn eine Zusammenfassung verbesserungswürdig ist, sagt das derjenige, der die Rückmeldung gibt.	☐	☐	☐	☐	☐
isgsn03	Der Trainer „hakt ein", wenn eine sehr schlechte Zusammenfassung gemacht wurde.	☐	☐	☐	☐	☐
isgsp	**Fragen**					☐
isgsp04	Es wird darauf hingewiesen, was eine gute Frage ausmacht.	☐	☐	☐	☒	
isgsp05	Wenn die Kinder eine knifflige Fragen zum Text stellen sollen, formulieren sie Unklarheiten.	☐	☐	☐	☒	
isgsp06	Die Kinder sagen selbst, wenn sie eine einfache Frage formuliert haben.	☒	☐	☐	☐	☐
isgsp07	Wenn eine Frage gestellt werden soll, überlegen die Kinder erst, bevor sie die Frage stellen.	☒	☐	☐	☐	
isgsp08	Die Fragen, die die Kinder stellen, werden in der Gruppe auch beantwortet.	☒	☐	☐	☐	
isgst	**Unklarheiten**					☐
isgst09	Die Kinder versuchen, sich gegenseitig die unklaren Wörter zu erklären.	☒	☐	☐	☐	
isgst10	Wenn die Kinder eine Unklarheit nicht selbst klären können, arbeiten sie mit dem Text.	☒	☐	☐	☐	☐
isgsp11	Der Trainer erklärt sofort eine Unklarheit, wenn die Kinder es selbst nicht wissen und es auch nicht anhand des Textes herausgefunden haben.	☐	☐	☐	☒	☐
isgsv	**Vorhersagen**					☐
isgsv12	Bei Vorhersagen wird versucht, auf die Hinweise im Text zu achten.	☐	☐	☒	☐	

3. Einschätzung der Arbeitsweise der Gruppe
3.1 Rollenübernahme – Identifikation der experimentellen Bedingung

Code	**Verhältnis Trainer - Kinder**					
	Item	trifft zu	trifft eher zu	trifft eher nicht zu	trifft nicht zu	kam nicht vor
isgy01	Der Trainer bestimmt, welches Kind wann was macht.	☐	☐	☐	☒	

INHALTSANALYSE

	Item	trifft zu	trifft eher zu	trifft eher nicht zu	trifft nicht zu	Kam nicht vor
isgy02	Der Trainer greift ein, wenn die Kinder ihre „Rolle" nicht erfüllen.	☐	☐	☐	☒	☐
isgy03	Der Trainer greift ein, wenn die Kinder die Strategien nicht richtig ausführen.	☐	☐	☐	☐	☒
isgy04	Der Trainer hält sich zurück.	☒	☐	☐	☐	
isgy05	Der Trainer weist darauf hin, wenn die Disziplin schlecht ist.	☐	☐	☐	☐	☒
isgy06	Der Trainer hat „die Fäden in der Hand".	☐	☐	☐	☒	
isgy07	Die Kinder werden vom Trainer ermutigt, sich gegenseitig zu helfen.	☐	☐	☐	☒	
isgy08	Eines der Kinder beendet den Textabschnitt.	☒	☐	☐	☐	
isgy09	Es wird ein Kind bestimmt, das die Leistung des anderen bewertet (nicht das Lehrerkind)	☐	☐	☐	☒	
isgy10	Ein Kind legt fest, wer eine bestimmte Strategie ausführt.	☒	☐	☐	☐	
isgy11	Die Kinder bestimmen wer vorliest.	☒	☐	☐	☐	☐
isgy12	Wenn Unruhe herrscht, bemerken die Kinder das von selbst und fordern sich gegenseitig auf, leiser zu sein.	☐	☐	☐	☐	☒
isgy13	Die Kinder achten darauf, dass der andere ausreden kann.	☒	☐	☐	☐	
isgy14	Die Kinder fragen vorher den Trainer, wenn sie aufstehen möchten.	☒	☐	☐	☐	☐
isgy15	Den Ablauf der Sitzung gestalten die Kinder.	☒	☐	☐	☐	

Code	Item	trifft zu	trifft eher zu	trifft eher nicht zu	trifft nicht zu	Kam nicht vor
isgu	**Rückmeldung**					☐
isgu01	Die Schüler geben sich gegenseitig Rückmeldung über die erbrachte Leistung.	☐	☒	☐	☐	
isgu02	Der Trainer gibt nach der Ausführung einer Strategie eine Rückmeldung über die Richtigkeit des Inhalts der Aussage.	☐	☐	☐	☒	
isgu03	Die Kinder fordern eine Rückmeldung ein, wenn keine gegeben wird (einschließlich Verbesserung).	☒	☐	☐	☐	☐
isgu04	Bei der Rückmeldung wird nur gesagt, ob es gut oder schlecht war.	☒	☐	☐	☐	
isgu05	Der Trainer lobt Schüler für **Leistungen**.	☐	☐	☐	☒	
isgu06	Der Trainer lobt Schüler für **Anstrengungen**.	☐	☐	☐	☒	
isgu07	Der Trainer lobt Schüler für **Leistungsverbesserungen**.	☐	☐	☐	☒	
isgu08	Der Trainer kritisiert Schüler für **Leistungen**.	☐	☐	☐	☒	
isgu09	Der Trainer kritisiert Schüler für **Anstrengungen**.	☐	☐	☐	☒	
isgu10	Der Trainer kritisiert Schüler für **Leistungsverschlechterungen**.	☐	☐	☐	☒	
isgu11	Der Trainer vergleicht die Leistungen von Schülern miteinander.	☐	☐	☐	☒	

INHALTSANALYSE

3.2 Der Trainer

Code	Der Trainer	trifft zu	trifft eher zu	trifft eher nicht zu	trifft nicht zu	Kam nicht vor
iscw	**Hinweise auf (organisatorische) Regeln**					☒
iscw01	Bespricht explizit Regeln.	☐	☐	☐	☐	
iscw02	Bespricht Folgen von nicht eingehaltenen Regeln.	☐	☐	☐	☐	
iscw03	verweist auf aufgestellte Regeln.	☐	☐	☐	☐	
iscw04	geht über deutliche Regelverstöße hinweg.	☐	☐	☐	☐	☐
iscq	**Instruktion**					☒
iscq01	fragt Wissen über Strategien ab.	☐	☐	☐	☐	
iscq02	erklärt die Strategien noch einmal.	☐	☐	☐	☐	
iscq03	arbeitet mit Veranschaulichungshilfen (Tafelbild...).	☐	☐	☐	☐	
iscq04	verweist auf Trainingsmaterial (Tafelbild...).	☐	☐	☐	☐	
iscq05	demonstriert laut.	☐	☐	☐	☐	
iscq06	bittet ein Kind eine Strategie zur Veranschaulichung zu demonstrieren.	☐	☐	☐	☐	
iscq07	bittet ein Kind, sich selbst zu bewerten.	☐	☐	☐	☐	
iscq08	fragt Vorwissen ab	☐	☐	☐	☐	
iscl	**Tempo**					
iscl01	lässt den Kindern Zeit zum Nachdenken bevor sie antworten	☒	☐	☐	☐	
iscl02	geht zum nächsten Kind, wenn ein Kind nicht sofort antwortet.	☐	☐	☐	☒	
iscl03	verweist auf die Zeit.	☐	☐	☐	☒	
iscl04	drängt die Kinder, schneller zu arbeiten	☐	☐	☐	☒	
iscl05	zeigt auf, was noch geschafft werden soll	☐	☐	☐	☒	
iscl06	erkundigt sich explizit, ob die Kinder den Abschnitt verstanden haben.	☐	☐	☐	☒	

4. globale Disziplin isgd

Kinder gut im Griff					Absolutes Chaos
1	2	3	4	5	6
☒	☐	☐	☐	☐	☐

INHALTSANALYSE

5. Eingreifen des Trainers hauptsächlich: iscx

Code	Der Trainer	trifft zu	trifft eher zu	trifft eher nicht zu	trifft nicht zu	Kam nicht vor
iscx01	Musste nicht eingreifen.					☒
iscx02	Griff nicht ein (obwohl es möglich gewesen wäre und nötig)					☐
iscx03	Sprach die Kinder sachlich an (vs. unsachlich)	☐	☐	☐	☐	
iscx04	Sprach die ganze Gruppe an (vs. nur ein einzelnes Kind)	☐	☐	☐	☐	
iscx05	Greift direkt ein (vs. indirekt)	☐	☐	☐	☐	
iscx06	Greift nonverbal ein (vs. verbal)	☐	☐	☐	☐	

6. Angemessenheit des Eingreifen des Trainers isck

Völlig angemessen					Völlig unange-messen
1	2	3	4	5	6
☒	☐	☐	☐	☐	☐

7. Arbeitsatmosphäre isgz

sehr gut					nahezu unmöglich
1	2	3	4	5	6
☒	☐	☐	☐	☐	☐

INHALTSANALYSE

8. Einschätzung der Rollenübernahme

8a) Rollenübernahme des Lehrer – Kindes (experimentelle Bedingung – Reziprokes Lehren)

Lehrer-Kind: 850 Ariane

Code	Item	trifft zu	trifft eher zu	trifft eher nicht zu	trifft nicht zu	kam nicht vor
zskxj01	Der „Lehrer" macht Vorschläge, wie der Schüler seine Antwort verbessern kann.	☐	☐	☐	☒	
zskxj02	Der „Lehrer" bestimmt, welches andere Kind eine Strategie ausführen soll.	☒	☐	☐	☐	
zskxj03	Wenn das „Schüler"-Kind nicht weiter weiß, hilft der „Lehrer".	☐	☐	☐	☐	☒
zskxj04	Der „Schüler" bekommt eine Rückmeldung.	☒	☐	☐	☐	
zskxj05	Das „Lehrer"-Kind spricht laut und deutlich.	☒	☐	☐	☐	
zskxj06	Die Anwendung der Strategie wird vom „Lehrer" bewertet. (Häufigkeiten)	☒	☐	☐	☐	
zskxj07	Der „Lehrer" fordert den „Schüler" zur Verbesserung auf.	☐	☐	☐	☒	☐

8b) Rollenübernahme des Monitor – Kindes (experimentelle Bedingung – Monitor)

Monitor-Kind: _____

Code.	Item	trifft zu	trifft eher zu	trifft eher nicht zu	trifft nicht zu	Kam nicht vor
zskxm01	Das „Monitor"-Kind schlägt vor, wie der andere seine Antwort verbessern kann.	☐	☐	☐	☐	
zskxm02	Wenn das „Schüler"-Kind nicht weiter weiß, hilft der „Monitor".	☐	☐	☐	☐	☐
zskxm03	Die „Schüler" bekommen eine inhaltliche Rückmeldung.	☐	☐	☐	☐	
zskxm04	Das „Monitor"-Kind spricht laut und deutlich.	☐	☐	☐	☐	
zskxm05	Die Anwendung der Strategie und nicht der Schüler wird vom „Monitor"-Kind bewertet.	☐	☐	☐	☐	
zskxm06	Das „Monitor"-Kind leitet den „Schüler" bei der Verbesserung an.	☐	☐	☐	☐	☐

INHALTSANALYSE

9. Einschätzung der Kinder

Code.	KIND 850 (Nummer) kx z.B. k01 k02 Ariane		Kam nicht vor
	Das Kind...		
ikxc01	bringt eigene Beiträge (ohne Aufforderung).	☐ ++ ☐ + ☐ - ☐ --	
ikxc02	schaut in der Gegend herum.	☐ ++ ☐ + ☐ - ☒ --	
ikxc03	Spielt mit Gegenständen.	☐ ++ ☐ + ☐ - ☒ --	
ikxc04	unterhält sich mit den anderen über Dinge, die nicht zum Training gehören.	☐ ++ ☐ + ☐ - ☒ --	
ikxc05	beteiligt sich aktiv.	☐ ++ ☒ + ☐ - ☐ --	
ikxc06	ist motorisch unruhig.	☐ ++ ☐ + ☐ - ☒ --	
ikxc07	steht auf und macht andere Dinge.	☐ ++ ☐ + ☐ - ☒ --	
ikxc08	Albert herum.	☐ ++ ☐ + ☐ - ☒ --	
ikxc09	meldet sich, um zu antworten, auch wenn ein anderes Kind schon aufgefordert wurde.	☐ ++ ☐ + ☐ - ☒ --	
Ikxc10	unterbricht ein anderes beim Antwortversuch	☐ ++ ☐ + ☐ - ☒ --	
ikxc11	hilft einem anderen (auch ohne Aufforderung).	☐ ++ ☐ + ☐ - ☒ --	
ikxc12	ermuntert ein anderes, auch bei Schwierigkeiten die Anwendung einer Strategie zu versuchen.	☐ ++ ☐ + ☐ - ☒ --	
ikxc13	Reicht einem anderen Arbeitsmaterialien.	☐ ++ ☐ + ☐ - ☒ --	
ikxc14	verweist auf Regeln in der Gruppe.	☐ ++ ☐ + ☐ - ☒ --	
ikxc15	hört den anderen aufmerksam zu.	☒ ++ ☐ + ☐ - ☒ --	
ikxc16	gibt negatives Feedback.	☐ ++ ☐ + ☐ - ☒ --	☐
ikxc17	formuliert Kritik sachlich.	☒ ++ ☐ + ☐ - ☐ --	☐
ikxc18	wiederholt Äußerungen für andere Kinder (ohne Aufforderung des Trainers).	☐ ++ ☐ + ☐ - ☐ --	☒
ikxc19	**Time on Task – Prozentsatz**	100 %	
ikxu	Kind erhält Feedback.	☒ nein	
ikxu20	Wird gelobt für seine LEISTUNG.	☐ ja	
ikxu21	Wird gelobt für seine LEISTUNGSVERBESSERUNG.	☐ ja	
ikxu22	Wird gelobt für seine ANSTRENGUNG.	☐ ja	
ikxu23	Wird kritisiert für seine LEISTUNG.	☐ ja	
ikxu24	Wird kritisiert für seine LEISTUNGSVERSCHLECHTERUNG.	☐ ja	
ikxu25	Wird kritisiert für seine ANSTRENGUNG.	☐ ja	

INHALTSANALYSE

9. Einschätzung der Kinder

Code.	KIND 781 (Nummer) kx z.B. k01 k02 Peter					Kam nicht vor
	Das Kind...					
ikxc01	bringt eigene Beiträge (ohne Aufforderung).	☐ ++	☒ +	☐ -	☐ --	
ikxc02	schaut in der Gegend herum.	☐ ++	☐ +	☐ -	☒ --	
ikxc03	Spielt mit Gegenständen.	☐ ++	☐ +	☐ -	☒ --	
ikxc04	unterhält sich mit den anderen über Dinge, die nicht zum Training gehören.	☐ ++	☐ +	☐ -	☒ --	
ikxc05	beteiligt sich aktiv.	☒ ++	☐ +	☐ -	☐ --	
ikxc06	ist motorisch unruhig.	☐ ++	☐ +	☐ -	☒ --	
ikxc07	steht auf und macht andere Dinge.	☐ ++	☐ +	☐ -	☒ --	
ikxc08	Albert herum.	☐ ++	☐ +	☐ -	☒ --	
ikxc09	meldet sich, um zu antworten, auch wenn ein anderes Kind schon aufgefordert wurde.	☐ ++	☐ +	☐ -	☒ --	
Ikxc10	unterbricht ein anderes beim Antwortversuch	☐ ++	☐ +	☐ -	☒ --	
ikxc11	hilft einem anderen (auch ohne Aufforderung).	☐ ++	☐ +	☐ -	☒ --	
ikxc12	ermuntert ein anderes, auch bei Schwierigkeiten die Anwendung einer Strategie zu versuchen.	☐ ++	☐ +	☐ -	☒ --	
ikxc13	Reicht einem anderen Arbeitsmaterialien.	☐ ++	☐ +	☐ -	☒ --	
ikxc14	verweist auf Regeln in der Gruppe.	☐ ++	☐ +	☐ -	☒ --	
ikxc15	hört den anderen aufmerksam zu.	☒ ++	☐ +	☐ -	☐ --	
ikxc16	gibt negatives Feedback.	☐ ++	☐ +	☐ -	☐ --	☒
ikxc17	formuliert Kritik sachlich.	☐ ++	☐ +	☐ -	☐ --	☒
ikxc18	wiederholt Äußerungen für andere Kinder (ohne Aufforderung des Trainers).	☐ ++	☐ +	☐ -	☐ --	☒
ikxc19	**Time on Task – Prozentsatz**	100 %				
ikxu	Kind erhält Feedback.	☐ nein				
ikxu20	Wird gelobt für seine LEISTUNG.	☒ ja				
ikxu21	Wird gelobt für seine LEISTUNGSVERBESSERUNG.	☐ ja				
ikxu22	Wird gelobt für seine ANSTRENGUNG.	☐ ja				
ikxu23	Wird kritisiert für seine LEISTUNG.	☐ ja				
ikxu24	Wird kritisiert für seine LEISTUNGSVERSCHLECHTERUNG.	☐ ja				
ikxu25	Wird kritisiert für seine ANSTRENGUNG.	☐ ja				

9

INHALTSANALYSE

9. Einschätzung der Kinder

Code.	KIND 679 (Nummer) kx z.B. k01 k02 Jonas					Kam nicht vor
	Das Kind...					
ikxc01	bringt eigene Beiträge (ohne Aufforderung).	☒ ++	☐ +	☐ -	☐ --	
ikxc02	schaut in der Gegend herum.	☐ ++	☐ +	☐ -	☒ --	
ikxc03	Spielt mit Gegenständen.	☐ ++	☐ +	☐ -	☒ --	
ikxc04	unterhält sich mit den anderen über Dinge, die nicht zum Training gehören.	☐ ++	☐ +	☐ -	☒ --	
ikxc05	beteiligt sich aktiv.	☒ ++	☐ +	☐ -	☐ --	
ikxc06	ist motorisch unruhig.	☐ ++	☒ +	☐ -	☐ --	
ikxc07	steht auf und macht andere Dinge.	☐ ++	☐ +	☐ -	☒ --	
ikxc08	Albert herum.	☐ ++	☐ +	☐ -	☒ --	
ikxc09	meldet sich, um zu antworten, auch wenn ein anderes Kind schon aufgefordert wurde.	☐ ++	☐ +	☐ -	☒ --	
Ikxc10	unterbricht ein anderes beim Antwortversuch	☐ ++	☐ +	☐ -	☒ --	
ikxc11	hilft einem anderen (auch ohne Aufforderung).	☐ ++	☐ +	☐ -	☒ --	
ikxc12	ermuntert ein anderes, auch bei Schwierigkeiten die Anwendung einer Strategie zu versuchen.	☐ ++	☐ +	☐ -	☒ --	
ikxc13	Reicht einem anderen Arbeitsmaterialien.	☐ ++	☐ +	☐ -	☒ --	
ikxc14	verweist auf Regeln in der Gruppe.	☐ ++	☐ +	☐ -	☒ --	
ikxc15	hört den anderen aufmerksam zu.	☒ ++	☐ +	☐ -	☐ --	
ikxc16	gibt negatives Feedback.	☐ ++	☐ +	☐ -	☒ --	☐
ikxc17	formuliert Kritik sachlich.	☐ ++	☐ +	☐ -	☐ --	☒
ikxc18	wiederholt Äußerungen für andere Kinder (ohne Aufforderung des Trainers).	☐ ++	☐ +	☐ -	☐ --	☒
ikxc19	**Time on Task – Prozentsatz**	100 %				
ikxu	Kind erhält Feedback.	☐ nein				
ikxu20	Wird gelobt für seine LEISTUNG.	☒ ja				
ikxu21	Wird gelobt für seine LEISTUNGSVERBESSERUNG.	☐ ja				
ikxu22	Wird gelobt für seine ANSTRENGUNG.	☐ ja				
ikxu23	Wird kritisiert für seine LEISTUNG.	☐ ja				
ikxu24	Wird kritisiert für seine LEISTUNGSVERSCHLECHTERUNG.	☐ ja				
ikxu25	Wird kritisiert für seine ANSTRENGUNG.	☐ ja				

10

C. Data analyses

C 1. Absolute agreement (intra-class-correlations) of raters for text comprehension scores

Measure		Text	The mandan-indians (pre) N	The mandan-indians (pre) ICC	A brown wave of ants (post) N	A brown wave of ants (post) ICC
Quantity	points on competence level 1		68	.929	72	.956
	points on competence level 2		68	.958	72	.904
	points on competence level 3		68	.782	72	.955
	total number of points		68	.958	72	.940
Quality	number of main ideas: very central		46	.987	47	.989
	number of main ideas: less central		46	.968	47	.901
	total number of main ideas		46	.986	47	.975
	number of details: more important		46	.943	47	.943
	number of details: less important		46	.951	47	.988
	total number of details		46	.943	47	.981

C 2. Factor loading matrix for items of session analysis on the level of group and trainer behaviour (N = 168)

Item		Factor					
Variable	Label	1	2	3	4	5	6
ZSGY11	Ch. decide who reads aloud	**.945**	.016	.175	-.015	.006	-.061
ZSGY10.	Ch. choose strategy	**.933**	-.008	.214	-.015	.029	-.030
ZSGY01	T. decides	**-.919**	.030	-.154	-.029	-.028	.086
ZSGY08	Ch. complete section	**.909**	-.054	.205	-.017	-.026	-.021
ZSGZ01	Ch. shape	**-.817**	.184	-.117	-.082	.064	.160
ZSGY04	T. takes a back seat	**.679**	-.208	.089	.008	-.043	-.019
ZSGU02	T. gives feedback	**-.634**	.104	.095	-.015	.020	-.048
ZSGZ04	clear rules apparent	.147	**.838**	-.175	.114	-.057	-.178
ZSCX04	T. ignores disruptions	.175	**-.803**	-.248	.116	-.035	-.096
ZSGY06	T. in command	-.358	**.739**	.286	-.012	.054	-.021
ZSCX10	enforcement of rules by T. is inconsistent	-.044	**-.676**	-.187	.158	-.043	-.070
ZSCX02	T. intervenes without interrupting	-.031	**.669**	-.247	.082	.145	-.089
ZSCW04	T. ignores rule breaking	.012	**-.553**	-.135	.051	.491	-.220
ZSGU01	Ch. give each other feedback	.221	.143	**.867**	.028	-.054	.010
ZSGY09	Ch. evaluates performance	.263	.128	**.819**	.045	.037	-.062
ZSGY07	T. encourages ch. to help	.115	.299	**.801**	.083	.045	-.002
ZSGY02	T. intervenes: role	.511	.044	**.609**	.130	.014	.190
ZSCX03	T. admonishes	-.019	.028	-.037	**.745**	-.067	.156
ZSGY05	T. mentions discipline	.072	-.021	.040	**.727**	.269	-.094

Item		Factor					
Variable	Label	1	2	3	4	5	6
ZSCX06	T. reprimands	-.087	-.210	.218	**.727**	.206	-.025
ZSCX07	T. annoyed	-.137	.004	.100	**.669**	.168	.014
ZSCX05	T. praises more than admonishes	-.168	.350	.162	**-.458**	.006	.078
ZSCX01	T. overreacts	.046	.181	-.039	**.437**	-.309	-.060
ZSGY12	Ch. notice noise and disturbance themselves	.165	-.233	.055	**.434**	.078	-.011
ZSCW03	T. refers to rules	-.064	.100	.035	.100	**.880**	.055
ZSCW01	T. discusses rules	.059	.162	.005	.065	**.850**	.199
ZSCW02	T. discusses consequences	-.054	.039	.015	.249	**.845**	.064
ZSTC02	T. explains strategies	.019	.180	.054	.041	.051	**.802**
ZSTC04	T. demonstrates using examples	.060	-.061	-.115	-.111	.070	**.719**
ZSTC03	T. uses aids	-.269	-.028	-.029	-.035	.067	**.684**
ZSTC01	T. taps knowledge	-.043	.316	.232	.162	.066	**.578**
Portion of variance explained		18.45	12.11	9.87	9.23	8.91	7.21
Internal consistency		.94	84	.86	.68	.89	.68

Legend

Factor 1	Co-determination of children
Factor 2	Effective enforcement of rules by trainer
Factor 3	Feedback in the group
Factor 4	Discipline management (trainer)
Factor 5	Clarity of rules (trainer)
Factor 6	Instruction (trainer)

C 3. Factor loading matrix for Fulfillment of Roles in Reciprocal Teaching Conditions for Session Analysis (N=170)

Item		Factor		
Variable	Label	1	2	3
ZSKJ06	"Teacher" evaluates strategy use.	**.805**	.280	.228
ZSKJ04	"Student" is given feedback.	**.789**	.345	.012
ZSKJ05*	"Teachers" speak loudly and clearly.	**.720**	.253	-.288
ZSKJ03	"Teacher" helps the student.	-.035	**.847**	.109
ZSKJ07	"Teacher" calls for improvement.	.178	**.744**	-.218
ZSKJ01	"Teacher" suggests ways for the student to improve his/her answer.	.447	**.638**	.171
ZSKJ02	"Teacher" decides which other child is to apply a strategy.	.018	-.025	**.945**
Portion of variance explained		28.89	56.61	72.58
Internal consistency		.78	.67	--
* Item dropped to improve internal consistency of scale				

Legend

Factor 1 Fulfilment of central tasks

Factor 2 "Teacher" provides help and guidance to improve

Factor 3 no scale constructed

C 4. Factor loading matrix for Fulfillment of Roles in Monitor Conditions for Session Analysis (N=170)

Item		Factor	
Variable	Label	1	2
ZSKM03	"Students" are given feedback.	**.792**	.107
ZSKM04	"Monitors" speak loudly and clearly.	**.769**	-.090
ZSKM01 *	"Monitors" suggest ways for the other child to improve his/her answer.	**.719**	.187
ZSKM02	"Monitor" helps when the "student" gets stuck.	.145	**.784**
ZSKM06	"Monitor" gives the "student" guidance to improve.	.265	**.740**
ZSKM05 *	"Monitor" evaluates strategy use.	-.124	**.514**
Portion of variance explained		30.69	55.39
Internal consistency		.60	.58
* Items dropped to improve internal consistency of scales			

Legend

Factor 1 Fulfilment of central tasks

Factor 2 "Monitor" provides help and guidance to improve

C 5. Factor loading matrix for items of paragraph analysis on the level of group and trainer behaviour (N = 495)

Item		Factor							
Variable	label	1	2	3	4	5	6	7	8
ISGY10	Ch. choose strategy	**.952**	-.010	-.017	-.039	.049	.086	-.032	.086
ISGY01	T. decides	**-.928**	.036	.030	.074	.008	-.069	.000	-.062
ISGY15	Ch. shape	**.899**	-.032	-.023	-.056	.042	.028	-.074	.041
ISGY08	Ch. complete section	**.865**	-.017	.004	-.050	.056	.051	-.005	.068
ISGY11	Ch. decide who reads aloud	**.863**	-.002	.010	-.082	.065	.093	-.025	.078
ISGY06	T. in command	**-.673**	.148	.029	.004	.149	.017	.013	.339
ISGY04	T. takes a back seat	**.564**	-.314	-.010	-.128	-.085	-.244	.327	-.070
ISGU02	T. gives feedback	**-.508**	.240	.005	.078	.389	.031	-.144	.174
ISCQ04	T. refers to material	-.053	**.826**	.066	.055	.053	.018	.030	-.032
ISCQ06	T. asks a ch. to demonstrate	-.054	**.810**	.017	.156	.059	.032	-.008	.028
ISCQ08	T. taps prior knowledge	-.048	**.797**	.073	-.044	.036	-.067	.148	-.079
ISCQ02	T. explains strategies	-.030	**.788**	.016	.018	.132	.087	-.129	.075
ISCQ01	T. taps knowledge	-.017	**.780**	-.061	.058	.218	.016	-.019	.003
ISCQ05	T. demonstrates	-.068	**.773**	.084	-.019	-.009	.091	-.077	-.070
ISCQ03	T. uses aids	-.029	**.764**	.141	.091	.035	-.031	.108	.007
ISCQ07	T. asks a ch. to rate his/her own performance	-.058	**.747**	.154	.039	.108	-.061	.085	.036
ISGY03	T. intervenes: strategy	-.274	**.446**	.153	.057	.236	.251	-.321	.212
ISCX03	T. intervenes: content	-.050	.139	**.929**	.076	.024	-.009	-.005	-.047
ISCX05	T. intervenes: directly	-.095	.078	**.868**	.073	.020	.043	-.044	-.108
ISCX04	T. intervenes: group	.005	.077	**.820**	.028	.052	.018	-.018	-.005
ISCX06	T. intervenes: nonverbally	.014	.071	**.779**	.001	-.079	-.065	.012	.027

Item		Faktoren							
Variable	Bezeichnung	1	2	3	4	5	6	7	8
ISGY05	T. mentions discipline	.028	.095	**.637**	.116	-.106	.216	-.052	-.363
ISCW02	T. discusses consequences	-.105	.058	.060	**.918**	.005	-.009	.008	-.046
ISCW03	T. refers to rules	-.081	.081	.072	**.866**	.060	-.025	.046	.055
ISCW01	T. discusses rules	-.062	.110	.122	**.860**	.029	-.009	.003	-.038
ISCW04 *	T. ignores rule breaking	-.070	.037	.003	**.770**	.001	-.018	-.019	-.001
ISGU01	Ch. give each other feedback	.087	.145	-.001	.060	**.828**	-.015	.020	.095
ISGY09	Ch. evaluates performance	-.363	.065	.033	.117	**.732**	-.133	.026	-.043
ISGY02	T. intervenes: role	.302	.190	.026	-.002	**.688**	-.016	-.066	.084
ISGY07	T. encourages ch. to help	-.040	.091	-.090	-.045	**.489**	.116	.059	-.051
ISCL05	T. points out what still has to be done	.061	-.034	.085	-.068	-.063	**.751**	.156	-.073
ISCL03	T. mentions time	.068	.070	.095	.002	-.095	**.709**	.230	.019
ISCL04	T. hurries ch.	.068	.038	-.071	-.022	.240	**.676**	.164	.012
ISCL01	T. lets ch. take time	.037	-.007	-.016	.081	-.034	.436	**.649**	.010
ISCL06	T. calls for feedback	.019	.082	-.008	-.030	.082	.111	**.629**	.073
ISCL02	T. moves on to next ch.	-.102	.019	-.057	.039	-.011	.300	**.623**	.031
ISGY13	Ch. let others finish speaking	.125	.003	-.121	.036	-.025	.027	.083	**.829**
ISGY12	Ch. notice noise and disorder themselves	.097	.052	.442	.097	-.126	.091	-.035	**-.558**
Portion of variance explained		14.50	14.41	9.64	8.11	6.29	5.49	4.31	3.74
Internal consistency		.92	.90	.87	.87	.68	.68	.54	.46
* was not used for constructing scale to be equal to session analysis									

Legend

Factor 1	Co-determination of children	Factor 5	Feedback in the group
Factor 2	Instruction by trainer	Factor 6	Time pressure (trainer)
Factor 3	Intervention by trainer	Factor 7	Time and pace management (trainer)
Factor 4	Clarity of rules (trainer)	Factor 8	Rules of conversation (children) (dropped because of very low internal consistency)

C 6. Factor loading matrix for items of paragraph analysis on the level of children behaviour (N = 1887)

Item		Factor		
Variable	Label	1	2	3
IKC08	fools about	**.775**	.088	.080
IKC06	fidgets	**.648**	.019	.245
IKC04 *	talks about other things	**.583**	.045	-.165
IKC03	plays with things	**.566**	-.044	.409
IKC07	stands up without permission	**.521**	-.194	-.051
IKC15	listens attentively to others	**-.507**	.323	-.347
IKC05	participates actively	-.054	**.890**	.000
IKC01	makes own contributions	.018	**.861**	-.007
IKC02	looks around	-.006	.026	**.913**
Portion of variance explained		20.43	17.32	15.29
Internal consistency		.69	.74	
* recoded to measure attention				

Legend

Factor 1	Attention
Factor 2	Participation
Factor 3	Looking around (no further analyses because of single-item).

C 7. Factor loading matrix for items of paragraph analysis on the level of children (performance feedback) (N = 1082)

Item		Factor	
Variable	Label	1	2
IKU20	praised for performance	**-.837**	.067
IKU23	criticized for performance	**.830**	.109
IKU24	criticized for deterioration in performance	-.063	**.787**
IKU21	praised for improvement in performance	.103	**.775**
Portion of variance explained		35.12	30.92
Internal consistency		.57	.19

Legend

Factor 1	Performance feedback to child
Factor 2	Feedback on change in performance (dropped from further analyses because of very low internal consistency).

C 8. Factor loading matrix for Fulfillment of Roles in Reciprocal Teaching Conditions for Paragraph Analysis (N=495)

Item		Factor		
Variable	Label	1	2	3
ZSKJ04	"Student" is given feedback on content.	**.955**	.145	.048
ZSKJ06	"Student" evaluates strategy use.	**.952**	.141	.100
ZSKJ07	"Teacher" calls for improvement.	.196	**.798**	-.178
ZSKJ01	"Teacher" suggests ways for the student to improve his/her answer.	.405	**.700**	-.020
ZSKJ03	"Teacher" helps the student.	-.052	**.657**	.209
ZSKJ02	"Teacher" decides which other child is to apply a strategy.	-.001	-.217	**.833**
ZSKJ05	"Teachers" speak loudly and clearly.	.150	.312	**.711**
Portion of variance explained		29.22	54.11	72.51
Internal consistency		.95	.62	--

Legend

Factor 1	Fulfilment of central tasks
Factor 2	"Teacher" provides help and guidance to improve
Factor 3	Internal consistency too low – no scale constructed

C 9. Factor loading matrix for Fulfillment of Roles in Monitor Conditions for Paragraph Analysis (N=495)

Item		Factor	
Variable	Label	1	2
ZSKM03	"Students" are given feedback on content.	**.826**	-.073
ZSKM05	"Monitor" evaluates strategy use.	**.748**	-.244
ZSKM01 *	"Monitors" suggest ways for the other child to improve his/her answer.	**.521**	.333
ZSKM04 *	"Monitors" speak loudly and clearly.	**.429**	.078
ZSKM06	"Monitor" gives the "student" guidance to improve.	-.014	**.821**
ZSKM02	"Monitor" helps when the "student" gets stuck.	.003	**.805**
Portion of variance explained		28.28	53.34
Internal consistency		.60	.58
* Items dropped to improve internal consistency of scales			

Legend

Factor 1	Fulfilment of central tasks
Factor 2	The "monitor" provides help and guidance to improve

C 10. Factor loading matrix for items of Post-Training-Questionnaire on the level of children (N=64)

Item		Factor		
Variable	Label	1	2	3
pt_a _09	Did you find the training program boring?	**-.750**	.031	.256
pt_a _10	Did you find the training program useful?	**.723**	.321	.095
pt_a _02	How much did you enjoy the training program?	**.721**	.109	-.070
pt_a _04	How much effort did you put into the training program?	**.657**	-.145	-.190
pt_a _03	How much did you enjoy working in a group with other children?	.088	**.828**	.061
pt_a _05	How good was it for you to study together with other children?	.008	**.794**	-.142
pt_a _08	How demanding did you find the training program?	-.050	.004	**.919**
pt_a _01	Do you think that what you have learned will help you in school?	.482	.156	**-.635**
Portion of variance explained		33.97	17.38	12.91
Internal consistency		.71	.56	-

Legend

Factor 1	Enjoyment of training program
Factor 2	Enjoyment of group work
Factor 3	Because the items do not match in terms of content, they were treated individually