Reciprocal Teaching: Investigation of its effectiveness as a method of whole class reading comprehension instruction at Key Stage Two

Frances Elsie Hampson-Jones

Supervisor: Professor J. Masterson
Institute of Education, University of London

Thesis submitted for the degree of Doctor of Philosophy

2014

Abstract

Reciprocal Teaching (RT) is a metacognitive training programme that was found to improve reading comprehension during the 1980s (Palincsar and Brown, 1984). Four strategies: predicting, clarifying, questioning and summarising are taught, then students gradually assume control of teaching within a heterogeneous small group until they are actively involved in constructing meaning from text. A review of the literature revealed that there is "very little" research on RT in the UK (Brooks, 2013), little research worldwide in whole class settings, and little evidence that RT is effective for children under twelve (Cain, 2010). Three intervention studies presented here investigated the effectiveness of RT in whole class UK settings with young readers, and the effect of incorporating visualisation as an additional strategy (RTV). Study 1 was delivered by the researcher with three Year 5 classes (N=50). Results revealed a significant improvement in comprehension scores for the RT groups over a normal instruction group after ten hours of training, but no difference between the RT and RTV groups. A second mixed-methods study (Study 2) involved Year 3 children (N=12) with ten weeks of RT instruction, followed by ten weeks of RTV, delivered by the class teacher. Think-aloud protocols and interviews revealed an increase in strategy use. There was a significant improvement in reading comprehension scores immediately after the intervention, and a one year follow-up assessment showed improvement close to significance (p=.09). The third study (Study 3) in a different school with Year 3 children (N=28), replicated the results from Study 2, but with a significant increase in reading comprehension scores at the one-year follow-up assessment. Overall, the results revealed that RT was effective in three different whole class settings, with children as young as seven. Qualitative measures indicated that the instruction worked by increasing strategy use.

2

I hereby declare that, except where explicit attribution is made, the work presented in this thesis is entirely my own.

....

Word count (exclusive of appendices and reference section): 79,703

Acknowledgements

I would like to thank my supervisor, Jackie Masterson, for her help and guidance, and for always giving me the feeling that this was achievable.

I would like to thank all the children and teachers who participated in these studies, and who gave so willingly of their time and enthusiasm.

I would also like to thank my family for their support and patience.

Finally, a huge thank you to my father for all the encouragement he always gave me. I am sorry he did not see the finished product, but I know he would have read every word

Contents

8

Figures

Tables

15

Appendices

Chapter 1: Introduction and rationale

1.1. Introduction

As a primary school teacher with experience from Reception (age 5) to Year 6 (age 10-11) as a class teacher, and as a provider of one-to-one instruction for children with learning difficulties, I have always been interested in how children learn to read and what makes some children successful while others struggle. I was particularly interested in those children who could read fluently but who seemed to have little understanding of what they had just read. When I decided to increase my knowledge of how children learn, I began post-graduate studies at the Institute of Education. I was surprised to discover that the lack of reading comprehension instruction reported by Durkin (1978) had hardly changed over the following two decades. In fact, researchers found that children aged nine to eleven spent more time having their reading comprehension assessed than they did being taught how to do it (Pressley, Wharton-McDonald, Mistretta-Hampton, & Echevarria, 1998). Nearly a decade and a half after that, Duke and Carlisle (2011) were still bemoaning the lack of comprehension instruction for the early years in the USA. In the UK the drive to teach phonics arising from the Rose Review (2006)[1] seemed to make the same situation inevitable here. In Brooks' reviews of literacy interventions (2002, 2007) he commented that most aspects of reading improvement in the UK were under-researched, but that comprehension skills were "the most under-researched of all"[2]. However, during the course of my studies I was introduced to an approach to teaching reading comprehension called Reciprocal Teaching which immediately struck me as a possible solution to help the children I taught who seemed to be able to decode, but did not comprehend what they were reading. It made me question what it is to be an effective teacher of reading, and to think more deeply about the possibilities that RT might offer in this quest.

---

[1] The Rose Review was an independent review on the teaching of early reading, which proposed that the Simple View of Reading replace the previous Searchlights model and that synthetic phonics should be the primary means of teaching decoding.

[2] Whilst the present research was being carried out, a further edition of Brooks' survey of effective literacy interventions (Brooks, 2013) has pointed to a proliferation of studies, indicating that, "from the evidence now available, children's comprehension skills can be boosted by suitable teaching" (p.27).

1.2. Thesis structure

This introduction will look at what reading is, briefly outlining the Simple View of Reading framework, before looking more specifically at the comprehension component of reading and the models which have been developed to explain it. There will then be an examination of research into the reasons for reading comprehension failure and what can be done to remedy those failures. One particular teaching method - strategy instruction - will be examined in more detail. Reciprocal Teaching (RT) will be described as one method of multiple strategy instruction. There will follow some discussion of how RT works, whether it has been recommended as an approach, and whether its effectiveness might be improved by the addition of the intervention technique of visualisation.

Chapter 2, the literature review, will examine previous research looking at the effectiveness of RT. It will involve discussion of the changing nature of the criteria used to select studies for review in intervention research. Drawing on the evidence, there will be a discussion of what we still need to know about RT, and whether it can be used in whole class settings. The review will be followed by details of, and the rationale for, the present studies.

Chapter 3 presents the first study of the present research, a comparison of three Year 5 classes, one of which received normal comprehension instruction, one which received RT, and one which received RT with the inclusion of an additional strategy – visualisation (RTV). This study was undertaken to test the viability of RT instruction in a whole class situation, to compare its effectiveness against normal instruction, and to compare the effectiveness of RT versus RTV.

Following an examination of the results, a new research question, about the changes which take place in strategy use during RT instruction, was addressed with novel research methods in a second study; this will be discussed in chapter 4. In addition to pre- and post-intervention testing using standardised reading tests, in this study concurrent methods were used. This allowed detailed examination of change in reading strategies in children in a Year 3 class as they experienced a year of RT. They were first taught the original RT strategies, and then introduced to visualisation. In this study a class teacher, rather than the researcher, delivered the programme. There was also a follow-up assessment a year later to investigate whether any gains in reading comprehension were maintained.

Chapter 5 presents the third and final study, where the evolving RT method was extended to a different teacher in a different school, with a larger class of Year 3 children.

Over a period of a year, RT was implemented and the children's concurrent responses were recorded together with pre- and post-intervention standardised reading test results. Again, there was a follow-up assessment a year later to look for maintenance of any gains in reading comprehension.

Chapter 6 presents a discussion of the results of the three intervention studies. It also covers the contribution of the present research, as well as limitations of the studies, suggested improvements and implications. The introduction will begin by looking at the Simple View of Reading.

1.3. What is reading? The Simple View, and is it too simple?

The Simple View of Reading (SVR), first outlined by Gough and Tunmer (1986), and later employed as a theoretical framework in the Rose Review (2006), maintains that reading is a product of decoding and language (listening) comprehension. The SVR has been criticised for over-simplifying reading (Barrs, Pradl, Hall, & Dombey, 2008) with the model seen as reductionist and omitting many aspects of reading, such as the role of the text, and the prior experience of the reader. Stuart and colleagues (2008) countered that the SVR provides a useful framework for focusing attention on the distinct components of reading, but argue that there is nothing *simple* about either component. Kirby and Savage (2008) concur; they note that how much we understand of a text will depend on a range of factors including, for example, how much we were read to as a child, our vocabulary knowledge and our memory resources. The processes of language comprehension and word recognition may be complex, but that need not preclude us from accepting the tenets of the Simple View. Indeed, it is a model that "most researchers are using to frame their studies" (Kirby & Savage, 2008 p.69).

The Simple View is not the only framework for thinking about reading; for example, Snow (2002) provides a heuristic which emphasises that to extract meaning from a text involves the reader, the text, and the activity and that these all occur within a specific socio-cultural context. Students' experiences are mediated by the socio-cultural context, and their experiences in turn have an impact on that context. Within that context, the reader, the text and the activity are interconnected and vary in importance according to whether we look at what takes place before reading, during reading and after reading. Before reading begins, the reader has various pre-existing characteristics relating to their cognitive capacities (e.g. attention, memory, and inferencing ability) their motivation (e.g. purpose for reading,

interest in the content) and their knowledge (e.g. domain specific knowledge, vocabulary and strategy knowledge). As a text is read, some of these characteristics may change - for example, new items may enter the reader's vocabulary- and after reading has taken place change may occur again, in that same characteristic or in another – for example, new knowledge may have been gained which will improve background knowledge for a similar passage when next read.

In addition to the personal characteristics of the reader, the text itself affects comprehension. Texts can be easy to read or difficult, and this can be related to the reader, the text itself, or the activity. Snow (2002) points out that text can vary in genre, form, discourse structure, sentence difficulty, vocabulary, syntax, content and interest. Efforts need to be made to match the correct text to readers at varying levels, and to ensure that texts are 'considerate'. The activity of reading is varied. We can read for pleasure, for information or because someone has told us to do it, and the purpose and motivation will vary accordingly. Reader, text and activity all need careful consideration when we think about reading comprehension. Much research has concentrated on the reader (Snow, 2002) but it is important to think about all the other inter-related factors when seeking to improve comprehension. Thus the choice of texts and the way those texts are introduced can be just as important as the characteristics of the individual readers concerned.

An earlier theorist, Adams (1990) had acknowledged the importance of decoding, but placed it within an interactive model which also stressed the importance of context, and which put meaning at the heart of the reading comprehension process. Adams (1990)outlined four processors – the orthographic, the phonological, the meaning and the context, and said they all act together continuously, receiving information and providing feedback to each other. Thus, successful reading depends on abilities in these four areas – automatic letter recognition, accurate phonemic processing, strong vocabulary knowledge and the ability to construct meaningful messages during reading.

Snow and Adams both emphasise the interactive nature of the processes involved in reading comprehension, which highlights a possible difficulty with The Simple View, in that it may be taken, mistakenly, to suggest that, in contrast to these other frameworks, decoding and language comprehension are separate components of reading comprehension rather than components which are actively combined as we read. The Simple View is not a model of processing, but rather a way of separating the two essential elements so that each may become a focus for teaching and research. Additionally, as the SVR was adopted by the Rose

Report, as we have seen, and as it subsequently formed a central part of the Primary National Strategy's view of how reading is learned and should be taught, it has been taken as the model which underpins the research in this thesis.

1.4. What is reading comprehension?

Reading comprehension consists of more than just reading the words on the page. Text reading, according to a cognitively based view, is an interactive process rather than just a set of skills (Dole, Duffy, Roehler, & Pearson, 1991). A good definition of reading comprehension might be: "the process of simultaneously extracting and constructing meaning through interaction and involvement with written language" (Snow, 2002 p. 11).

In seeking to understand the comprehension of a text as a whole, mental model theories have been influential. According to these, when we read, or when we listen to someone speaking, we construct a mental model of what we have read or heard. This model is not the text itself, but a representation of it (Johnson-Laird, 1983). Using Johnson-Laird's framework, van Dijk and Kintsch (1983) proposed the term 'situation model' to mean a model which goes beyond the text representation and integrates inference and background knowledge to construct meaning. According to McNamara and Magliano (2009), this first major processing model of comprehension has provided the foundation for most of the subsequent models. Kintsch developed the theory further, and proposed a two-stage Construction-Integration model, whereby as words or phrases are read (or heard) they set off associations in a spreading fan (Kintsch, 1988). These associations are refined by the context until a stable interpretation emerges. Later this Construction-Integration model was modified to become more interactive (Kintsch, 1998). Instead of all associations being activated in a bottom-up fashion, they are activated by the strength of the associations in memory and, as the reader proceeds; the context strengthens those associations which are more appropriate and inhibits those which are less likely. When the reader makes more connections within the text and generates more prior knowledge then a stronger situation model is made, which in turn leads to a stronger representation in long term memory (McNamara & Magliano, 2009).

McNamara and Magliano (2009) reviewed the principal models which build on the Construction-Integration model to a greater or lesser extent, i.e. the Structure-Building, Resonance, Event-indexing, Causal Network, Constructionist and Landscape models. They argue that the Construction-Integration model, the Structure-Building model and the

Landscape Model attempt to describe the basic and overall comprehension processes, whilst the other four models are concerned with the processes that go on beyond the information in a target sentence, e.g. retrieving background knowledge and making inferences. The authors concluded that the models differ largely in terms of their foci and the fact that they describe different comprehension situations. Thus, the Construction-Integration model is concerned more with explaining the comprehension of more challenging expository texts, whilst other models explain the comprehension of easier narrative texts by readers who do not have any difficulties (e.g. the Structure-Building model). McNamara and Magliano go on to argue that the models are not therefore contradictory, but they may be too limited, in that they fail to account for differences in the reader and the text. Consideration is needed of the differences between skilled and less skilled readers and of how comprehension is affected by the reader's goals and the social context (according to Snow, 2002).

Another way of differentiating between the various models of comprehension was that proposed by Graesser (2007), who distinguished between those models which use a bottom-up explanation (e.g. Construction-Integration) and those which are strategy-driven (e.g. the Constructionist model) and therefore see comprehension as more of a top-down process. The Constructionist model (Graesser, Singer, & Trabasso, 1994) is based on reader goals, coherence of text and explanation. Reader goals are one of the aspects of comprehension that McNamara and Magliano considered to be lacking in comprehension models generally and they suggest it may be because much research in the past has been carried out in the laboratory where readers are given a text and told to read it, thus leaving out any element of free will. Coherence refers to the way that the reader seeks to make a situation model which makes sense at the local and global level, and explanation refers to the way skilled comprehenders try to explain what is happening in the text. In contrast to the laboratory based research McNamara and Magliano consider to characterise other models of reading comprehension, the Constructionist model is the one they consider to be the most applicable to the study of comprehension in authentic educational settings, since it seeks to describe " a highly-motivated, strategic reader who routinely engages in goal-directed, effortful processing during comprehension" (p.330). According to McNamara (2007), the strategies a reader uses, their metacognitive awareness and their goals, are aspects which have large effects on comprehension. Furthermore, McNamara and Magliano feel that "the reader is often left out of the equation." (2009, p. 363). The present research aims to try and understand what the reader is doing and how what they do can be affected by strategy instruction; that is, how are readers simultaneously extracting and constructing

25

meaning through interaction and involvement in the text, and can *how* they are doing that be changed by instruction.

1.4.1. Why does comprehension fail?

Having considered what reading comprehension is this section will look at why reading comprehension might fail. If we know the reasons for failure it may be possible to put measures in place to remediate them. According to Oakhill and Cain (2007) reading comprehension failure could be at the level of printed word processing, seen as a failure of lower level skills, or it could be at the level of sentence or text, seen as a failure of higher level skills. Failure at the word level could be due to slow decoding, or to a lack of understanding arising from poor vocabulary knowledge. For reasons of space, difficulties with decoding are not discussed here, but once decoding is automatic and does not interfere with the processing capacity demanded for understanding, then decoding no longer constitutes a constraint on comprehension. If, at this stage, reading comprehension is below levels of spoken language comprehension, then we should be considering ways of improving reading comprehension through instruction, and it this instruction that is of interest in the present research.

Comprehension will therefore fail if a child cannot decode, whilst a focus on decoding at the expense of reading for meaning is a sign of a beginning, or less skilled, reader (Mokhtari & Reichard, 2002). With the emphasis on single words, the novice reader may fail to monitor their understanding of phrases and sentences and so fail to detect errors in their comprehension. Such readers may show this by failing to self-correct and by making mistakes in accuracy, as well as being unable to summarise what they have read or answer questions about the text correctly. Such readers have mastered the basic skills of reading, but not the higher order processes necessary for comprehension. They have learnt to read, but are not reading to learn.

There is much early research showing that younger and poorer readers are more concerned with reading as decoding than as constructing meaning from the text (Clay, 1973; Denny & Weintraub, 1963, 1966; Johns & Ellis, 1976; Myers & Paris, 1978). One of these early studies (Canney & Winograd, 1979) combined interview techniques with the experimental manipulation of text. In the interviews, the majority of responses made to the question "What is reading?" by second and fourth graders (ages 7 to 8 and 9 to 10) were related to pronouncing or understanding individual words, whilst only a small percentage of

the responses related to understanding at the sentence or text level. Poorer comprehenders were also found to be concentrating more on the mechanical aspects of reading than their higher performing peers at both these grade levels and at sixth grade (age 11 to 12). When these children were presented with texts which had been altered semantically, so that they no longer made sense, none of the higher level comprehenders identified them as readable, but the majority of the lower comprehenders at all three grade levels, thought the passages could be read. Garner (1981) used the detection of inconsistencies in text to show that younger children were more likely to use a lexical standard than an internal consistency standard. Baker (1984), using similar methods, again showed that younger readers (in this case aged 5 and 7) identified more problems with lexical consistency than either internal or external consistency. Baker and Zimlin (1989) later showed that children could be taught to use more than one standard - for example, after training, fourth graders could identify what they called "macrostructure" inconsistencies as well as the lower level "microstructure" inconsistencies, which included word level evaluations; but certainly for some children, decoding seemed to remain the dominant goal. Garner felt that it could be connected to an over-reliance on narrow instruction and the use of graded texts, an observation made by Palincsar and Klenk (1992).

In the UK, there has been an emphasis on the teaching of phonics in early reading instruction. The Rose Review (2006) may have stressed the importance of the Simple View of Reading and the need to improve speaking and listening skills, but the message in the Review that made the most headlines was the need to teach phonics systematically. Since the Review was published the Government has sought to enhance the drive for quality phonics instruction; a White Paper (Department for Education, 2010) made a commitment to ensure there is support available to every school for the teaching of systematic synthetic phonics, as the best method for teaching reading, and to provide funding for high-quality training and classroom teaching resources for all schools with Key Stage 1 pupils. In addition, the Government offered matched funding schemes to primary schools to buy resources (Press notice date 06 April 2011) and has shown its faith in phonics by introducing a new phonics-based screening check for six-year olds.

Whilst the ability to decode is a necessary prerequisite for comprehension, it is not sufficient. There is a need to address higher level skills too. Kendeou and colleagues (2014) point out that research has demonstrated the critical roles of inference making, executive function and attention-allocation abilities, and that by understanding the source of

comprehension failure it is possible to develop effective remedial measures (Gersten, Keating, Yovanoff, & Harniss, 2001). However, research into comprehension failure amongst readers with good word reading skills - 'poor comprehenders'- has shown that they are not a homogenous group, and that remediation may not be simple (Cain & Oakhill, 2006; Cain & Oakhill, 2007; Nation, Adams, Bowyer-Crane, & Snowling, 2004). Cain (2010) suggests that poor comprehenders as a whole have difficulty in making a situation model and that although there may be different reasons for this, it is possible to improve comprehension performance by training some of the skills which are impaired and subsequently form part of the reason for the failure to comprehend. Cain points to studies in inference making (Yuill & Oakhill, 1988) and story structure (Williams, Hall, & Lauer, 2005) as examples of such skills training. But she also highlights instruction aimed at making children more active readers, which is aimed less at specific skills, than at tuition in how to read for meaning. Such instruction is termed instruction in general comprehension strategies, and as it is such an important part of RT, strategy instruction will be covered in more detail in the following section, beginning with an outline of the early research.

1.4.2. Comprehension strategies

According to Dole, Duffy, Roehler and Pearson (1991) early research into reading comprehension strategies was linked to the behavioural sciences and based on the view that reading was a skill which could be subdivided into a set of component skills, such as, sequencing events, finding the main idea and drawing conclusions. These skills could be taught until mastery was achieved, and reading comprehension would result (Smith, 1965). However, as research has progressed, reading comprehension has come to be seen as a far more complex activity (as the examination of the mental models theories has shown) with an interactive and constructive nature. According to this cognitively based view of comprehension, strategies are more important than skills.

Graesser (2007) defines a reading comprehension strategy as "a cognitive or behavioural action that is enacted under particular contextual conditions, with the goal of improving some aspect of comprehension" (p. 6). Early strategy research in the 1970s and the 1980s investigated the effectiveness of a single reading comprehension strategy – for example summarising or generating questions (for reviews see Pearson & Fielding, 1991; Rosenshine, Meister, & Chapman, 1996). However, since reading is a complex process, it is not surprising that the strategies involved in achieving understanding are complex too. In a

review of think-aloud studies carried out with expert readers, Pressley and Afflerbach (1995) summarised their findings as follows:

> Skilled readers know and use many different procedures (strategies) in coming to terms with text: They proceed generally from front to back of documents when reading. Good readers are selectively attentive. They sometimes make notes. They predict, paraphrase, and back up when confused. They try to make inferences to fill in the gaps in text and in their understanding of what they have read. Good readers intentionally attempt to integrate across the text. They do not settle for literal meanings but rather interpret what they have read, sometimes constructing images, other times identifying categories of information in text, and on still other occasions engaging in arguments with themselves about what a reading might mean. After making their way through text, they have a variety of ways of firming up their understanding and memory of the messages in the text, from explicitly attempting to summarise to self-questioning about the text to rereading and reflecting (p.79).

Following the findings of the research with expert readers and the number of strategies they used when reading, commercial reading schemes in the USA adopted strategies to such an extent that Block and Paris (2008) identified 45 of them which were proposed by publishers between 1978 and 2000. However, when Congress convened a national panel to review the scientific literature and to determine the most effective way to teach children to read, their report (National Institute of Child Health and Human Development, 2000) suggested that only eight of these instructional procedures had a scientific basis. The National Reading Panel seems to use the words 'strategy' and 'procedure' interchangeably. Indeed, in the summary they define comprehension strategies as "specific procedures that guide students to become aware of how well they are comprehending as they attempt to read" (p. 4-5). Pressley and Afflerbach also used the terms as synonyms in the passage above. The Panel identified six procedures or strategies which could be taught, and two procedures, or strategies, which related to the methods used to teach them. Thus, the six effective strategies which could be taught were: the use of graphic and semantic organisers; question answering; question generation; story structure; summarisation; and comprehension monitoring. The effective teaching strategies were cooperative learning and multiple strategy teaching. Cooperative teaching will be discussed in a separate section, whilst multiple strategy instruction has seen further research that has confirmed that the most successful comprehension training involves combining a number of strategies and teaching

them alongside metacognitive awareness (Cummings, Stewart, & Block, 2005; Palincsar, 2006).

Block and Pressley (2007) have added the construction of mental images, instruction in inferencing skills and activating prior knowledge to the list of procedures having a sufficiently strong scientific basis to be included since the publication of the original report. Thus, there is recognition that comprehension instruction can improve comprehension, and there is evidence to show that there are effective methods of doing so. One such method of improving reading comprehension through a multiple strategy approach is RT, which is described in the following section.

1.5. Reciprocal Teaching - strategy teaching combined with metacognitive awareness

Palincsar and Brown developed RT as a "multicomponent, metacognitive training package" (1985) which aimed to combine explicit instruction about comprehension with metacognitive awareness. It was initially introduced in the United Stated in the 1980s, in an effort to improve reading comprehension in poor comprehenders in junior high schools (age 12 to 14). The history of the research into its implementation will be covered in the literature review in chapter 2, but the recommendations for its use in everyday classroom instruction in reading comprehension are of interest here. In the USA, The National Research Council (2003) identified RT as one of several research-based practices suitable for broad-scale implementation. The Reading Next report (Biancarosa & Snow, 2006) included direct teaching of comprehension strategies as well as text-based collaborative learning amongst its list of the fifteen elements needed to improve middle and high school literacy. More recently, in the USA, RT has been included in a guide to improving reading comprehension for younger children, in an Institute of Education Sciences Practice Guide (Shanahan et al., 2010). The guide, published by What Works Clearinghouse, offers educators evidence-based recommendations. The guide makes five recommendations; but only the first of these had strong evidence – that is, to teach students how to use reading comprehension strategies. Within the recommendation to teach comprehension strategies, RT is given as one of four examples of multi-strategy formats. In recent years, the Common Core State Standards (National Governors Association Center for Best Practices: Council of Chief State School Officers, 2010) initiative in the USA has moved away from making strategy teaching

30

mandatory, in its drive to mandate standards rather than methodology. However, the importance of strategy teaching has not gone unnoticed, and the use of strategies is still mentioned as one of the tools with which teachers need to equip their students (p.4).

In the UK, in 2005, RT was recommended as a means to improve reading comprehension in the new Primary National Strategy Guidance for Teachers in a series of leaflets entitled *Understanding Reading Comprehension* (Department for Education and Skills, 2005a, 2005b, 2005c) which were aimed at primary head teachers, literacy coordinators and teachers at Key Stage One and Two. The recommendation was reiterated in the revised Primary Framework in 2006, with reference to RT as a 'classic method' for teaching reading comprehension strategies (Department for Children Schools and Families, 2006). In this guidance, the role of the teacher is seen as crucial in explicitly encouraging the use of comprehension strategies. Further, the guidance continues, comprehension improves when teachers provide explicit instruction in comprehension strategies.

Despite the recommendations, there is little evidence of RT being used in UK schools. One example is provided in an evaluation of the EveryChildaReader programme (Tanner et al., 2011). EveryChildaReader (ECaR) was developed by a collaboration between the KPMG Charitable Trust and the Institute of Education and Government. It was rolled out nationally in 2008, with the intention that by the academic year 2010-2011, 30,000 pupils a year would access reading support through ECaR. Although ECaR is primarily concerned with delivering Reading Recovery (a short-term one-to-one literacy intervention for primary school children who are struggling with reading and writing after their first year of school) it also supports a number of literacy interventions, such as Early Literacy Support, Talking Partners and Better Reading Partners. Another one of these supported interventions is RT. However, according to Tanner and colleagues, these interventions were delivered in a more fragmented way than Reading Recovery. As for how many schools implemented RT, it is impossible to tell as it is included in the 'other' category during the interviews in the report. Thirty-one per cent of ECaR schools implemented an intervention in this category; therefore some schools may be implementing RT in line with the recommendations, but the vast majority are not. Anecdotal evidence from teachers with an interest in early literacy, studying at the Institute of Education, suggests it is rarely implemented. Similarly, a Google alert for mentions of RT on the internet over the past eighteen months has produced nearly daily bulletins, but without exception they have been links to sites based in the USA. Possible explanations may be that much of the research was carried out across the Atlantic, and is

not seen as relevant in the UK, or that much of the research was carried out with small groups of children. Research-based interventions will only be adopted if it is practicable to do so in the classroom setting (Pressley, Graham, & Harris, 2006), and small group work has implications for staffing which makes such interventions difficult to implement.

Having looked at how RT has been recommended as an approach, both in the USA and the UK, the following section will look in more detail at RT, describing which strategies are taught and the methods by which the instruction proceeds. It will be followed by a discussion of why RT might be effective.

1.5.1. Reciprocal Teaching – the content

RT was developed to consist of content and method. The content involves four reading comprehension strategies – predicting, clarifying, questioning and summarising - while the method consists of a dialogue, initially between the teacher and pupils, and then (through giving pupils the role of teacher in turn) between pupils, so that they practise comprehension strategies through discussion. The four strategies as described by Palincsar, David and Brown (1989) are given below.

Predicting

Predicting requires the reader to hypothesise about what the author might talk about next in the text. It requires the reader to recall what they already know about a topic, and as they read on they have their hypothesis confirmed or disproved, encouraging them to link new knowledge with their existing background knowledge. It takes place not only at the beginning of a text, but throughout. Predicting also helps the reader to learn about the structure of written texts.

Clarifying

Clarifying means attending to the many reasons why text may be difficult to understand, and may be of particular importance to those readers accustomed to believing that the purpose of reading is merely to say the words correctly, not to make sense of them. Clarifying entails asking themselves and one another for help when new vocabulary, awkward structure, unclear referent words and unfamiliar or difficult concepts occur. Readers are taught to be

alert to such stumbling blocks and to reread, read ahead, ask for help, discuss or take any other steps needed to restore meaning.

Questioning

Readers can be taught to create questions about the text at many levels, from questions that can be immediately answered by using the text just read, to questions that require an inference to be made or to apply information from the text to new problems or situations, which may not have an immediate answer.

Summarising

Summarising is identifying, paraphrasing and integrating important information in the text. Readers are instructed to ask, "What is the most important information? What is this text mainly about?" Most importantly, they are then instructed to put the answers to these questions into their own words, assuring themselves that they understand what is happening in the text.

In a review of self-directed learning Biemiller and Meichenbaum (1992) claimed that ten years of research had shown that one source of difference between the highest and lowest achieving students is the degree to which they become self-regulators of their own learning. RT involves teaching the four strategies, but it also aims to enable students to gain the awareness of their own thinking processes so that failures in comprehension are recognised and corrected by applying the strategies used by expert readers.

1.5.2. Reciprocal Teaching – the method

RT, as noted above, was intended to combine content and method. The content, namely the four strategies of predicting, clarifying, questioning and summarising, has been covered in the previous section. We will now look at the methods central to its implementation; modelling and cooperative learning. Then we will look at some possible explanations as to why RT is effective; that is, through increasing motivation, and through the models of participation and self-regulation.

Although there has been a variation in whether the four strategies are made explicit before group discussions follow (what Rosenshine and Meister 1994 called "ET-RT") or whether group dialogues proceed first and the strategies are modelled during group discussions (what Rosenshine and Meister called "RTO") it is clear from a training manual

33

produced by Palincsar, David and Brown (1989) that although the original studies had used RTO their preferred method became ET-RT; that is, explicit teaching of the strategies followed by discussion in groups. Both ET-RT and RTO have a particular method however, regardless of whether the strategies are introduced first or as the intervention proceeds. This method is that of a dialogue between teachers and children, with each person acting in response to the other. In an appendix to the manual, the authors make the basis of the teaching explicit, saying: "Underlying the model of RT is the notion that expert-led social interactions play an important part in learning and can provide a major impetus to cognitive development" (p.44). The RT intervention programme is defined as ''an example of socially mediated instruction in which the teacher and students engage in dialogue for the purpose of constructing meaning from text'' (ibid). In the initial stages the teacher is the class teacher, but later this role is taken by all the children in turn. RT is based on the writings of Vygotsky, with social interaction seen as playing a vital part in developing cognition (Vygotsky, 1978). Using the zone of proximal development, the teacher guides children to understand concepts they are unable to understand on their own, but which are not beyond their understanding when they are helped by a 'more capable other'. According to the manual, the teacher has an important role at the beginning in modelling strategy use and scaffolding the children's learning by providing the structure of each session and encouraging responses and giving corrective feedback. As the children become more confident with RT this support is increasingly withdrawn, until the leadership of the group has been taken over completely by the children.

The idea of scaffolding instruction is from the writings of Bruner (Bruner, 1978; Wood, Bruner, & Ross, 1976) and is based on a scaffold as metaphor, in that it represents a temporary structure which gives support, but which can be adapted and adjusted as the child's needs change. In a review of research on scaffolding (Van de Pol, Volman, & Beishuizen, 2010) the authors acknowledge that although there have been decades of research on the subject no consensus exists with regard to its definition. They cite Puntambekar and Hübscher's contention that "the scaffolding construct is increasingly being used synonymously with support" (2005 p.1), but maintain that following Stone (1998a; 1998b) scaffolding is an interactive process in which the teacher and the student must both participate. Van de Pol and colleagues distinguish three key features- contingency, fading and transfer of responsibility. Contingency implies that the teacher adapts support in one way or another according to the pupil's level of learning; fading occurs as the teacher decreases the level of support and the transfer of responsibility occurs as the pupil

34

increasingly takes responsibility for their own learning. Scaffolding in RT can be implemented by the class teacher, by the child acting as the teacher, or by other children in the group. The role of collaborative learning is therefore an important part of the RT method and one which will be considered in more detail in the following section.

1.5.2.1. The role of collaborative learning

If pupils and teachers are to work *together* to construct meaning from text then a move away from traditional teaching with its initiation-responds-feedback exchanges may be needed. In traditional teaching, exchanges between pupils are not encouraged (Rojas-Drummond & Mercer, 2003), and pupils work as individuals. In collaborative settings pupils are encouraged to work together to construct meaning, to help each other as they proceed and to arrive at an understanding of what they are seeking to achieve as a group (Johnson & Johnson, 1990). Effective teachers of literacy have been shown to use more small group work than whole class instruction (Taylor, Peterson, & Pearson, 2002). Collaborative learning requires more than just putting pupils into groups, however. Properly collaborative learning involves a move away from the traditional asymmetry of power between the teacher and the pupil (Garner, 1992). If text is to be constructed jointly then the teacher may no longer have all the answers. This may not always be easy for teachers to accept. In research in Australia (Gillies, 2006) 26 teachers were given the same initial instruction about cooperative learning; from observations of teacher and pupil behaviours however, it was apparent that from the complete data set of 24 teachers, 13 had implemented cooperative learning which fulfilled the criteria they had set out, and 11 had implemented group work on a more ad hoc basis, without the necessary key elements outlined by Johnson and Johnson. Gillies found that the teachers who had been using collaborative learning in a structured way engaged in more mediated-learning behaviours and used fewer disciplinary remarks, than those teachers in the group work only condition. These teachers did not differ in the beliefs they held about how pupils learn (according to a questionnaire completed beforehand), nor did they differ in the training they had received, but they implemented that training in different ways through the way they interacted with their pupils. Although Gillies does not explain why some interactions were more successful than others, one possible explanation is that some teachers find it hard to accept that children may have an important part to play. From transcripts of RT dialogues (Palincsar & Klenk, 1992) Garner (1992) noted that the children get to do a lot of the talking and share their own experiences and insights; there is an opportunity for "true conversation" to take place - "information flow in the reciprocal

teaching transcripts was clearly *not* unidirectional from texts to readers or from teachers to readers; much information came from readers" (p.228).

Collaboration may be a very important part of RT. A year-long programme which involved explicit teaching of the strategies without the collaborative context (Antoniou & Souvignier, 2007) failed to achieve a significant improvement in reading comprehension for LD students in the fifth to eighth grades[3] in Germany at the immediate post-intervention assessment. The researchers noted that as peer tutoring had been seen to be effective with LD students (Fuchs, Fuchs, Mathes, & Simmons, 1997) as well as typically developing students (Cohen, Kulik, & Kulik, 1982), then combining the explicit teaching methods with cooperative learning techniques would be an interesting avenue to explore. Collaborative learning methods may also increase motivation (see Slavin, 1996 for a review). The following section examines the importance of motivation.

1.5.2.2. The role of motivation

Motivated readers not only read more (Guthrie et al., 1996; Oldfather & Wigfield, 1996) but they will have more positive attitudes towards reading (McKenna, Kear, & Ellsworth, 1995). Eccles and Wigfield (2002) reviewed the literature on motivation and established that there are developmental differences in the way that different factors affect motivation. Younger children are more likely to think that they will fail at a task based on their general competence. They do not use a task by task approach as older children would, taking account of how difficult that particular task is (Wigfield, Eccles, Schiefele, Roeser, & Davis-Kean, 2006). This would make younger children particularly vulnerable to thinking that if they are bad at reading then they cannot read anything, so they will not be motivated to read at all.

A recent study on the effects on vocabulary of reading less (Cain & Oakhill, 2011) highlights the differences between those children who read a lot and those who do not. Children who read a lot (98th percentile) will read more than 4 million words a year, whilst those who read very little (10th percentile) will only encounter about 50,000 words (Anderson, Wilson, & Fielding, 1988). Given the age of this study, this differential may now be even greater as children spend more time on the computer playing video games than they would have done twenty years ago. This differential would not only affect vocabulary development, as Cain and Oakhill show, but it also shows just how big a gap there is in the

---

[3] In this study, Learning Disabled students were identified as having an IQ above 85 and a reading deficit of scores between two to three grades below expectations.

amount of reading practice and knowledge building (Stanovich, 1991 as quoted by Cain and Oakhill 2011). Children who are poor at reading may also exacerbate this difference by continuing to read books which are less challenging and by reading less – the 'Matthew effect.'[4]

However, Cain and Oakhill also show that evidence for Matthew effects is "elusive" (p.433) and cannot be found in every study or for every measure within a study. Protopapas and colleagues (2011) similarly failed to find the expected Matthew effects in their study of 587 Grade 2 to 4 children, followed over 2 years. However, even if Matthew effects are elusive, the fact remains that children who have below average reading skills find it hard to make the progress of their higher achieving peers. Any intervention which can be shown to improve reading motivation will have an effect on reading, but in particular, if younger children can be motivated to read more they will avoid possible Matthew effects. However, it is not only an increase in the overall amount of reading that is important, since as Guthrie et al. report (2007) there is extensive research showing an association between motivational and cognitive variables in reading comprehension. Thus, strategy use has been shown to be associated with internal motivation (Pintrich, 2000). When students want to understand what they are reading, when they are interested in processing the material deeply and when they enjoy what they are learning, then they are more likely to use strategies which facilitate this, such as summarising, questioning and monitoring. Taboada and Guthrie (2006) in a study of fourth graders (children between the ages of 9 and 10) found that measures of prior knowledge, questioning and internal motivation had unique correlations with reading comprehension, and they concluded that motivation should be made explicit in instruction aimed at improving reading comprehension. These results suggest an important role for motivation in improving reading comprehension; in addition to the question of which strategies to teach, we should be considering how to improve motivation, and therefore engagement with the text. If RT can improve motivation it may help to explain its effectiveness. The next two sections look at two models which may help to explain why RT works: (1) participation and (2) self-regulation.

---

[4] The Matthew effect coined by Stanovich (1986) - after the Gospel according to St. Matthew-describes how in reading, the rich get richer and the poor get poorer. Good readers read more and thus get better at reading, whilst the poorer readers read less and thus fall even further behind.

1.5.2.3. The participation model

In a book about reading development and difficulties Cain (2010) postulates that RT is successful because it encourages children to be active readers and that teaching comprehension skills encourages children to focus on content and become more engaged with "the process of meaning construction" (p.185). These ideas are reiterated in an article which questions the model that strategy teaching results in an internalisation of strategies and replaces it with the idea that participation in strategic reading leads to a view of reading as a strategic process (Davis, 2011). Theories of internalisation of strategies have their roots in the writings of Vygotsky, (1978). What begins as an interpersonal process transforms into an intrapersonal one through the internalisation of socially based and historically developed activities. Thus, strategy instruction originates from the

teacher, from where through modelling and then scaffolded instruction the strategies become part of the communication between the teacher and the pupil, until finally they become part of the pupil's own reading process. Davis (2011) proposes that instead of strategies becoming internalised in this way, a more illuminating model may be that which characterises the development as a process whereby the pupil gradually learns to be a particular kind of reader. Through engagement in reading as a problem-solving activity, pupils begin to understand that reading is about constructing meaning rather than just being about decoding. Pupils begin as what Lave and Wenger call 'legitimate peripheral participants', and by gradually taking a fuller part in the process of reading strategically they become full participants (Lave & Wenger, 1991).

Twenty years ago Garner (1992) noticed that RT brought about a shift in children's responses. In the initial stages they would begin with a statement such as "this is my question", but eventually this would become a spontaneous interjection, and they would even interrupt a teacher's reading of the text. Garner feels that children are no longer using RT strategies as "mental recipes for action" (Rogoff, Gauvain, & Gardner, 1987, p. 303), but they are inventing new strategies out of the old. They are becoming active readers. Garner argued that rather than testing children post-intervention to see if they are summarising, for example, it would be more informative to see if these children were still stopping themselves during reading to make interjections, and to see if they were questioning their understanding.

It may be helpful then, to think about a difference between strategy instruction and becoming a strategic reader, as it appears that the former does not always result in the

latter. Recent research in the USA, on a large scale, has looked at pupils who received strategy instruction, but who did not improve their reading comprehension, i.e. they did not become strategic readers (James-Burdumy et al., 2010; James-Burdumy et al., 2009). The first report from the study was based on the first year of data collected in 2006-2007, for the first cohort of 6,349 fifth-grade students and was released in May 2009. It examined the implementation of four curricula designed to improve reading comprehension – Read for Real, Reading for Knowledge, ReadAbout and Project CRISS. Using random assignment of schools, they found no statistically significant positive impacts of the interventions compared to control schools which did not receive the intervention. Indeed, one programme (Reading for Knowledge) was found to have a negative impact on the composite test score. The second year of the study involved 4,142 new fifth-grade students and revisited 176 of the original 252 schools to test for any delayed impact. The new fifth grade students were taught using three of the original programmes (Reading for Knowledge being excluded). No significant effects were found a year after the interventions had ended, and again, no significant effects were found for the intervention groups over the control in those schools implementing the programmes for a second time. The only support for any of the supplemental programmes was for ReadAbout teachers implementing the curriculum for a second time. It would seem then that teaching strategies is not enough to improve reading comprehension. However, the study does suffer from including data from schools where there was a low level of implementation. From the first year of the study, for example, classroom observations showed that 80% of teachers reported using the curricula they had been assigned, which means that from the outset, 20% of the teachers were not implementing the program they were supposed to be. Add to this the observations of behaviour deemed important, which showed teacher implementation of between 55% and 78%, and there are obviously some problems with fidelity to the programmes. It would appear that these programmes were teaching strategies but were not teaching strategic reading.

In the USA, since the publication of the NRP report (National Institute of Child Health and Human Development, 2000) and the listing of strategies that were found to work and that should be taught, books have been written to encourage teachers to use comprehension strategies with their classes (Harvey & Goudvis, 2007; Keene & Zimmermann, 1997; Miller, 2002). Reading programmes have included strategy use in their teachers' manuals. However, as we have seen, the evidence that these approaches are successful is lacking. In an examination of core reading programmes (Dewitz, Jones, & Leahy,

2009), it was found that strategies were taught in isolation and any connections which could and should be made by the teachers and pupils were left to chance. The instruction was found to be wide, but not in any depth (Marcell, DeCleene, & Juettner, 2010). It is not enough just to know about strategies, students need to know how and when to employ them as well. Paris, Lipson and Wixson (1983) identified three types of knowledge about strategies. Declarative knowledge is knowing what a strategy is, procedural knowledge is knowing how to apply the strategies which are known, and conditional knowledge is knowing when to use a particular strategy.

Further criticisms of strategy instruction have been made by Hirsch (2006), who called it "formalistic" and strongly objected to programmes which "persist, unit after unit, in asking students to 'predict', 'summarise', 'infer, etc. – as if endless use of these strategies will increase students' reading comprehension ability". Fisher and Frey (2008) also feel that strategy instruction in the USA is in danger of being 'curricularized'; that is, the strategies have become more important than what is being read. In a recent article (Palincsar & Schutz, 2011) some of these criticisms of strategy teaching were addressed. Palincsar and Schutz contend that this is separating strategy instruction from its theoretical roots. It is not the strategies themselves which are important, but the way they are used to enable a reader to construct a situation model. Getting children to practise predicting because they have drawn the name of that strategy from a pack (as witnessed by Palincsar and Schutz and the kind of activity that may well arise from the basal programs Hirsch was criticising) will not encourage engagement with the text as an active process unless that text has something in it to enable a prediction to be made by connecting with prior knowledge, or by the content or structure.

Pearson (2011) was also critical of those opposed to strategy instruction, saying that often these criticisms are levelled at "some hypothetical caricature of strategy instruction" (p. 250). He did acknowledge however, that strategy instruction in the USA has been put into practice by the use of basal readers and kits which "may breed an excessive reliance on abstract, content-free, metacognitive introspection" (p. 251) and encourage strategy instruction as a goal in itself rather than as facilitators to knowledge acquisition and insight. The participation model is helpful therefore, as it emphasises the role that the pupil has in the intervention and ensures that they practise the strategies in a meaningful way. Firstly they see the teacher model the strategies, then they practise them themselves and then they evaluate how others use them. They participate in being an active, strategic reader. A

further explanation of the effectiveness of RT may be that it promotes self-regulated learning which will be discussed next.

1.5.2.4. Self-regulated learning

Self-regulated learning is the process whereby learners personally activate and sustain cognitions, affects, and behaviours that are systematically oriented toward the attainment of learning goals (Zimmerman & Schunk, 2011). According to Zimmerman (2002) these processes occur at three different stages – the forethought, performance and self-reflection phases. In the forethought phase, self-regulated learners set personal goals and engage in the planning necessary to meet those goals. In the performance phase, self-regulated learners monitor their progress in achieving these goals, and control their feelings, actions and thoughts so that learning can take place, whilst in the self-reflection phase they compare what they have achieved with what they set out to achieve so that they will perform better in the future. Proposing a sequence of developmental levels, Zimmerman and colleagues (Schunk, 2001; Schunk & Zimmerman, 1997; Zimmerman, 2000; Zimmerman & Kitsantas, 1997) suggest that self-regulated learning begins with observation, continues with emulation and culminates in internalisation and enactment. It has been noted that these self-regulated procedures fit in very well with the RT strategies and the way they are taught (Schünemann, Spörer, & Brunstein, 2013). RT, which as we have seen, has an emphasis on modelling and scaffolding, can be seen to be following the same model of instruction. Indeed, even in the early research, Palincsar and Brown (1984) made it clear that the strategies were not an end in themselves, but were an aid to what they identified as the mental processes inherent in text comprehension.

We have seen how RT works and why it might be effective, but there may be ways to increase its effectiveness, and one possible addition is a fifth strategy. As we saw in the section on recommendations, there is evidence to support mental imagery as a strategy to improve reading comprehension, so the possibility of its inclusion is considered next.

1.5.3. Why might visualisation help?

Pressley (1976), in one of the very first experimental studies of a single strategy instructional programme, found that eight year olds instructed in imagery use could correctly answer significantly more questions about a short story than a control group instructed to "do whatever you can or have to" to remember what they read. More than twenty years later, a study with random assignment to groups, found a large and statistically significant positive

effect of visualisation on reading comprehension in children with an average age of 7 years 6 months (Center, Freeman, Robertson, & Outhred, 1999). Brooks (2007) reports that although the British studies he evaluates have used very small numbers, they did report significant beneficial results of the use of visualisation.

Imagery and the use of pictorial organisers have been shown to be beneficial, particularly in the case of poor comprehenders. Oakhill and Patel (1991) compared the effect of imagery training on reading comprehension with two groups of 11 good and poor comprehenders at aged 9 and 10. The poor comprehenders improved their reading comprehension scores relative to the controls, whilst there was no change for the good comprehenders. Visualisation has also been added successfully to other multi-strategy interventions and has been found to improve their effectiveness. A study of Transactional Strategy Instruction with second grade children found that adding visualisation resulted in a positive and statistically significant improvement in reading comprehension (Brown, Pressley, Van Meter, & Schuder, 1996). Johnson-Glenberg (2000) compared the effects of verbalising versus visualising, in children aged 8 to 11, and found benefits for both, concluding that a programme which combined the two might be highly effective.

The next section considers some other multiple strategy instructions, and discusses why RT was chosen as the basis for the research reported in thesis, in preference to the alternatives.

1.5.4. Alternatives to Reciprocal Teaching

There are alternative multi-strategy reading instruction programmes besides RT, for example Transactional Strategies Instruction (El-Dinary, Pressley, Coy-Ogan, & Schuder, 1994) Informed Strategies for Learning (Paris, Cross, & Lipson, 1984) Concept-Oriented Reading Instruction (Guthrie et al., 1998; Guthrie, Van Meter, et al., 1996) Peer Assisted Learning Strategies (Fuchs & Fuchs, 2005; Fuchs, et al., 1997; Fuchs et al., 2001) and Collaborative Strategic Reading (CSR, Klingner, Vaughn, Arguelles, Hughes, & Leftwich, 2004; Klingner, Vaughn, & Schumm, 1998; Vaughn et al., 2000). For reasons of space these alternative programmes cannot be discussed here[5]: they are not the focus of the present research, as they have not been recommended in the guidance to teachers, in the same way that RT has. RT was chosen because it is a strategy that has been recommended in the Primary National

---

[5] For a discussion of the differences between Reciprocal Teaching and the programme which is closest to it - Collaborative Strategic Reading- see Appendix A.

Strategy (DCSF, 2006) and was recommended in the series of leaflets published as guidance as previously mentioned, and because it is an approach which seems achievable by a classroom teacher without extensive resources or training (unlike Transactional Strategies Instruction for example, which requires an initial 4 half days of in-service training and ongoing support (El-Dinary & Schuder, 1993)). Additionally, The Literacy Hour already contains within it a time for Guided Reading. Guided Reading was intended to replace the traditional 'listening to children read' practice which had been shown to be ineffective (Wheldall, Colmar, Wenban-Smith, Morgan, & Quance, 1992) with lessons where the teacher works with a group to "enhance their reading strategies" (Hobsbaum, Gamble, & Reedy, 2002 p.5). RT can easily be used within the time allocated for this. However, despite the use of Guided Reading in the UK and internationally as a way to teach reading comprehension to older children (for an example of its recommended use for Secondary pupils see Department for Children Schools and Families, 2010) there appear to be no experimental studies which have tested its efficacy for fluent readers, and no evidence to show it is more effective than any other approach (Tennent, 2011)[6].

RT was chosen specifically for its ease of implementation by the Highland Park school district in Michigan (Carter, 1997), showing that is considered to be an intervention which can be introduced by class teachers with little previous experience. This is not to say that it is an approach without difficulties; for example, an article by Seymour and Osana (2003) shows the very different thought processes of two teachers as they underwent training in RT. In this study, each teacher had misinterpreted a strategy, with the result that their students were not fully engaged in the RT procedure. It is important that any intervention is examined closely as it proceeds to ensure it is being delivered correctly. Previous studies have been criticised for not providing enough information about fidelity of implementation (Rosenshine & Meister, 1994). The three studies in the present research all contain information about how RT was implemented, and the opinions of the teachers about delivering the intervention will be discussed.

This chapter has introduced RT, and discussed the content and method of instruction. It has been shown to be an approach which is recommended on both sides of the Atlantic and reasons for its use and a possible improvement have been proposed. The literature review section which follows will look at the original RT studies and then at subsequent research which cites Palincsar and Brown and refers to RT as a method of

---

[6] This doctoral thesis did go on to show Guided Reading was effective, but not as effective as RT.

improving reading comprehension. If RT is advocated for use in the UK we need to know more about its effectiveness and whether it can be applied in a whole class situation.

Chapter 2: Literature Review

In chapter 1, RT was described and the features of the content and the method of its implementation were discussed. This chapter looks at the evidence for its effectiveness in terms of the research and commentaries on the intervention which have been carried out since Palincsar and Brown began their studies more than thirty years ago.

2.1. The early studies

The early studies in the USA are covered with reference to two reviews, that of Moore (1988) and Rosenshine and Meister (1994). The first review by Moore provides a useful overview of the initial studies carried out by Palincsar and Brown, but contains little in the way of critical evaluation.

The first study which Moore refers to (Palincsar & Brown, 1984), focused on one-to-one teaching for 24 seventh grade poor comprehenders. The children were selected as being at least two years behind in reading comprehension as measured by a standardised test, but they had decoding skills assessed as sufficient to read grade level texts at 80wpm or more, with two or less errors. Although the children were taught one-to-one, the data from the 24 children provided data for four groups of six children, with one group having received RT, one an alternative intervention (locating information instruction), one doing the daily assessments (to rule out practice effects), and the fourth acting as a control group. The control group received the same pre- and post- assessments as the other three groups, but they remained in their class for normal reading instruction. The RT group made what Moore calls "quite substantial gains" (op. cit p.8) increasing their scores on the daily comprehension assessments from 30% to over 80%. This contrasted with no gains by the practice group and the control group, and minimal gains by the locating information group. In terms of statistical analysis, there were found to be significant effects for group and time, and the interaction between the two was significant, since the difference between the RT group and the other groups increased as the intervention progressed

Moore considered that this initial study raised questions about five different areas, which Palincsar and colleagues addressed in later studies and which are discussed in the sections that follow.

2.1.1. The effects of strategy instruction

A further study (Brown & Palincsar, 1985) tried to determine which was more important in RT, the method or the content, by keeping the content constant (i.e., the same four strategies were taught) but varying the method. Three different groups were used: one group received RT, involving instruction in the strategies, modelling by the teacher and then practising the strategies through an interactive dialogue with one child acting as the teacher; a second group received just modelling by the teacher, but they did not practise the strategies themselves, and the third received explicit instruction in the strategies and subsequently completed worksheets. The RT group was the only group which showed a large and reliable gain. The paper also describes a further study, comparing RT, explicit instruction, a practice group, and a group using scripted dialogues. Again, the RT group showed the most improvement, with the group who practised answering comprehension questions showing no gains: the group receiving explicit instruction in the four strategies and the group who were taught using scripted dialogues made some gains, but these were not as dramatic as the RT group.

Thus, the content alone does not seem to be responsible for the observed gains, since all groups received the same content, viz. learning about the four strategies, but only the way in which they were taught differed. Although Moore does not speculate on the reasons for this, Palincsar and Brown state that the success of RT lies in the gradual transfer of control to the student, in response to the student's changing 'region of sensitivity to instruction' (Wood & Middleton, 1975 cited by Palincsar and Brown, 1985).

2.1.2. Teacher and one-to-one effects

Palincsar and Brown were careful to extend their studies to include teaching by more than the researchers, in order to counter any argument that the improvements could be attributed to teaching by one 'expert'. Moore seems to have confused the initial pilot study, reported by Palincsar and Brown where the instruction was on a one-to-one basis, with Study 1 in the same article, where the instruction was in groups of two. Palincsar and Brown were also careful to extend their research beyond the initial one-to-one teaching (in Palincsar's doctoral research, 1982), to one-to-two in Study 1 in 1984, to groups of 4 to 7 in Study 2 in the same article, and eventually to groups of between 7 and 15 (Palincsar, 1987). Very similar results were reported across all the studies.

46

2.1.3. Reciprocal Teaching and groups of different ability

The majority of the studies of Palincsar and Brown were conducted with groups of children identified as poor comprehenders with reading comprehension scores which were at least two years delayed, and decoding scores of 80 to 100 words per minute with two or fewer errors. However, in a study briefly reported (in Brown & Palincsar, 1985), the students were far more disparate on scores of decoding (ranging between 64 words per minute with 6 words incorrect to 145 words per minute with no errors). We are not given any indication of the children's reading comprehension skills, but the sample would seem to constitute a wider range of readers than in the previous research. The authors state that in this sample of 70 children the instruction 'proved effective'. The same report mentions that interventions were also carried out with a group of gifted third graders and learning disabled college students, but as Moore notes, the researchers do not include details of the results or the design, which make it impossible to judge the effectiveness.

2.1.4. Peers as teachers

To extend the examination of the effectiveness of RT, the researchers investigated using peers as tutors after the children were introduced to the strategies by a teacher. Although the intervention was shorter than previous experiments (three days whole class introduction followed by 12 days peer tutoring) both tutors and tutees showed improvement in their reading comprehension scores, although the tutees were not so successful on transfer tasks of summarising. The researchers attributed this to the lack of time. For purposes of comparison it would have been preferable for all these interventions to have taken place over the same time scale, but it is appreciated that where interventions are taking place in the real world that the length of the intervention may sometimes have to be compromised for practical reasons.

This same study addresses the question Moore raises about whether peers could teach without adult intervention. Palincsar, Brown and Martin (1987) said that although there was no taped evidence to show teachers taking part in the discussions, they were there to monitor the sessions and played an important role in maintaining momentum and direction and resolving conflicts. The question of the relative importance of the various aspects of the teacher's role is one that will be examined more closely later in the literature review section as research in Germany has looked at this in detail (Demmrich, 2005).

Moore continued his review with further questions which he did not feel Palincsar and colleagues had answered as fully, which will be discussed in the sections that follow.

2.1.5. Does Reciprocal Teaching work for all students?

Moore considers that the answer to this question is a qualified "no", citing the example of a student in Study 1 in 1984 (Palincsar & Brown, 1984) who did not show the same dramatic effects as the other five students in the intervention group. This student was considerably delayed in her comprehension skills, being some four years behind on a standardised test, and having a full scale IQ of 67. However, a low IQ was not a barrier to another student in the same study, who increased his standardised reading comprehension score by 20 months over the same time scale. The response to RT therefore, does seem to depend on the individual, and although Moore does not explore this question further, it is one which remains to be answered.

2.1.6. Decoding proficiency

Owing to the nature of the intervention, and a belief in a hierarchy of reading skills, Palincsar and Brown confined their studies to students with adequate decoding abilities. Moore accepts the logicality of this, but does wonder if decoding could be taught in a similar manner. One much more recent study (Pilonieta & Medina, 2009) which uses RT for younger children (First Grade, aged 6 to 7) has used elements of this idea. As the children were less proficient at decoding, because of their age, Pilonieta often used the clarifying element to teach phonics strategies. There is no data as yet to show if this helped with single word reading, and as mentioned earlier, single word reading is not the focus of this thesis, but it remains an interesting avenue for further research, and it is discussed in more detail in chapter 6.

2.1.7. Peer tutoring and misinformation

Moore rightly highlights the problems which could be associated with peer tutoring if incorrect information is taught or inappropriate strategies are modelled. Palincsar and colleagues quote Allan (1976) as reporting that some 30% of errors made by tutees in drill and practice sessions went uncorrected, and they themselves noted some instances of misinformation of facts (e.g., "falconry" was described as a television programme) but as RT is designed to impart processes rather than content, this is not seen to be important. There were no instances in Palincsar and colleagues' study of inappropriate strategies being

modelled. This shows that the tutors were well prepared, but Moore's request for careful monitoring of what tutors teach is justified.

2.1.8. Reciprocal Teaching and other skills

As the researchers themselves have tried to separate content and method, and concluded that the content alone cannot account for the improvements, then it does follow that the method is successful and it would be interesting to see if it can be used in other contexts. Moore does tell us that the technique has been used to teach pre-reading skills to deaf children (Andrews, 1985), and to teach listening skills to first graders (Palincsar, 1986).

It should also be noted that although the experiments in Moore's review took place over twenty-five years ago - the word 'experiment' alone being a clue as to their age - the rigour with which they were conducted and reported means they still have validity today. Moore concluded that RT is a "highly successful way of increasing comprehension scores of students with comprehension deficits" (p.13) and that the effects generalised to other settings, transfer of the skills learned took place, and the gains were maintained. The claims for maintenance are rather over-stated however; for example, in the second study in the 1984 article (Brown, Palincsar, & Armbruster, 1984) the Gates-MacGinitie Reading Test was administered after a period of 3 months, but only six children were involved, whilst in the study of peers as tutors (Palincsar, Brown, & Martin, 1987) the final post test was conducted just one week after the end of the programme.

2.2. Rosenshine and Meister 1994

A second review of research published in 1994 (Rosenshine & Meister) has been the most cited of all the reviews. A search of the ISI Web of Knowledge in June 2014, revealed a total of 187 citations (in contrast with the previous review which did not feature at all) and it is still being cited today (e.g. Compton, Miller, Elleman, & Steacy, 2014). The criteria for inclusion in the review were strict. Studies had to use the term "Reciprocal Teaching", they had to make a reference to the work of Palincsar and Brown (1984), and they had to have random allocation of participants or ensure that the comparison groups were equal on initial measures of reading comprehension. Studies were also excluded if they had taught only one of the strategies rather than the four used by Palincsar and Brown; this left 16 studies, listed in Table 1. In these studies, improvement in reading comprehension with RT was significantly better than in the control groups, with a median effect size of 0.32 on standardised tests, rising to 0.88 when researcher-devised tests were used. These 16 studies

included 12 which used the same 4 strategies, and 4 studies which used 2, 3, or even 10
strategies.

Table 1. Effect sizes in the studies cited by Rosenshine and Meister (1994).

Author/s Year	Country	Grade*	Norm-referenced test	Experimenter devised test	Strategy test
Brady (1990)	United States	ES 7th	.36	.87	
Dermody (1988)		LE 4th	-.32	3.37 (summarising)	
Fischer and Galbert (1989)		LE 3rd/4th/5th			
Jones (1987)		EE 3rd			.34
Labercane and Battle(1987)		LE 5th			
Levin (1989)		LS			
Lonberger (1988)		LE 4th and 6th		1.24	
Lysynchuk et al (1990)		LE 4th ES 7th	.55 .68		
Padron (1985)		ES 3rd to 7th	-.55		
Palincsar (1987)		LE/ES Middle school		.68(est.)summarising 1.08 (est.)	
Palincsar and Brown (1984)		ES 7th		1.0 (est.)	
Rich (1989)		AD		1.74 1.10 (recall)	
Rush and Milburn (1988)		PS			
Shortland-Jones (1986)		EE/LE Grades 1-6	.77	-.02 (summarising)	
Taylor and Frye (1992)		LE 5th and 6th	.07	.85 (summarising)	
Williamson (1989)		EE 3rd	.32		

*Age of participants categorized according to the same standards as a later review
(Galloway, 2003) to allow comparisons: EE - Early elementary Grades 1-3 (age 5 to 9), LE -
Late elementary Grades 4-6 (age 9 to 12), ES - Early secondary Grades 7-9 (age 12 to 15), LS -
Late secondary Grades 10 -12 (age 15 to 18), PS - Post secondary college and junior college
(age 18 to 21), AD - Adult

The authors sometimes found it difficult to tell from the studies which form of RT had been used (RTO or ET-RT). This is not because the distinction is itself difficult, but because some studies did not make this clear in their methodology. The review makes a plea for authors to include more information about the implementation of instructional procedures. However, Rosenshine and Meister used separate analyses to see if there was any difference in the effectiveness between RTO and ET-RT, and concluded that there was little difference between them.

The review is important in that it analysed the findings in so many ways – not only by the type of instruction (RTO or ET-RT) as we have already seen, but also by the outcome measures used (researcher-designed, standardised, or both), the type of student (all students, poor comprehenders only, or generally poor readers with no reference to decoding abilities) and setting (grade level, number of sessions, size of group, person providing the instruction, and type of control group). Other than the difference between the results achieved on standardised tests as opposed to experimenter designed tests reported above, in their summary the reviewers report little difference for any of these factors, despite the number of analyses. Thus, the results remained significant whether the number of strategies was 2,3,4,or 10, whether the students were aged 9 through to adults[7], whether the groups consisted of 2 or 23 participants, whether they had 6 sessions or 25, and whether the instruction was delivered by a researcher or a teacher. The results remained significant and of the same order. There were differences when the control group was considered however, with results on standardised tests being lower when the control group received traditional instruction rather than when they took part in an activity unconnected with reading, for example computer exercises (Brady, 1990). This is to be expected, since in the traditional instruction group some reading comprehension and practice was taking place.

The reviewers also looked critically at the studies and grouped them by quality. This meant that of the 16 studies included in their analysis, one was considered to be "low quality" (Williamson, 1989), whilst two were of "uncertain quality" (Padron, 1985; Rush & Milburn, 1988). The reviewers justified the inclusion of the studies on the grounds that they represented RT in action, but they went on to show that their exclusion did not have any effect on the results overall. Those studies which were considered to be of high quality were the original study by Palincsar and Brown (1984), a further study including work by the same authors three years later (Palincsar, et al., 1987), a study in the intervening years (Shortland-

---

[7] The evidence for children in Grade 3 (aged 8 to 9) was inconclusive.

Jones, 1986), and two later studies (Dermody, 1988; Taylor & Frye, 1992). These studies were considered to be of high quality as they were well designed, contained sufficient information on the study for a judgement to be made, assessed learning of the strategies as well as reading comprehension in general, and assessed the quality of the RT instruction provided. Sixteen years later, these criteria were echoed by the National Reading Panel (National Institute of Child Health and Human Development, 2000) in their methodological overview of teaching reading in general. The concerns remained that studies needed to have sufficient description to ensure fidelity and that interventions be described in sufficient detail to allow replicability. Outcome measures need to be described and characteristics of the participants given. The panel stressed the need for a control group or a multiple-baseline method. Where they differed from Rosenshine and Meister was in their stipulation that all research must have appeared in a refereed journal. This criterion would exclude 75% of the studies in the Rosenshine and Meister review, since nine studies were unpublished doctoral dissertations and three were conference proceedings. Only four of the studies reviewed were published in journals.

A further criticism of the Rosenshine and Meister review might be the restriction to North American sources. Although the initial ideas about RT had been disseminated from the USA, research into its effectiveness had been carried out in other countries. The search procedure clearly states the databases that were used (ERIC and Dissertation Abstracts International, both dominated by North American research) but it is important to look at research carried out elsewhere. Rosenshine and Meister only mention one non-USA article, and that was one they discounted (Gilroy & Moore, 1988), as it compared a group of poor readers in the intervention group with a control group of average readers. This is an unfortunate exclusion, as using RT for these 3 intervention groups (Standard 4, Form 1 and Form 2, equivalent to Years 6, 7 and 8 in the UK) enabled them to achieve comprehension scores consistent with both an average and above average comparison control group of the same ages. After 13-15 days teaching, gains in reading age were seen in 9 out of 10 participants, with half of them achieving gains of over a year on a standardised test. This is in contrast to the majority of Rosenshine and Meister's reviewed studies, which showed small gains on standardised tests.

An additional feature of the review was a list of the criteria necessary for instruction to be classed as RT - as the National Reading Panel has criticised the lack of fidelity measures as we have just seen, the provision of a checklist provides such a measure. Thus, for an

intervention to be called 'Reciprocal Teaching' it must adhere to these guidelines (listed in Table 2). The guidelines were also used as a check for the RT instruction employed in the studies in the present research.

Table 2. Checklist for the quality of the description of Reciprocal Teaching, from Rosenshine and Meister (1994).

1. Students are instructed in a repertoire of strategies (two or more) that they can use to help them understand better what they read.
2. The teacher models each of the activities.
3. Students are invited to make comments regarding the modelling and the passage, such as, "Was there more important information?" or "Does anyone have anything more to add to my prediction?"
4. Students are provided with guided assistance as they participate at whatever level they are capable in carrying out the strategies.
5. The teacher supports each child's participation in the dialogue through specific feedback, praise, prompting, additional modelling, paraphrases, coaching, hints and explanation.
6. The teacher invites students to initiate discussion and to react to other students' statements. Such participation can include (a) suggesting other questions, (b) elaborating upon a summary, (c) commenting on another's prediction, (d) requesting clarification of material they do not understand, (e) offering additional comments on the content, and (f) helping to resolve misunderstandings.
7. During the RT procedures, there is a gradual shift from the teacher doing much of the work to the child taking over the major thinking role. The teacher gradually transfers control of the dialogues to the students and becomes a supportive observer.
8. During the dialogues, instruction is provided on why, where, and when these strategies might be applied.

2.3. Research since 1994

In order to examine more research from outside the USA and to cover research since 1994 when Rosenshine and Meister's review was published, an online search was carried out. Searches were made of ERIC and PsycINFO databases, along with ProQuest. ERIC is the largest database on education in the world. It is produced in the United States, but its coverage is not limited to North America. ERIC covers journal articles, books, conference papers, technical reports, policy papers, research syntheses and other education-related materials. PsycINFO is the American Psychological Association's resource for scholarly articles, book chapters, books and dissertations within the fields of behavioural science and mental health. ProQuest Dissertation and Theses database is the world's database of dissertations and theses and was used to ensure research from outside of North America was included. Only research published in English was included, and the search was restricted to those studies involving school age children and those which were connected with reading comprehension (as, for example, the term 'Reciprocal Teaching' has also been used in Physical Education). As the ERIC search alone revealed nearly 300 sources (after excluding post-secondary education) the databases were searched for any reviews or meta-analyses as starting points. This search revealed one meta-analysis of RT (Galloway, 2003) and one review (What Works Clearinghouse, 2010).

2.3.1. Galloway's (2003) meta-analysis

Meta-analysis is a quantitative way of examining a number of studies and summarising the findings through effect sizes. It has been shown to provide a more robust measure of significance of findings when compared to a traditional literature review (Cooper & Rosenthal, 1980). Galloway gives the guidelines about effect sizes from Cohen (1988). That is, .2 is a small effect, .6 a moderate effect and .8 is considered to be a larger effect. Galloway's meta-analysis was conducted using guidelines for evaluating research which had been published in 2002 (Kratochwill & Stoiber). These guidelines were developed to evaluate all kinds of evidence-based research in the fields of education and psychology. Each research design is evaluated under three headings – general characteristics, key features and supplemental descriptive information. Under general characteristics, studies are investigated for soundness of theory, empiricism and methodology. Under key features there are eight headings, which enable judgements about internal and external validity and connections to real-world situations. The supplemental descriptions allow for evidence that cannot be statistically evaluated, since Galloway points out that previous meta-analytic

research has been criticised for not including qualitative information (Hunter & Schmidt, 1990).

Galloway's search of the PsycLit, ERIC and PsycINFO databases revealed 22 dissertations, 1 book chapter and 24 peer-refereed journal articles. Seven articles were added from the reference lists of the original search items. On closer inspection, Galloway eliminated three studies which she felt did not teach the four strategies of predicting, clarifying, questioning and summarising. The first of these (De Corte, Verschaffel, & Van De Ven, 2001) was not included as the content taught did differ from RT, although the method remained the same. The strategies taught were activating prior knowledge, clarifying difficult words, making a schematic representation of the text, and formulating the main idea. In addition, a metacognitive component was taught – regulating one's own reading process. In excluding this study Galloway has provided an example of where a meta-analysis and a literature review may differ. By applying strict criteria in the interests of statistical analysis, a study had been ignored which could potentially contribute information about how RT could be altered or improved.

The second study which was excluded was that of Klingner et al. (1998) which is a study about Collaborative Strategic Reading. CSR is a method of comprehension instruction which owes much to RT. However, it is largely unknown in the UK, and for that reason the discussion of the research and the comparison with RT can be found in Appendix A.

The third study excluded was that of Tomeson and Aarnouste (1998). This study again uses the method of RT but differs in content. In effect, only one strategy was taught, that of clarifying, and as such Galloway was correct to exclude it from her analysis. A further study (Braun, Rennie, & Labercane, 1986) did not use students as dialogue leaders, a crucial element of RT. Finally, Galloway excluded the study of Marston and Deno (1995) for not assessing reading comprehension directly. On examination of this study it was found that only word reading was measured and not comprehension, so Galloway was right to exclude it. Table 3 shows the studies Galloway listed and the effect sizes she calculated, with the addition of the country in which the study was carried out and the age group involved.

Table 3. Studies cited by Galloway (2003) with reported effect sizes.

Study author/s year	Country	Age	Effect sizes				
			Norm-referenced test	Experimenter devised test	Strategy test	Transfer tests	Follow-up tests
Aarnoutse et al (1997)	Netherlands	LE				-.70	
Aarnoutse et al (1998)	Netherlands	LE				-.06	
Alfassi (1998)	Israel, but data collected in United States	ES	-.26	.81		-.15	.52
Brand-Gruwel et al (1998)	Netherlands	LE	.20		.26	.08	
Dermody and Speaker (1995)	United States	LE	.30	.91	.81		
Hart and Speece (1998)	United States	PS	.42	.61			
Hodge et al (1992)	United States	PS	.62			.12	
Johnson-Glenberg (2000)	United States	LE	.26	.34	.46	-.07	
King and Johnson (1999)	United States	LE	.96				
Lederer (2000)	United States	LE		.49	.82		.89
Lovett et al (1996)	Canada	ES		.51	1.72	.31	
Lysynchuk et al (1990)	Canada	LE ES	.62	.42		.02	
Miller et al (1988)	United States	ES		1.15		.52	
Taylor and Frye (1992)	United States	LE		-.03	.40		
Walraven and Reitsma (1995)	Netherlands	LE	1.49		.71		1.34
Westera and Moore (1995)	New Zealand	ES	.67				1.89

56

Study author/s year	Country	Age	Effect sizes				
			Norm-referenced test	Experimenter devised test	Strategy test	Transfer tests	Follow-up tests
Small N studies							
Kelly et al (1994)	New Zealand	EE		1.03	1.35		1.62
Palincsar and Brown Study 1 (1984)	United States	ES		1.55	.62	.30	2.07
Palincsar and Brown Study 2(1984)	United States	ES		2.3	.94	-.48	2.88
Palincsar et al (1986)	United States	ES			.31		

*Studies in bold were included in Rosenshine and Meister (1994)

In contrast to Rosenshine and Meister (1994), Galloway reports no significant difference in effect sizes between norm-referenced and experimenter designed measures; the overall effect size is .74 for the 667 participants covered by the meta-analysis. However, although the difference between the two may not be significant (F(1, 24) = 2.84, p=.105) there is still a difference, and it is very similar to the Rosenshine and Meister findings, that greater gains are recorded for experimenter devised tests than standardised tests. Galloway reported an average effect size on norm-referenced tests of .52, compared to the .32 reported in Rosenshine and Meister, and on experimenter devised tests, the effect size was .92, compared to the effect size calculated by Rosenshine and Meister size of .88. Some of the difference, Galloway concludes, may be attributable to the small N size of some of the studies she reviewed, which reported significantly greater effect sizes.

Galloway calculated effect sizes for strategy use measures separately, showing that in the nine studies which tested strategy use (not including questionnaires and strategy knowledge interviews) there was a medium overall effect size of .72. However, there was significant heterogeneity of variance suggesting some strategies are better used than others.

One study reported by Galloway that is of specific interest in this thesis is that of Johnson-Glenberg (2000), since here RT is compared to visualising. In an intervention for poor comprehenders between the ages of 8 and 11, there were three conditions, a verbally based RT group, a visually based verbalising/visualising group, and an untreated control

group. The verbalising/visualising (V/V) intervention was based on a programme developed by Bell (1991), in which children are taught strategies to develop their use of mental imagery. In this programme, children are taught to use four categories and twelve terms to facilitate their visualisation. The categories are: what, when, where and how and the respective terms the children are taught to consider are: number, size and shape, and colour for 'what' they can see; background and perspective for 'where' they can see it; the time, for' when' they can see it; and any movement, mood, sound, smell and texture, for the 'how' of what they see. Initially the children practise using the categories to describe a picture in front of them. Then they practise visualising a personal image (e.g., a family pet) followed by visualising a familiar noun (e.g., a boat or a clown). At each step, the emphasis is on providing a detailed description using the terms. From working with a single word children then move on to single sentence visualisation. When visualisation becomes more automatic children are asked about the main idea of a paragraph and their interpretation of what they had read.

Johnson-Glenberg found that both the visual and verbal strategies were successful. The intervention groups made significant pre-test to post-test gains in 11 measures, whilst the control group only made significant gains on one, the Detroit test of Learning Aptitude-Following Instructions (Baker & Leland, 1959). More importantly, Johnson-Glenberg claimed that the intervention groups made significant improvements in their performance on several key measures associated with decoding, reading comprehension and cognitive processing, and the measure most sensitive to gains in reading comprehension – answering open-ended questions – also showed significant improvement. Using Rosenshine and Meister's bench-mark effect size of .32, the intervention groups outperformed the controls on seven measures: prediction, question generation, recall-main ideas, explicit and implicit open-ended questions, listening recall-expository (RT only) and working memory-linguistic processing. When comparing the RT and V/V groups however, Johnson-Glenberg did not find any substantial differences between the two. She explains this by saying that in Paivio's dual code theory (1986), there are two distinct systems for representing and processing information - the verbal and the visual - and that the two groups had strengthened their text comprehension but each through a different route. In conclusion, Johnson-Glenberg suggested that the combination of the two strategies might therefore be "very powerful" (p.781). It is just such a combining of strategies that the present research seeks to examine.

There is one further point of interest in the Johnson-Glenberg article, not mentioned by Galloway or by the original author, outside the table of results. There was a *negative* effect of time. All the students were reading more slowly post-test, with the difference compared to pre-test being significant for both the intervention groups. This may be because when students read more actively they take more time to think about what they are reading and therefore take longer to read a passage. This possibility was explored through the assessment of reading rate in the studies in the present research.

2.3.2. The What Works Clearinghouse review (2010)

The latest review of RT literature was from 2010 (What Works Clearinghouse). The What Works Clearinghouse (WWC) was established in 2002 as an initiative of the Institute for Education Sciences at the U.S. Department of Education, with the aim of being a resource for informed education decision making. The WWC identifies studies that provide credible and reliable evidence of the effectiveness of a given practice, programme, or policy and disseminates summary information and reports on them. This review exemplifies the drive to be more scientific, by adopting the randomised controlled trial as the gold standard, as recommended in the National Reading Panel report (National Institute of Child Health and Human Development, 2000). Unfortunately, this strict criterion has the result of narrowing the literature examined, to the extent that the 2010 review only contains six studies out of one hundred and sixty-four reviewed. Five studies met the strictest criteria, whilst a sixth study met them 'with reservations'. These studies are listed in Table 4.

Table 4. Studies cited in What Works Clearinghouse (2010) with reported effect sizes.

Meets WWC evidence standards							
Study author/s year	Origin	Age	Effect sizes				
			Norm-referenced test	Experimenter devised test	Strategy test	Generalisation/ transfer tests	Follow-up tests
Brady (1990)	United States	LE/ES	0.27	0.90 (average)		0.37 (vocab.)	
Dao (1993)	United States	LE	n/a but sig.*				
Leiker (1995)	United States	LE		0.14			
Lysynchuk et al. (1990)	Canada	LE	0.50		0.12 (retelling) .85(questions)	0.54 (vocab.)	
		ES	0.36		-0.12(retelling) 0.27(questions)	-0.62 (vocab.)	
Martin (1989)	United States	LS/PS	0.26				
Meets WWC evidence standards with reservations							
Westera and Moore (1995)	New Zealand	ES	-0.48				

* The authors report that they were unable to calculate effect size in the same way as for the other studies as no unadjusted standard deviations were included
Studies in bold appeared in Galloway's review (2003). Of these, only one, Lysynchuk et al. (1990), appears in Rosenshine and Meister's review as well. Brady (1990) appears in Rosenshine and Meister, but not in Galloway.

The WWC report classifies RT as an intervention with mixed effects in the domain of reading comprehension, since it did not meet the criteria for potentially positive effects. These criteria, taken from the WWC handbook (2008) are: 1) at least one study showing a statistically significant or substantively important positive effect, and 2) no studies showing a statistically significant or substantively important negative effect. According to the WWC

handbook, a study must meet both criteria to be considered as having potentially positive effects. RT met the first criterion as three of the six studies in the review showed statistically significant or substantively important positive effects, but it failed to meet the second criterion since one study showed substantively important *negative* effects. This study was that of Westera and Moore (1995), the study that met WWC criteria for inclusion 'with reservations'. The reservations were because it was a quasi-experimental study as there was no random allocation of children to groups. However, I would have reservations about this study as it consisted of a very short duration, only 6 to 8 sessions of 30 minutes each. This compares unfavourably with the 15-20 sessions recommended by Palincsar and must be considered to be too short to implement the programme in any detail. The average length of intervention (in the 19 studies which give this information) in Galloway's review is 678 minutes which is very nearly four times the length of the training provided in Westera and Moore's study. Three hours of intervention would likely not result in enough improvement to be detected with a standardised test, which was the only outcome measure. In including this small scale (N=15 students) and very short duration study the authors have caused the intervention to be classified as 'mixed' rather than 'potentially positive' which must have an impact on the teacher who is looking for a way to improve reading comprehension and this report is aimed at practitioners.

The WWC reviews have been criticised by Slavin and colleagues (2008; Slavin, Lake, Chambers, Cheung, & Davis, 2009), for including very short interventions and failing to weight for sample size, which they maintain has led to many of the WWC conclusions being based on atypical effects. This may well be the case with RT, since as we have seen, the Westera and Moore study was indeed a very short intervention.

A further criticism, as intimated in the introduction to this section, concerns only including studies that involved randomised controlled trials. Given that random allocation is not always possible, since doing research in the real world of classrooms means groups are often pre-selected, and that there is a need to conduct intervention research in real world environments (Kratochwill & Stoiber, 2002), then evidence from other studies needs to be considered, even if it does not meet the strictest criteria. Galloway (2003) reported that less than 25% of the studies she included in her meta-analysis used some form of randomisation in selecting participants, yet when the results of those studies that included randomisation and those did not were compared there was no statistically significant difference between them, suggesting that there is no bias in favour of the intervention in studies which did not

use random selection. Additionally, many studies were excluded from the What Works review as they did not include a comparison group, yet, as Brooks (2002) has shown, there is another means of determining whether a treatment group has made better than expected progress – analysis of ratio gain. Ratio gain is a group's average gain in reading age or spelling age in months divided by the time in months between pre- and post-test.

The WWC review also excluded any studies which varied the number of strategies employed from the original four of predicting, clarifying, questioning, and summarising. This follows the need to only include replication studies (as outlined in Kratochwill, 2002), but as Galloway (2003) points out, all studies that were aiming to replicate the original Palincsar and Brown studies by using the same instructional programme and only differed by the addition of strategies, need to be included to investigate whether the programme can be improved.

2.3.2.1. Lysynchuk et al. (1990)

There is one study which met the strictest criteria of the What Works Clearinghouse report and was included in the original Rosenshine and Meister review, as well as being in Galloway's meta-analysis. This study stands out as an exemplar of how to determine the effectiveness of RT, and so it will be studied here in more depth.

Lysynchuk et al. (1990) recruited 36 Grade 4 students (aged 9:02 to 10:10) and 36 Grade 7 students (aged 12:05 to 14:01) from six schools in the case of the Grade 4 students, and two schools in the case of the older students. The participants were poor comprehenders as nominated by their teachers. All could decode at least 80% of the words expected for the grade level, using the Diagnostic Reading Scales (Spache, 1972), whereas all were below the fiftieth percentile on standardised reading comprehension tests. As stated by the authors, the testing was designed to be superior to that carried out in the original Palincsar and Brown study (Palincsar & Brown, 1984), where the pre-testing had been conducted using a group test and the post-tests had been administered individually. After testing for reading comprehension, the participants were paired on the basis of their scores and one child from each pair was randomly assigned to either the RT or the control group. The RT group and control group were treated as alike as possible, except for the strategy instruction, thus they met the same number of times with the same instructor, and they used the same materials and took the same tests. The aim was to try to ensure that any differences observed in improvement in comprehension scores could not be due to a

difference in attention or exposure to training materials between the two groups. During the intervention phase, grade-appropriate expository passages of 300 to 900 words long were used for training and 26 shorter passages (about 200 words long) were used for daily assessments. Assessments were alternated on a daily basis between answering a set of 10 questions and retelling.

The strategy training took place over 13 days, with most groups being of three to four students with the same instructor each time. Unfortunately, it is not clear from the paper, how long these sessions lasted. Each instructor (three in all) took both RT and control groups. The strategy training followed that outlined by Palincsar and Brown (1984) and took the form of four days of explicit instruction followed by expert scaffolding of student leaders who made predictions, formulated questions, requested clarifications and summarised the text. The students took it in turns to be the leader and received support from other students as they talked about the text. The instructors gradually released control and allowed the students to make the transition from external control to internal self-regulation of the strategies. At the end of the session students were given the daily assessment, and given feedback by way of a graph showing their performance.

The authors report significant gains in reading comprehension on a norm-referenced test for the RT groups, both as a whole (t(68)=5.14, p<.001) and by grade (both ps <.01). For the experimenter devised comprehension tests, the RT groups outperformed the controls, and in retelling, the RT students recalled significantly more than the controls. Rosenshine and Meister give an effect size of .55 for fourth grade and .68 for seventh grade, which is equivalent to the .62 effect size Galloway reports (the meta-analysis gives an average where two age groups are studied). The What Works review however, gives an effect size of 0.50 for fourth grade and 0.36 for seventh grade. This is explained in a footnote by the authors who say a correction was applied for multiple comparisons.

2.3.3. Reviews of research in the UK

Although there have not been any reviews of RT in the UK as such, there have been reviews of a more general nature which have contained references to the intervention. Thus in 2002, when Brooks carried out a survey for the Department of Education and Skills of the effectiveness of reading and spelling interventions for children with literacy difficulties (Brooks, 2002), he pointed to the large amount of research on RT in North America, and contrasted it with "scarcely any" (p.37) in the UK. Brooks reported that RT was used in two

London boroughs, Westminster and Haringey. In the former, a study had been published (Greenway, 2002), but it was based on only six children. Data had been supplied to him from 16 children in Haringey, but that research was unpublished. The data did show a substantial gain in reading accuracy however, and an even stronger gain in reading comprehension (ratio gain of 6.4). The Greenway study, although very small, involved an implementation of the intervention within the framework of the Literacy Hour in a Year 6 class. It was focused on a group who had been identified as poor comprehenders. RT was introduced to the whole class on an ad hoc basis, by using the small group to demonstrate a strategy that they had been practising. There was no attempt to use RT for the whole class, nor was there a control group for comparison (owing to a lack of time and the intended quasi-experimental approach). The children's comprehension scores on the NARA (Neale, 1989) improved significantly over the intervention period of one term, and the intervention group also showed a significant increase in self-rated confidence scores.

Greenway argued that it was unclear which of the RT strategies was the most effective (also highlighted in the reviews of Rosenshine and Meister and Galloway), and what effect sharing the teacher's role had (again a common theme). In a later edition of his study (Brooks, 2007), Brooks omitted the Westminster data as involving too few pupils, but included additional unpublished data from Haringey (collected by Christa Rippon). This data, on 88 children, showed a useful gain in reading accuracy, and a substantial gain in reading comprehension. To determine the amount of progress made when a control group is not included in a study's design, Brooks uses the ratio gain measure. As noted above, ratio gain is a group's average gain in reading age or spelling age in months divided by the time in months between pre- and post-test. According to Brooks, a gain of 1.4 or more represents a gain that is more than standard progress, and therefore educationally significant, whilst a ratio gain of 2.0 or more represents not just satisfactory, but useful extra progress. After 10 hours RT instruction, the children in the Haringey study had made 'useful' progress in accuracy and 'substantial' gains in comprehension (Brooks, 2007, p.47) showing ratio gains of 2.4 and 3.7 respectively. The data was included in the fourth edition (Brooks, 2013), when Brooks again stated that despite all the research in the USA, there is still "very little in the UK" (p.48).

However, there is some evidence in the UK that programmes containing certain elements involved in RT are effective. Thus, Yuill and Oakhill (1988) carried out a study on inference training with small groups of poor comprehenders aged 7 to 8 who received

instruction about lexical inference, question-generation and prediction (although only one session was devoted to this latter strategy). They found that the children improved significantly more on a standardised test of reading comprehension than a control group who were given decoding practice. However, the intervention children did not improve significantly more than a group given comprehension exercises. The authors suggested this may have been because the children in the comprehension exercise group often corrected each other and discussed their answers. As a result, the inference group were receiving explicit instruction, but the comprehension exercise group were receiving implicit instruction, with both perhaps being effective. The authors also suggested that further studies were needed to determine which approach would be the most effective. It could be argued that RT involves explicit and implicit instruction, since the strategies are explicitly taught, but their use is also implicit through the use of modelling and discussion.

An extension to the inference training programme introduced by Yuill and Oakhill has been developed for use in Leicestershire schools, and was reported by Brooks (2013). In another small group intervention, carried out with children at Key Stages 2 and 3 (ages 8 to 15 years), the developers have added prior knowledge, word definitions, visualisation and summary skills to the previous inference training, question generation and predicting elements (Whatmuff, n.d.). Whatmuff developed the intervention and Brooks reports on an evaluation by a group lead by Lockley, which shows remarkable gains in both accuracy and comprehension, but there are no published data available. The programme would appear to have much in common with RT, but the lack of publications outlining the programme (other than those available to teachers undergoing the training provided) makes it difficult to compare its implementation.

Additionally, there is a UK doctoral thesis about the importance of inference in reading comprehension which includes two studies that involve RT (Tennent, 2011). In common with much RT research however, it is concerned with small group tuition, rather than whole classes. Three groups of six Year 3 children (mean age 8 years) were given five hours of instruction in either RT or Guided Reading sessions. These sessions took place over ten weeks for one RT group, and over five weeks for the other, the Guided Reading instruction took place over ten weeks. The difference in intensity was designed to compare the same approach delivered over two different time scales. However, there was no significant difference between the RT groups' results and so their data was combined for some of the analyses, and then compared with the Guided Reading group. Using

standardised test (Neale, 1989) significant improvements in reading comprehension and accuracy were observed for both types of instruction, and there was no significant difference between them. Using the gains in reading age given by Tennent, it is possible to calculate ratio gains, of 3.61 for the first RT teaching group, 3.35 for the intensive RT group, and 3.49 for the Guided Reading group. These are impressive gains, but as Tennent points out, the reading assessment used in the study has been criticised for only assessing literal comprehension (Hurry & Doctor, 2007). For this reason, Tennent developed his own measure for testing inference. There were problems with ceiling effects and non-normal distributions for some of the sub-tests, but for the measures which did not have these problems, i.e., overall inference scores and written responses, it was found that the RT approach was significantly better, in both forms (i.e., the intensive and the non-intensive). For total inference scores, there was a main effect of group (F (2,18)=8,48, $p<.001$) with the ten-week RT programme producing higher scores than the Guided Reading group ($p=.019$) and the five-week programme also producing better scores than the Guided Reading group ($p = .001$). For the written responses, there was also a significant main effect of group (F(2.18)=7.53, $p=.006$); again the RT groups outperformed the Guided Reading group - the first group compared to the Guided Reading group was better at writing responses to questions involving inference ($p=.018$) and the intensive RT group compared to the Guided Reading group was also significantly better ($p=.002$).

Tennent concluded that his study showed that Guided Reading is effective (which had not been shown before) but that RT is more effective, and may be particularly effective in supporting productive (i.e., written) responses. His study does show that interventions with this age group can be effective in improving inference making, and that RT can be effective for normal readers as well as the poor comprehenders previously included in RT research. However, Tennent's study remains a small group intervention, and as has been noted, this has attendant difficulties of time and resources. Another small group intervention which was carried out in the UK – the Reading for Meaning Project - will be covered in the next section.

2.3.3.4. The Reading for Meaning Project

One study which includes RT in a UK setting (though in small groups, not in a whole class situation) was published in 2010 (Clarke, Snowling, Truelove, & Hulme). In this randomised controlled trial, three interventions were utilised in an effort to improve text comprehension in poor comprehenders (although, according to the authors, only 84 out of the 160 children

involved met the strict criteria of 1 SD discrepancy between word reading and comprehension, as measured by the TOWRE and NARA respectively). One hundred and sixty children (mean age just below 10 years at the time of screening) were taught for three sessions of 30 minutes a week (in two sessions children were taught in pairs and one involved individual instruction) for 20 weeks. The interventions were designed to be orally based, text based and a combination of the two, as shown in Figure 1.

Figure 1. Programme components for the ReadME project (Clarke, Truelove, Snowling, & Hulme, 2011).
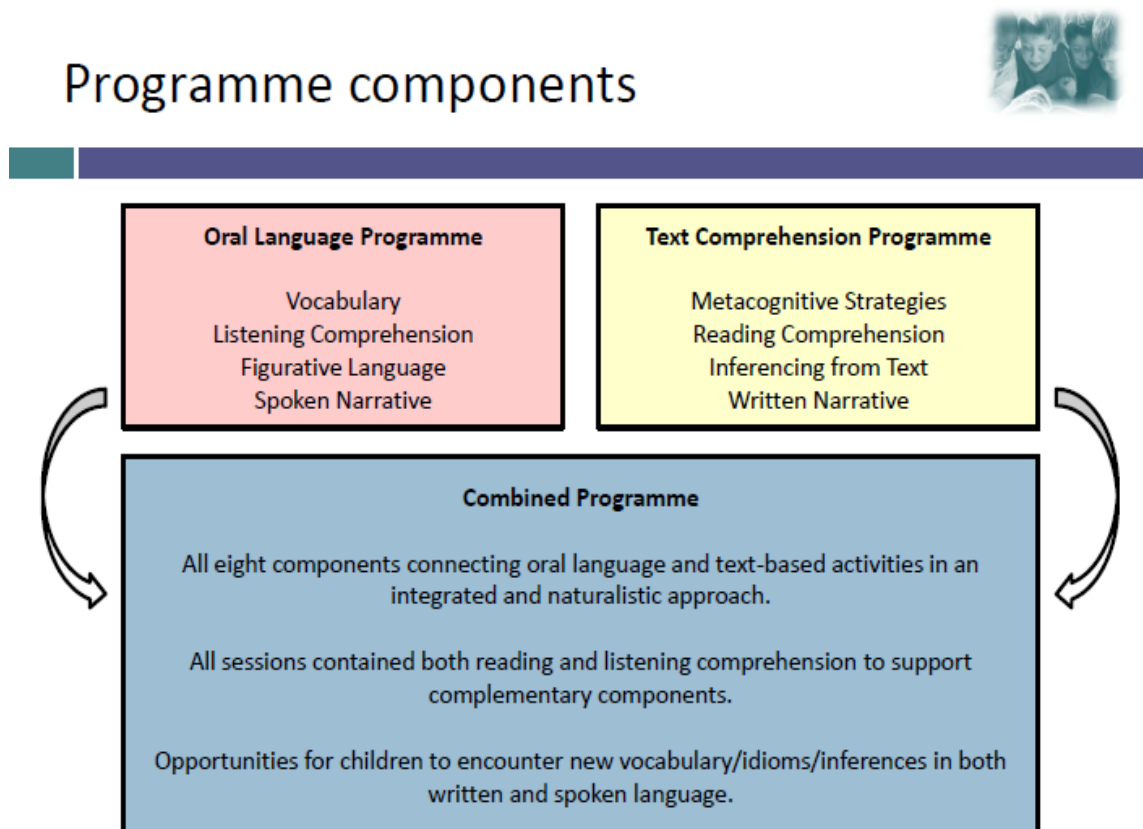
According to the authors, RT was "at the core" of each programme (p.1108) and they referenced both Palincsar and Brown's work (1984) and the Rosenshine and Meister (1994) review. The four strategies of RT - predicting, clarifying, questioning and summarising - were taught, as shown in Figure 2. However, although the figure indicates that the teacher and students took turns in being the teacher, the vital role of peer feedback discussion must have been missing, since children were taught individually for a third of the time and in pairs for the remaining two thirds. Indeed, the authors say RT "centres around discussion

between children and a tutor" but do not mention the important role the programme's originators give to discussion with other children. Additionally, the RT part of the lesson took place during seven minutes according to the lesson plans given, which leaves very little time for the scaffolding, feedback and practice which are an inherent part of the RT process.

Figure 2. The inclusion of Reciprocal Teaching in the ReadMe project (Clarke, et al., 2011)

Reading

RT refers to an instructional activity that takes place in the form of a dialogue between teachers and students regarding segments of text

The teacher and students take turns assuming the role of the teacher in this dialogue

Palinscar & Brown (1984)
Review by Rosenshine & Meister (1994)

Questioner Asks questions about a passage

Clarifier Finds difficult words or ideas and look for clues to help explain them

Predictor Makes guesses about what a passage might be about or what might happen next

Summariser Uses own words to explain key ideas

Assessments were carried out pre-intervention, mid-intervention, post-intervention and at an 11 month follow-up. Results indicated that all the intervention groups made significant improvements in reading relative to an untreated control group. However, it would appear that this intervention is not RT as it would be understood by Palincsar and Brown. In an address to the National Reading Conference in 2006, Palincsar talked about what Haertle (1986) had called "lethal mutations", that is, enacted designs becoming quite different from those envisaged by the originators. She goes on to talk about the danger of teaching the four strategies stripped of the dialogue that she considered so important. The Reading for Meaning project has not redesigned RT as direct instruction, but because of the

limited number of participants and the short time allocated to this part of the intervention, it cannot be considered to be RT as it was intended.

The next section will consider a study which did examine RT as defined by Palincsar and Brown. This research was carried out in Germany, and was particularly concerned with examining the way in which RT involves sharing the teacher's role.

2.3.4. Research in Germany

A doctoral thesis in Germany (Demmrich, 2005) set out to separate the importance of sharing the teacher's role from the other elements involved in RT. As there is so little research outside North America and as Demmrich set out to answer some of the questions raised by the previous reviews, the thesis will be examined in some detail. The research questions Demmrich asked were:

1. Are metacognitive competencies responsible for the improvements in text comprehension seen in RT interventions?
2. Which features of Reciprocal Teaching make it so effective?

Having identified that in RT the child as teacher has some roles which are organisational (such as selecting the strategies to be used and who will apply them, and when to move on in the text) and others which are content-based (such as giving feedback, helping to improve the answer by modelling) Demmrich set out to separate the roles, by investigating different intervention conditions. These involved: 1) RT as in Palincsar and Brown (1984), that is, children took it in turns to fulfil the organisational and content-related tasks, 2) a monitor condition where children had the content-related tasks but not the organisational role, this was taken over by the trainer; thus, the trainer led the discussion, but did appoint a child to give feedback on another's use of a strategy, and 3) a student condition where children had a more traditional role, and the trainer was responsible for organisational and content-related tasks.

The study involved children in the fifth grade (aged 10 to 11) at four different schools. Eighty-six children were recruited to a control group, and fifty-five children made up the intervention groups. These fifty-five children were all volunteers, meaning that the intervention group were self-selected to be below average, since as Demmrich explains, the children with lower verbal ability were the ones must likely to be put forward by their parents and teachers as 'volunteers'. The pre-test differences between the intervention

groups and the control group were controlled for in the analyses. The training took place after the normal school day, and comprised 16 sessions of 45 minutes each, which took place over 4 to 5 weeks. The children worked in small groups of 4-6, and all received three introductory sessions about strategy use and the four strategies. Demmrich gives transcripts of dialogue for all three groups, so we can see how they differed in their operation. This satisfies one of Rosenshine and Meister's earlier criticisms that some studies do not allow you to judge how RT had been carried out.

In terms of analyses, Demmrich carried out three comparisons: firstly, she compared the RT and monitor conditions with the control group on measures of strategy knowledge and performance. Then she compared the three intervention groups, and finally, she compared the student group with the control group. The rationale behind this was Demmrich's belief that RT works by improving metacognition and it is not so much what is taught, but how it is taught that is important. Thus, if there was no difference between the reciprocal and monitor condition, but there was a difference between these two groups and the student group, then it would not just be teaching the strategies that is important, as the student group had also been taught the strategies. Similarly, if there was no difference between the student condition and the control group (who spent their time watching TV, doing sport or homework, or playing with their friends) then Demmrich would have again shown that strategy teaching alone is insufficient.

A major difficulty with the control group in this experiment, as we have seen, and one acknowledged by Demmrich, is that the groups were not equivalent on reading and reading-related measures at pre-test. As the training programme was voluntary it was taken up primarily by children with poor verbal and decoding skills. Demmrich controlled for the differences by using a mean verbal ability score in the analyses. This may account for variation in ability, but it does not account for the variation in motivation, which is always considered to be a problem when any study involves volunteers, rather than random allocation (Cohen, Manion, & Morrison, 2007).

A further problem with the control group was that all children involved were not only from the same four schools, but they spent the normal school day in groups with the participating children, thus, although they were not participating in the training they may have come across the strategies through interaction in other literacy lessons. It is not made clear if the children discussed or used the strategies in other literacy contexts during the school day. Given that these problems are nearly always present in including an untreated

control group in educational research these concerns are not major ones, but it would have been helpful to have had the concerns addressed. However, Demmrich's study is otherwise well designed and provides much information about how RT works to promote metacognition and motivation.

To assess strategy use, Demmrich used several measures. Declarative knowledge about the strategies of summarising and clarifying was assessed by asking the children to write down the features of a good summary and then what they would do if something they were reading was unclear. The clarifying strategies suggested were then divided into external (the use of a dictionary, or asking for help) and internal (e.g. re-reading, reading on, inferring the meaning of a word from the context etc.). Conditional knowledge was assessed using the Index of Reading Awareness Questionnaire (Jacobs & Paris, 1987), whilst what Demmrich calls 'relational knowledge' was measured by asking children to rank the usefulness of different strategies when the task was to comprehend and remember the content of a text (Schneider & Schlagmüller, 2002). Demmrich also used what she called 'performance measures' which were writing a summary and the assessment of text comprehension. For the latter there was an experimenter-designed test - a longer passage (500 words) followed by four "relatively difficult" (Demmrich, 2005 p. 167) open-ended questions. Additionally, a questionnaire was used post-intervention, which assessed motivation and involvement using Likert scales. A large proportion of the lessons were also video-taped and analysed, giving qualitative information about how the children were taught the strategies and how they applied them.

2.3.4.1. Comparing the intervention groups with the control

In terms of strategy knowledge and the difference between the intervention and control groups, there were large highly significant effects of training for summarising and clarifying, but there was no difference between the training groups and the control on the more general measures of metacognitive knowledge, as measured by the questionnaire. However, for relational knowledge, all children, including the control group, knew more about the relative usefulness of strategies at post-test. Although the effect of the interaction between time and training just failed to achieve significance, the intervention groups did show some signs of having caught up with the control group, who had started ahead of them.

For writing a summary, medium to large effects were observed for the training groups. The trained children wrote shorter summaries, with relatively more important

content than the control group. However, when it came to text comprehension there was no difference. The control children started out better at text comprehension and they stayed that way, and even when verbal ability was controlled for, pre-test scores were always predictive of post-test scores, regardless of whether or not a child took part in the programme.

2.3.4.2. Comparing the training groups

Besides comparing the intervention groups with a untrained control group, Demmrich also compared the three intervention groups with each other, in order to explore the hypothesis that the content-related tasks associated with being the 'teacher' would lead to an improvement in metacognitive knowledge and skills. Declarative knowledge about summarising and clarifying showed a significant improvement, with large effect sizes, for the reciprocal and monitor groups over the student group. For relational knowledge, there was another difference in favour of the reciprocal and monitor groups (effect size .80) who improved their knowledge, whilst that of the student group decreased slightly.

When it came to performance however, there were no significant differences between the three groups on any of the measures. Although Demmrich acknowledges that choosing which measures to use is difficult, she does not offer any explanation as to whether the lack of significant results could be attributable to the measures she did use, arguing instead that the measures of general strategy use were not significantly different owing to the small sample size. The measures used need careful consideration beforehand as Demmrich intimates, but they also need careful consideration after selection, and some discussion about how they reflect strategy use is necessary.

Differences were seen however; when it came to the quality of the dialogue, an analysis of the video recordings showed that the children in the reciprocal and monitor groups improved their questioning skills, whereas the children in the student group stayed the same.

2.3.4.3. Comparing the student group with the control group

Demmrich found that the student group performed only marginally better than the control group when the post-intervention results were compared. Indeed, she concludes that they gained very little in comparison to the children who had spent the time on homework, watching TV, doing sports or playing with their friends. It would seem that the

more traditional teaching with an asymmetry of power in favour of the teacher, as has been discussed previously, was not as effective as when children were given more control.

2.3.4.4. The effect on motivation

One justification for separating the teacher's role in the reciprocal and monitor conditions was Demmrich's hypothesis that although the organisational role does not play a part in learning strategies or improving metacognition, it would improve children's motivation and self-efficacy, leading to better performance. The reciprocal and monitor groups both reported significantly greater enjoyment than the student group. It was also hypothesised that those groups which incorporated responsibility for the group (the reciprocal and monitor conditions) would report an improvement in co-operation. When asked to write down what they felt they had learned from the programme, half of the children in the monitor condition wrote that they had learned to cooperate with other children. This was also reported by a quarter of the reciprocal group, whereas no children in the student condition made similar comments.

Demmrich's study did show that the important part of RT, in improving children's metacognition, seems to be the content-based aspect of the teacher's role and that teaching strategy use alone does not seem to improve strategy knowledge and application. It would seem then that the children do not have to be the teacher in terms of organisation and control in order to benefit but they do have to perform the monitoring roles, and evaluate and regulate the performance of other children in their group However, despite improvements in strategy knowledge and improvements on those measures most closely aligned with the training provided, Demmrich found no evidence of transfer having taken place, since the more distal measure of text comprehension showed no significant advantage for the intervention groups. Demmrich suggests that this may be due to the "moderate" number of training sessions (p.165). Yet 16 sessions is comparable with the number in other studies which have reported greater effects of training. The more important observation Demmrich goes on to make is that the reading assessment itself may not have been related to the kind of training which took place. In using longer, expository text with relatively difficult comprehension questions, under timed conditions without any instruction to use the strategies that had been taught, Demmrich suggests that a demonstration of the transfer of skills taught may not have been possible.

2.3.4.5. Measurement of metacognition

Although Demmrich used several measures of metacognition she suggests a weakness in that she only used measures of declarative knowledge, and that procedural metacognition was not assessed. Demmrich concludes that this is difficult, and whilst attempts have been made to use "think-alouds", she feels such techniques tap controlled strategy use rather than automatic processes (Shriffin & Schneider, 1977). However, she used a rather old reference to substantiate this point, and there is more recent research that has gone some way to addressing these concerns about how to measure metacognitive strategy use. In the conclusion Demmrich states:

> 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 … but potentially also to provide useful information on how to help students acquire metacognitive knowledge and develop effective routines to support their cognitive processing. (2005 p.183)

The present research used think-alouds in Studies 2 and 3 to obtain this information. Demmrich's research, in common with many of the studies reviewed so far, used small groups to deliver RT. However, it has already been stated that in order to be used in the classroom, RT needs to be seen as a programme which can be used in a larger setting. The next section will look at those studies which have used an intact class rather than a small group.

2.4. Research on Reciprocal Teaching within a whole class setting

RT continues to be seen largely as a small group intervention. A recent article aimed at promoting active involvement in classrooms (Conderman, Bresnahan, & Hedin, 2011) champions several methods of achieving engagement, all of which are research-based. However, RT is confined to the section on small group instruction and no mention is made of the possibility of using it as a whole class technique. An examination of the studies found in the literature search showed only 18 studies which took place within a whole class situation, even when research before 1996 was included. These studies are shown in Table 5.

Table 5. Reciprocal Teaching studies carried out within whole classes.

Author and date	Country	Sample	Comments
Alfassi (1998)	USA	75 poor comprehenders Aged 14-15	
Carter and Fekete (2001, data from 1993-1995)	USA	Grade 3 Aged 8 to 9	Comparisons with previous years, but no equivalence established
Coley et al. (1993)	USA	Grades 1,4, and 7 Ages 6 to 7, 9 to 10 and 12 to 13	Qualitative study
De Corte et al. (2001)	Belgium	4 Grade 5 classes Aged 10 to 11	Instruction in Flemish
Hacker and Tenent (2002)	USA	17 teachers in 2 schools ages 8 to 12	Qualitative study of implementation
Halberstam (2008)	USA	Grade 3 Aged 8 to 9	Grouped by ability
Hashey and Connors (2003)	USA	Grades 3 to 8 Aged 8 to 13	Qualitative
Kelly et al. (1994)	New Zealand	Aged 9 to 11	2 classes 18 in each class did RT, remainder worked independently
King and Johnson (1999)	USA	Age 10 to 11	5 classes of 25
Lederer (2000)	USA	Ages 9 to 12	6 classes of 19 to 22, Large no. of LD students

Author and date	Country	Sample	Comments
Marks et al. (1993)	USA	Grade 1 (age 6 to 7) Grades 6,7,8 (remedial classes aged 11 to 14) Grades 9,10,11,12 (aged 14 to 17)	Qualitative comparison study. Only one teacher used RT for whole class (Grades 9,10,11,12)
Miller et al. (1988)	USA	Grade 7 (aged 12 to 13)	Each group had a leader who was psychology student or graduate
Myers (2005)	USA	Kindergarten Listening comprehension	Data on four children, case study
Pilonieta and Medina (2009)	USA	Grade (aged 6 to 7)	No data given
Sarasti (2008)	USA (unpublished thesis)	Grade 3 (aged 8 to 9)	Ability groups
Takala (2006)	Finland	Grade 4 and 6 (Ages 9 to 10 and 11 to 12)	Teaching conducted in Finnish RT group had 2 teachers
Taylor and Frye (1992)	USA	Grade 5 and 6 (ages 10 to 12)	No standardised tests used No data to show equivalence of group
Van Keer (2004) 2004	Belgium	Grade 5 (ages 10 to 11)	Only whole class (no interaction) or dyads

On reading these studies however, it soon became apparent that some which purported to involve RT within a whole class situation had not implemented RT as described by Palincsar and Brown. For example Van Keer (2004) included three conditions, one

involved whole class instruction but only individual practice, and the other two conditions used dyads – none of the interventions used peer interaction within small groups. Other studies used more than one teacher in the classroom (for example, the study of Takala, 2004), or had graduate students as leaders for each group (Miller et al., 1998). Three of the studies, although published in English, were conducted in other languages (De Corte, et al., 2001 in Flemish; Takala, 2006 in Finnish; Van Keer, 2004 in Dutch), which not only makes it difficult to assess the materials used, but makes generalisability even more difficult. Brooks, for example, in his assessment of interventions that work, did not include any studies outside the UK on grounds of generalisability, let alone those conducted in another language.

One study grouped the students by ability (Halberstam, 2008), whilst Palincsar and Brown had noted that heterogeneous groupings were the most effective (Palincsar, et al., 1989). This is borne out in other studies, which have shown that when readers are grouped by ability the gap between poorer readers and their higher achieving peers widens (Condron, 2008), the quality of instruction for the poorer readers is itself poor (Swanson, 2008), and the poorer readers spend less time interacting with text (Allington, 2009).

Kelly et al. (1994) did use heterogeneous groupings, but despite saying they wanted to show that RT could be used in "a regular primary school classroom", they only used the technique with six children in each of the two sixth grade (aged 10 to 11) and seventh grade (aged 11 to 12) classrooms. The remaining children in very large classes (an additional 30 in each class) continued with "assigned reading tasks". Thus, although the results showed that the experimental groups made significant gains on a standardised reading test post-intervention and at an eight week follow-up, it does not show that RT can be implemented for a whole class, but only that it was possible for one group within that class. However, Kelly et al. did strengthen their claims for the programme's success by including an attention-only comparison group, showing that RT has an effect over and above that produced by teacher-directed small group work. And as the results were obtained, as the authors state, "within the constraints of the regular classroom, with no additional resourcing for the teachers and students involved, but with the usual distractions, noise and organisational disruptions associated with a regular classroom programme", this study does demonstrate that RT within a whole class might be feasible.

Using Brooks' checklist of requirements, it was apparent that not one study met all the criteria. Some studies were purely qualitative and did not provide any data (Coley,

DePinto, Craig, & Gardner, 1993; Hacker & Tenent, 2002; Hashey & Connors, 2003; Marks et al., 1993; Pilonieta & Medina, 2009). Other studies provided some quantitative data, but failed to show pre-test comparability of the groups (Carter & Fekete, 2001; Lederer, 2000; Taylor & Frye, 1992), one showed that the groups were not equivalent at base-line (King & Johnson, 1999) and one only provided case study material (Myers, 2005).

One study which did fulfil the majority of the criteria however, was that of Alfassi (1998). There were no effect sizes reported, but there was enough data given for Galloway to calculate them in her review (Galloway, 2003). Thus, for Alfassi's study of 14 to 15 year olds, there was a large effect size (.81) on the experimenter-devised test, but a small negative effect (-.26) on the norm-referenced test, with a follow up effect approaching the moderate range (.52)[8]. Alfassi explains the difference between the test results as an artefact of the assessments used, and this is discussed more fully in the next section.

Thus, there is only one study conducted in English which fulfils the standard of comparability of groups at the outset, which provides statistical analysis from which to calculate effect sizes, and which implements RT as Palincsar and Brown intended, viz. using the four strategies of predicting, clarifying, questioning and summarising, and giving students turns in fulfilling the teacher's role during group interactions with text. However, although Alfassi contends that this is a naturalistic situation, the students were assigned to teaching groups (some as small as 8) which meant a much smaller class than would be expected in most school situations (the class sizes in the three studies in the present research were 16/17 In Study 1, 28 in Study 2 and 27 in Study 3). This is only a whole class intervention therefore, in the sense that it is a whole class of remedial readers, already removed from the naturally constituted group. All the students were classed as poor comprehenders by Alfassi, and additionally they were aged 14 to 15. Therefore, even this study has little similarity with the target group of the present research - primary aged children in a heterogeneous class in the UK - which therefore remains unresearched. There are no whole class studies that have covered the age groups sampled in Study 1 (9 to 10 years) and Studies 2 and 3 (7 to 8 years).

---

[8] Galloway states that there was no statistical difference between the follow-up results on standardised and experimenter-designed assessments, so she only gives the follow-up effect sizes for a combination of the two.

2.5. The crucial role played by assessment

Differences between results obtained by norm-referenced tests and experimenter designed tests have been a recurring theme in the RT literature. One study which examined this difference in depth is that of Alfassi (1998), reviewed in the previous section. The study involved 53 high school students in a remedial class, who were taught in groups of 8-15. Results revealed that after 900 minutes of intervention on a daily basis, there was a significant improvement on an experimenter designed test, but no significant improvement on a norm-referenced test. As this was a pattern found by Rosenshine and Meister (1994), Alfassi tried to explain why this might have happened in her study. Table 6 shows the differences between the conditions of the standardised test and those of the test designed by Alfassi.

Table 6. Differences between the standardised test and the experimenter-devised test in Alfassi 1998 (table adapted from Alfassi, 1998).

Differences	Standardised test (Gates-MacGinitie)	Experimenter-developed test
Number of words per passage	60-130	335-550
Average number of words per passage	88	400
Text available for consultation by student	Yes	No
Number of passages administered	14	2
Narrative passages	64%	None
Expository passages	36%	100%
Passages with topic questions	36%	100%
Types of question	Multiple-choice	Short answer
Number of questions per passage	2-9 (48 questions per test)	10 (20 questions per test)
Vocabulary dependent questions	40%	None
Questions requiring inference	50%	None

It is clear from the information given by Alfassi, that the tests are examining very different things. The standardised test (Gates-MacGinitie) requires the ability to search the text and recognize the correct answer, strategies which are not taught in RT. The experimenter-designed test required the ability to understand and recall ideas, and it could be argued that this is what we want children to be able to do when reading – to be able to learn from text is more important than being able to locate the correct answer from a given list, which is a very artificial situation. Thus, it is very important that the difference between the two types of tests and their requirements is acknowledged in articles which report results from different sources. And the question of what we are measuring in comprehension tests is critical.

2.6. Criticisms of Reciprocal Teaching and Palincsar's response

As we have seen, Hirsch's article in American Educator (Hirsch, 2006) was particularly critical of the "deadening activities" of strategy instruction. As Palincsar acknowledged (2006) this was "uncomfortable" to read (p.1). However, what Hirsch was referring to was not RT as understood by Palincsar and Brown, but the "lethal mutations" which reduce the intervention to the direct instruction of strategies and disconnect them from their purpose of promoting active, strategic reading through scaffolding, feedback and peer discussion (Brown & Campione, 1996). Palincsar reiterated that RT is a dialogic approach in 2002, when, in a presentation to the CEIRA Institute, she defined RT as "dialogic instruction in which comprehension monitoring activities are introduced and practiced for the purpose of understanding in the context of reading" (slide 10). Hirsch describes teachers whose time is devoted to skills training which drives content out of the classroom, but Palincsar argues that a strategic reader is using knowledge through conscious processing of the text. In forming a coherent situation model the reader must engage with the text through activating prior knowledge. This is particularly important when reading about an unfamiliar topic, as connections will not be made automatically. The reader must search for links and seek to understand how parts of the text relate to one another, and then decide which parts are important. To develop a deep understanding the reader needs to engage in problem solving both within the text and within their own knowledge. Without this active process they will be left with only a textbase level of understanding.

Palincsar feels that strategy instruction "has been extracted from the theoretical context in which it was originated, rendering it anaemic" (Palincsar & Schutz, 2011 p.88). Strategies need to be embedded in their purpose of understanding text and not presented

discretely. A parallel might be drawn here with research about deep and surface learning, where attempts to induce deep learning by the application of techniques exhibited by deep learners actually resulted in an increase in surface learning. Marton (1976) attempted to inculcate a deep learning approach in an experimental situation by asking a group of students the kinds of questions that would be posed by someone trying to understand a text. Among the subsections of a chapter he introduced questions about relationships and summarising. A control group was given no such guidance. Surprisingly the control group out-performed the "coached" group on tests taken immediately afterwards and after a time delay of two months. Marton argued that the predictability of the questions had induced a surface approach as the coached students invented a way of answering which fitted the demands, but at the same time made the task rather trivial and mechanical.

To ensure that RT is not mechanical and over-concerned with process, Palincsar outlined three principles which she felt would prevent decontextualisation and encourage active processing of text – choose related texts, remain close to the bone, and focus on knowledge building. She elaborates on these three themes, stressing that strategies need to be used for an authentic purpose and with the aim of supporting students "to build understanding by engaging in activities that lead to self-regulation and can be internalized over time" (slide 32). The idea of self-regulation was explored in a recent study on RT by (Schünemann, et al., 2013), which compares RT groups with groups instructed in RT plus explicit instruction in self-regulated learning (RT + SRL). However, some of the SRL strategies were those previously well-documented in RT, and which were taught in all three studies in the research for this thesis. For example, the use of a "help sheet", which "expressly described what to do and how to help the students whilst acting as a group captain" (p. 21) equates to the bookmarks given to the children which outlined how to use each strategy effectively (Oczkus, 2003). The authors of this study state that the SRL strategies they taught, viz. goal-setting, self-monitoring and self-evaluation "fit very well with Palincsar and Brown's reciprocal strategies programme" (p.9). Indeed, they fit so well, that they can be considered part of RT itself. By giving feedback the 'teacher' in each group is learning to monitor and evaluate, and these skills are also modelled by the 'teacher'. The results section of the paper shows that there was no significant difference between the two types of treatment (RT and RT + SRL) on strategy-related task performance. The authors explain this by outlining the many common features of the two programmes. However, eight weeks after the intervention, maintenance testing did reveal a significant difference between the groups (in favour of the RT + SRL condition) in strategy-related task performance and

reading comprehension. The authors claim this is in line with other studies which show the effectiveness of self-regulated learning activities, which can enhance training effects and prevent the loss of newly acquired strategic skills (Glaser & Brunstein, 2007; Hager & Hasselhorn, 2000; Souvignier & Trenk-Hinterberger, 2010). However, I contend that self-regulated learning strategies are an integral part of RT, and that by trying to deliberately exclude elements of self-regulation, Schunemann and colleagues are in danger of using a mutation. RT is therefore concerned with learning the "what" of the four strategies, but it is also concerned with learning the "how" and "when" of using them, that is, the declarative, procedural and conditional knowledge of Paris et al. (1983) (discussed previously in section 1.5.2.3.); and in learning the how and the when, self-regulated learning would appear to be integral.

2.7. What we still need to know about Reciprocal Teaching

Despite the amount of research into RT, there remain unanswered questions. The following sections will examine these in turn.

2.7.1. How does Reciprocal Teaching work?

Despite the positive effects on reading comprehension of RT that we have outlined in this literature review, relatively little is known about the specific mechanisms that lead to the improvements (Schünemann, et al., 2013). An increase in the use of reading strategies might be perceived as a mediator of training success, but direct evidence of this is limited (Sporer, Brunstein, & Kieschke, 2009). The qualitative data in Studies 2 and 3 will provide evidence of how reading strategy use changes and develops over the training period.

2.7.2. Does it work in the UK?

With very limited data from classrooms in the UK, and none of that from whole class situations, there is a need to discover if RT can be effective in this country when it is delivered at the level of a whole class. As Brooks (2007) stated at the beginning of his report on what works for pupils with literacy difficulties, information from other parts of the world often meets the objection, "How do we know it will work here?" (p.13). Brown and Campione (1996) pointed out that the effectiveness of a design in one setting is no guarantee of its effectiveness in other settings, but there is more chance of its being repeated if the settings are as alike as possible. This means that there is a need to study RT within the UK.

82

2.7.3. Does it work for younger children?

As Galloway (2003) reported, there is not a single study which has focused on the elementary school years, which met her strict criteria, and of those which looked at late elementary grades (Grades 4 to 6) many were carried out in special schools. Some of the original research studies, however, did focus on younger children. Palincsar and Klenk (1992) reported that some 300 children who were first to third graders (ages 6 to 9) had participated in their research since 1981. Eighty per cent of these students, although commonly below the 40th percentile on standardised measures of reading comprehension at the outset, had improved their scores on independent measures of text comprehension from 30% to 75% or more after 3 months of instruction. These gains had also been maintained on follow-up measures between 6 months and a year later (Brown & Palincsar, 1989; Palincsar & Brown, 1984; Palincsar & Brown, 1989).

Despite the positive results from the early research with younger children, Cain (2010) argued that "it is not clear that younger children would get as much advantage from what is essentially a rather cognitively demanding task" (p. 185). RT has been adapted for children as young as kindergarten, however. Myers (2005) used interactive think-alouds to demonstrate reading strategies to children aged 4 to 5 in a whole class situation. Although this small action-research project only reports anecdotal evidence for four case study children, the author does feel that there was an increase in the children's ability to think metacognitively, to ask questions, to comprehend and to be involved. There was also evidence of some transfer to other situations, leading Myers to conclude that kindergarten children can benefit from RT, even when these very young children have difficulty with expressive language or are second language learners.

Pilonieta and Medina (2009) have also written about using RT with first graders (aged 6 to 7). Unfortunately, we only have a 'how-to' article, and not the data, but the authors say that the students learned the strategies, applied them to new content and texts, and learned how to apply the strategies in sequence in small groups within a whole class structure. Furthermore, the declarative, conditional and procedural knowledge of this modified form of RT was retained six months later when the children had moved on to second grade. Although the tangible outcomes are not yet available, what the authors call "the intangible outcomes" were visible in an increase in engagement and motivation, and students were observed actively participating in higher order thinking as they discussed the texts they were reading. Children as young as first grade were, according to the authors,

"able to learn, coordinate and apply comprehension strategies and work in collaborative groups" (p. 128).

2.7.4. Does Reciprocal Teaching work in a whole class situation, without the need for the additional resources which small group work implies?

Kratochwill and Stoiber (2002) recommended that educational interventions should be in line with educational practices in schools to make them viable. Since most teaching takes place within the unit of the class, then it is within such a unit that this thesis sought to examine the effectiveness of RT. As has been discussed earlier, previous work on RT has focused on small groups, so by using extant classes this thesis will make a contribution to determining whether RT can be effective in whole class situations.

2.7.5. Could Reciprocal Teaching be improved by the addition of visualisation?

Visualisation has been shown to be effective by research which has taken place over decades. Linden and Wittrock, (1981) showed that drawing attention to text by constructing verbal or imaginal elaborations improved comprehension in fourth graders, whilst Oliver (1982) devised a set of three investigations in elementary schools and concluded that visualisation enhances comprehension. Other studies reached similar conclusion (Gambrell & Bales, 1986; Gambrell & Bales, 1987; Gambrell & Koskinen, 2002; Kulhavy & Swenson, 1975; Sadoski, 1983). These studies support the findings of Paivio (1986) who suggested that there is a nonverbal dimension to the processing of discourse, and this route can aid comprehension. Imagery training has been reported to help poor comprehenders (Oakhill & Patel, 1991), even in an intervention lasting for just three sessions. More recently, imagery training has been shown to help children with specific language impairment (Joffe, Cain, & Maric, 2007). Cain (2010) suggests it can be of benefit to children who have a language-based problem since it is a nonverbal strategy. By adding mental imagery to the strategies of RT, children are offered a further way to become actively engaged in processing text, which is the ultimate aim. As Cain also suggests, nonverbal supports such as imagery training are suitable for young readers, and it may therefore be particularly appropriate to add it to the strategies when introducing RT to younger children, as was the aim in the present research.

2.8. What did reviews tell us a new study on Reciprocal Teaching needs to incorporate?

2.8.1. A randomised trial including a control group

The NICHD (National Institute of Child Health and Human Development, 2000) and the What Works Clearinghouse (Shanahan, et al., 2010) argue that strict standards need to be applied before evidence can be accepted. The Practice Guide on Improving Reading Comprehension, published by WWC (Shanahan, et al., 2010), considered 812 studies, but only 27 studies met their standards without reservations. Whilst there is strong preference for RCTs (one large RCT with positive effects is sufficient to provide "strong evidence", in the absence of contradictory evidence) quasi-experimental studies are considered acceptable if well designed. However, the highest rating a well-implemented QED can receive is "meets evidence standards with reservations". There may be compelling reasons as to why randomisation is not possible, for example in Brown et al. (1996) the amount of time required to become a Transactional Strategies Instruction teacher precluded randomly assigning a group of teachers to either an instruction or control group. Conversely, the authors felt that TSI teachers were committed to strategy instruction in such a way that they would not be able to alter their teaching for a whole year in order to teach a control group.

The WWC contend that where a component of the design lines up with a unit, for example a teacher or classroom, then there may be other unobservable confounding factors such that even if observable characteristics have been established as equal, then such a study cannot be used as evidence of a programme's effectiveness. Unfortunately, if the experiment is designed to examine the effectiveness of an intervention for a whole class, then it becomes impossible to fulfil these criteria. The best that can be done therefore is to ensure that groups are as comparable as possible on all critical characteristics. However, it can also be argued that if an intervention is to be used in the real world of classrooms in schools, then no amount of laboratory based randomised controlled trials will provide the evidence that the intervention can be transferred to a natural classroom setting. Complete naturally constituted classes are needed to show applicability in the real world (Van Keer, 2004).

In the UK, Brooks (2002, 2007) has also been critical of intervention studies. His reviews of what works for literacy instruction comment on the literature reviewed as ranging from "the meticulous to the very weak" (p. 110). The "meticulous" is the RCT, which is again upheld as the gold standard. And in common with the WWC, far more studies were

rejected from his reports than included, although the exact figures are not given. Some studies did not include any quantitative data at all, whilst others had data that was unclear, or based on very small samples. Brooks (2007, p. 111) recommended that all evaluations include "as a minimum" the following:

- the date when the evaluation was carried out (in addition to the date of reporting);
- the exact age range of the children involved;
- salient characteristics of the children involved, for example, whether they had any special educational needs;
- the numbers of children in the experimental and control/comparison groups and in any alternative intervention groups;
- how the children were assigned to the different groups, for example, randomly or by matching;
- the nature of any alternative intervention;
- the exact length of any intervention;
- the tests used;
- the pre- and post-test average standardised scores and standard deviations for every group involved; this would make it, strictly speaking, unnecessary to report the amount of gain, but this might be interesting in itself;
- the statistical significance of the differences between groups at pre-test, so that the initial equivalence of the groups can be shown, or the statistical handling of any significant pre-test differences can be explained;
- the statistical significance of the differences between pre-test and post-test scores for each group, so that it can be seen whether or not the absolute value of any gains was statistically significant; this is easier for standardised scores than for reading/spelling ages;
- the effect size, so that the impact of the approach can be compared with others;
- any follow-up data that is available.

It was intended that Brooks' guidelines would be followed in the initial study for this thesis (Study 1) but since random allocation was to be at the level of the class as a unit, it was the case that the study would not meet the stricter guidelines of the WWC, since as stated previously, the intervention was being investigated at the level of the class, which the WWC precludes. There is a body of thought however, which holds that although the RCT may be the 'gold standard', there are circumstances where it is not appropriate in

education. The National Research Council in the USA (2002) stated that " a variety of legitimate scientific approaches exist in education research" (p.98) and that "research that explores students' and teachers' in-depth experiences, observes their actions and documents the constraints that affect their day-to-day activities provide a key source of generating plausible causal hypotheses" (p.109). Thus there is a belief that RCTs should be supplemented with other methods, which includes in-depth qualitative approaches that can illuminate important nuances, and provide additional sources of evidence. It was the intention in the current research to utilise such qualitative methods in Study 2 and Study 3. After the initial investigation of the effectiveness of RT in a whole class setting in Study 1, in the subsequent studies the aim was to find out more about the "how" and "why" of what is happening when RT is taking place. To understand the how and why, a different approach is required beyond the quantitative, so a mixed methods approach was considered to be appropriate. Collins, Onwuegbuzie and Sutton (2006) identified four different rationales for conducting research in this way: participant enrichment, instrument fidelity, treatment integrity and significance enhancement. Using quantitative and qualitative measures together in the subsequent two studies fulfils the latter three of these rationales. Interview and think-aloud data maximise the appropriateness of the instruments used, since the quantitative assessments of reading comprehension (which focus on the end result) are supplemented by an examination of the ongoing processes involved in reading. Treatment integrity is enhanced since the differing assessments provide triangulation and there is significance enhancement since the mix of techniques maximises the interpretation of the data.

2.8.2. Controlling characteristics of the comparison group

If groups are to be compared in an experimental situation then steps must be taken to ensure that they are as comparable as possible on relevant pre-test measures. In what has been called a well designed and executed study (Pressley, Graham, & Harris, 2006) Brown et al. (1996) took steps to ensure that all groups were also given quality instruction by ensuring that the control group's teachers all had reputations as excellent educators. In Study 1 of the present research, data is presented to show that the groups were equivalent at the pre-test stage, and the control group was instructed by experienced teachers.

2.8.3. A way of establishing fidelity to the Reciprocal Teaching programme

As Rosenshine and Meister contend, evidence is needed to show that RT is being correctly implemented. Information concerning the content of lessons and the way the programme is administered enable the reader to determine if RT is taking place in the way in which the Palincsar and Brown intended, or if it is a "lethal mutation" (Brown & Campione, 1996 p.292). Some studies have done this by using a checklist of the activities which should be observed in a RT lesson (Andreassen & Braten, 2011; Shortland-Jones, 1986) or by using fidelity of treatment protocol (Hart & Speece, 1998). Despite the recommendations by Rosenshine and Meister however, Galloway (2003) reports that only 14% of the studies she reported used a form of coding to score fidelity, and only Hart and Speece provided a fidelity rating. Galloway also reports that some studies have provided sample dialogues as evidence, but that these are not often evaluated in a way that permits comparison across studies. It is necessary therefore to provide evidence of fidelity in a manner which permits replication, and which can be used for comparison. In the three studies presented in this thesis, fidelity to RT was assured by following the guidelines of Rosenshine and Meister as to what constitutes RT, and the delivery of the instruction was either by the researcher, or by a class teacher in the presence of the researcher.

2.8.4. Careful consideration of outcome measures

Since considerable differences were observed in the effectiveness of RT according to whether standardised tests or experimenter-designed tests were used in the studies included in Rosenshine and Meister's review, the decisions about which tests to use should be carefully made. There will be more discussion about this in the measures section of the chapter for each study.

2.8.5. Outcome measures examining strategy application as well as reading comprehension

If RT works through increased application of strategies, then a measure of strategy use needs to be included alongside a measure of reading comprehension. The measures used will be discussed in the relevant chapter for each study. Additionally, the present research was concerned with the process of reading, rather than just the end results, since as Janssen, Braaksma and Rijlaarsdam (2006) report, "teachers and textbooks tend to focus on reader response as the end result of a reading process, and not so much on the process

itself" (p.35). Thus the second and third studies included measures looking at strategy use as it occurs during the process of reading.

2.8.6. Follow-up measures

If an intervention is delivered, then we need to know not only about any immediate gains, but also about possible longer lasting effects. For this reason Rosenshine and Meister (1994) and Brooks (2007) maintain that follow-up measures are necessary and such measures were incorporated in the present research.

2.9. Aims of this research

The author of this thesis is a primary school teacher with an interest in improving reading comprehension in her students. It was intended that a form of action research be conducted with a convenience sample of three Year 5 classes in the school in which she taught. If the effectiveness of RT could be demonstrated, and the effects of adding another strategy, visualisation could be determined, then a second study would look in more detail at the effects of RT on individual children. Thus, the first study was to involve largely quantitative data, whilst the second and third studies would adopt a more mixed methods approach, by using qualitative methods in addition. The qualitative approach would also ensure that the reader is not left out of the equation (McNamara & Magliano, 2009) when comprehension is being considered. The aim of including the qualitative data was to make clear the reader's use of strategies and level of metacognitive awareness, and how these may change during RT instruction. The aims of each study are given below.

Study 1:

The aim was to investigate whether RT could be an effective method for improving reading comprehension in a whole class situation, in the UK, for children in Year 5 (aged 9 to 10).

A supplementary aim was to examine whether the addition of a fifth strategy, visualisation, would increase the effectiveness of RT in improving reading comprehension.

Study 2:

The aim of Study 2 was to investigate the changes which take place in reading processes and motivation when RT is used in a whole class situation.

A supplementary aim was to examine whether the addition of a fifth strategy, visualisation, would increase the effectiveness of RT in improving reading comprehension.

A further aim was to investigate how RT was implemented by a class teacher, as opposed to the researcher, but within the same school.

Finally, by using follow-up assessments the aim was to ascertain if any improvements in reading comprehension or strategy use were maintained.

Study 3:

The aim of Study 3 was to investigate the changes which take place in reading processes and motivation when RT is used in a larger whole class situation.

A supplementary aim was to examine whether the addition of a fifth strategy, visualisation, would increase the effectiveness of RT in improving reading comprehension.

A further aim was to examine a second class teacher's implementation of RT, within a different school.

Finally, by using follow-up assessments the aim was to ascertain if any improvements in reading comprehension or strategy use were maintained.

Chapter 3: Study 1 - Testing the effectiveness of Reciprocal Teaching in children aged 9 to 10

"I learned how to help me read."

3.1. Research Aims

The aim of Study 1 was to test the effectiveness of RT as a method of improving reading comprehension within a whole class setting. A supplementary aim was to see if the intervention could be improved by adding a fifth strategy, visualisation, to the four strategies originally taught; predicting, clarifying, questioning and summarising.

3.2. Study 1 Methodology

The study was conducted between January and May, 2008. The design consisted of pre-test assessment, intervention and post-test assessment. Following the increasing need for intervention studies to include a control group, Year 5 was chosen as a convenience sample within Key Stage Two at the school where the researcher taught, as it was the only year group with three classes. This enabled three groups, an RT group, an RT plus visualisation group, and a waiting control group. It was not possible to make this a fully randomised controlled trial, since the children were already in three classes, but there had been random allocation of children to these classes on entry to the school and since then children had been moved to ensure the class sizes remained equal and to ensure that the three classes were balanced in terms of achievement and behaviour. The pre-intervention testing sought to show that the groups were comparable on critical variables. There was random assignment of the three classes to the three interventions.

3.2.1. Participants

The participants were nine- to ten-year-old children in three parallel Year 5 classes of a home counties UK primary school where the researcher worked as a special needs teacher. The school is an independent junior school which follows the UK National Curriculum and takes part in the National Curriculum Standard Assessment tasks. Owing to its charitable status, the school has a more socially diverse intake than the average independent school. The school is non-selective in an area of highly selective schools resulting in an intake that is also more diverse in terms of ability.

The classes consisted of 17 or 18 children each, equally divided between genders. Two classes had been joined just before the study began by a child with English as an additional language (EAL). These two children were not included in the present study; they received specialist EAL teaching outside the classroom during the RT sessions. This left 17 children in each class, each comprising 9 girls and 8 boys. The results for one child in the RT group were excluded from the analyses as she was found to have word reading skills far below that of the other children in the sample. Her score on the standardised TOWRE word reading test (Torgesen, Wagner & Rashotte, 1999) was more than two standard deviations below the mean for the group.

The three Year 5 classes were randomly allocated to one of three conditions: Reciprocal Teaching (RT), Reciprocal Teaching plus visualisation (RTV) and a normal instruction control group (NI). Table 7 gives a summary of participant characteristics for the three groups in terms of chronological age in months, and standardised scores on the Cognitive Abilities Tests (CATs, Lohman et al., 2001). The CATs are administered as part of school policy to all children in Year 4. One child was not present at the school at the time, so data was available for 49 out of the 50 children for whom data were collected.

Table 7. Mean chronological age and standardised scores on Cognitive Abilities Tests for the children in the three groups in Study 1 (standard deviations are in parentheses).

	RT	RTV	NI
N	16	17	17
Chronological age (months)	117.25	118.82	118.71
	(2.82)	(3.41)	(3.40)
CAT verbal	110.31	106.06	104.19
	(11.19)	(11.92)	(10.78)
CAT non-verbal	110.50	106.12	107.63
	(14.15)	(10.64)	(11.86)
CAT numerical	108.38	104.41	103.31
	(13.71)	(8.46)	(11.09)

Note: CAT = Cognitive Abilities Test

Analyses of the data revealed no significant difference between the groups in chronological age ($F(2,47)=1.20$, $p>.05$), CAT verbal reasoning ($F(2,44)=1.55$, $p>.05$), CAT non-verbal reasoning ($F(2,44)=1.80$, $p>.05$), or CAT numerical ability ($F(2,44=.96$, $p>.05$).

3.2.2. Ethical issues

As the study involved participants under the age of 18, ethical approval for the study was obtained from the Department of Psychology and Human Development, Institute of Education, London. Letters describing the study were sent out to the parents/carers of all Year 5 children and the opportunity was given for any parent/carer to opt out of the study. A copy of the letter can be found in Appendix B. Data was anonymised, and all computer records referred to numbers. The key to the names was kept in a locked drawer in the researcher's office.

3.2.3. Pre- instruction measures

Assessments of single word and non-word reading ability and of reading comprehension were administered to the children in the three groups prior to the instruction in January-February 2008 (term 2 of Year 5). The tests employed were the Test of Word Reading Efficiency (TOWRE Torgesen, Wagner, & Rashotte, 1999) and the Neale Analysis of Reading Ability (NARA, Neale, 1989). The TOWRE assesses single word and non-word reading, while the NARA assesses reading comprehension, accuracy and reading rate. According to the Simple View of Reading, decoding is fundamental to comprehension so it was necessary to ensure that all pupils were able to decode sufficiently well that decoding could not be considered to be a problem for comprehension – i.e., that it fell within the normal range for age. Similarly, it was important to select a reading comprehension test which was not strongly dependent on decoding. The NARA allows the tester to supply the correct word when the child cannot read a word during reading of the text passages, or else does not self-correct. Additionally, the NARA relies on spoken questions and answers for assessing comprehension, thus ensuring that the reading comprehension assessment is not confounded by writing ability, as can be the case in tests requiring written answers (Jenkins, Johnson, & Hileman, 2004). In the NARA, children read aloud from a series of short narratives and are then asked questions about what they have read. Testing continues until more than sixteen reading errors are made in one passage. It was an appropriate test to use in this study since it is closely aligned to the way RT was taught, in that the passage is read

aloud and then questions are answered orally, which is the same format as that used in the intervention.

The TOWRE assesses reading accuracy and fluency for words and pronounceable non-words. Children are given 45 seconds to read as many items as possible for each list of words and non-words. Standardised scores were obtained for comprehension, accuracy, and rate measures in the NARA, and for sight word reading and phonemic decoding measures in the TOWRE. Table 8 provides a summary of these pre- instruction reading scores for the children in the three groups.

One of the criteria for inclusion of studies in the influential review of RT conducted by Rosenshine and Meister (1994) was that where random allocation of participants to treatment and control groups was not possible (as in the present study) then classes involved should be similar on initial measures of reading comprehension. The scores for the assessments were therefore analysed using one-way ANOVAs, in order to ascertain whether the three groups were comparable on pre- instruction reading measures. The results from the NARA revealed no significant effect of group for comprehension (F(2,47)=0.42, p=0.66), accuracy (F(2,47)=0.18, p=.84, or rate (F(2,47)=1.13 p=.33). In addition, the results for the TOWRE revealed no significant effect of group for sight word reading (F(2,47)=0.60, p=.55) or phonemic decoding scores (F (2,47) = 0.60, p=.55).

Table 8. Mean standardised scores on the NARA and TOWRE reading assessments prior to intervention for the children in the three groups in Study 1 (standard deviations are in parentheses).

	RT	RTV	NI
NARA Comprehension	100.44	102.72	103.88
	(11.15)	(10.58)	(11.16)
NARA Accuracy	100.44	102.71	101.29
	(11.51)	(10.23)	(11.36)
NARA Reading rate	109.31	112.94	108.29
	(9.47)	(9.15)	(9.70)
TOWRE Sight word	108.88	108.52	106.88
	(9.26)	(7.45)	(8.39)
TOWRE Phonemic decoding	109.19	111.46	107.94
	(12.38)	(9.66)	(15.58)

The children's self-reported use of reading comprehension strategies was assessed using an adapted version of the Metacognitive Strategies Index (MARSI, Mokhtari & Reichard, 2002). Participants responded to 30 statements (e.g. "I try to make pictures in my head of what is happening") using a 5-point Likert scale, ranging from 1 ("I never or almost never do this") to 5 ("I always or almost always do this"). The questionnaire was adapted for present purposes from the original, as it had been developed for use with students in Grades 6 to 12, and it was felt that some of the questions were inappropriate for the younger participants in the present study. For example, the statement "I think about whether the context of the text fits my reading purpose" was omitted, whilst other statements were amended when the vocabulary was felt to be too difficult; for example, the original questionnaire "I preview the text to see what it is about before reading it", was amended to "I look over the text quickly to see what it is about before I start reading". The questionnaire used can be found in Appendix C. The questionnaires were scored by calculating an average of the points checked on the scale. The overall pre-instruction scores indicated a medium level of strategy use by the children, according to the MARSI authors' classification (RT group $M = 2.75$, $SD = .73$, RTV group $M = 2.92$, $SD = .79$, NI group $M = 3.01$, $SD = .57$). A one-way ANOVA revealed no significant effect of group for pre-instruction level of strategy use ($F(2,47)=.57$, $p=.57$).

3.2.4. Post-instruction measures

The NARA was administered again at the end of the instruction period in order to see whether gains occurred in scores across the three groups. Parallel forms are available for the test. Form I was used prior to the instruction period and Form II following the instruction period. The adapted MARSI questionnaire was also re- administered following instruction in order to see whether there were any changes in self-reported reading comprehension strategy use.

Finally, a short questionnaire (reproduced in Appendix D) was given to the RT and RTV groups to enable them to evaluate the instruction. All the post- instruction assessments and the evaluative questionnaire were administered in May 2008 (in term 3 of Year 5). Following completion of the study, the two RT classes gave a presentation on their experience of the instruction to Year 5, their teachers and the head teacher of the school, based on a transcript of a RT session. The transcript can be found in Appendix E.

3.3. Procedure

Following the administration of the pre- instruction assessments during January-February 2008 the study ran from February to May, with the Easter school holiday in April. The study covered a period of 14 weeks. The RT and RTV groups received an hour a week of instruction, delivered by the researcher, for a total of ten sessions. Ten hours of instruction is a usual amount based on previous RT research (e.g., in Galloway's (2003) meta-analysis there was an average of 678 minutes of instruction across studies reviewed) and fits in well with a term's timetabling. The RT group were taught for an hour on Monday morning, while the RTV group were taught for an hour on Thursday morning.

3.3.1. Details of the Reciprocal Teaching intervention

The present study followed the criteria for RT from Rosenshine and Meister (1994) as outlined in chapter 2. The first session involved an introduction to RT, and the importance of the children being involved in the teaching and learning. Rules for group discussion were also established. Then the strategies were introduced, at the rate of one a week, with the use of think-alouds and modelling by the researcher. Thinking aloud provides valuable information about how a reader builds a situation model. The teacher models what is going on in their head as they read by predicting from the title, headings and illustrations, and then by stopping to add thoughts as they occur throughout the text. The teacher might clarify a word or a sentence, ask questions about the text and summarise each paragraph. An example of a teacher thinking aloud is provided below (Oster, 2001), with the teacher's comments in brackets:

> The day Professor Herbert (Professor – that must mean this is a college) started talking about a project for each member of our general science class; I was more excited than I had ever been. ('I'- this is first person narration, the main character in the story is telling it himself. Also, he must like science. Well, it might not be a he?). I wanted to have an outstanding project, to be more unusual than those of my classmates. (He has high hopes! Or, he's ambitious.) I wanted to do something worthwhile, and something to make them respect me. (I wonder why he needs their respect?) (Stuart, 1979).

Think-alouds have been used in the past as teaching tools (Davey, 1983; Oster, 2001) for analysing what readers do as they read (Pressley & Afflerbach, 1995) , and for assessment (Klingner, et al., 2004; Leslie & Caldwell, 2011). In Study 1 they were used as a teaching tool,

and in Studies 2 and 3 they had an additional role as an assessment. Rosenshine and Meister (1994) noted that think-alouds were included in a later study by Palincsar (Palincsar, 1987) but not in the earlier reports (Palincsar, 1982; Palincsar & Brown, 1984). They did form an important part of the instruction in the present research, being utilised in modelling the procedures.

The introduction to the strategies followed the explicit teaching before RT (or ET-RT) as defined by Rosenshine and Meister (1994). This was the approach taken in some of the later studies by Palincsar and colleagues (Palincsar, et al., 1987; Palincsar, David, Winn, Stevens, & Brown, 1990). Rosenshine and Meister (1994) found that out of the seven studies using ET-RT, six of them reported significant improvements in reading comprehension scores, suggesting this is an effective method. Practice and feedback took place as whole class and group activities. Children were also given two bookmarks with the four strategies and their key ideas, adapted from Oczkus (2009). The use of cue cards was one of the adaptations implemented in Collaborative Strategic Reading and it was felt to be suitable for a younger age group who might need more support.

Both RT groups used a novel as the text for their group work - *Cue for Treason* (Trease, 2002). This was chosen as it fitted with the history curriculum. Different sections of the text were studied each week as material for practising strategy use. The researcher also used sections to model think-alouds, and read or summarised the chapters in between so that the novel was covered in its entirety by the final week.

After the initial sessions introducing the strategies, each lesson began with a brief recap and then the children, in mixed ability groups of four or five (which remained constant throughout the intervention) worked on a page of text, taking it in turns to be the teacher. Mixed ability groups were used as previous research has shown that peer interaction is most successful when one child is more skilled than the other (Cohen, 1994; Fuchs, 1996; Webb & Palincsar, 1996). Each session usually involved two sections of text, providing the opportunity for two children to be leaders. Records were kept to ensure that every child took the role of teacher on at least two occasions. The roles of predicting, clarifying, questioning and summarising were also undertaken on a rota basis. The researcher moved from group to group providing feedback as necessary, and led a plenary session at the end of each session. Usually a group who had worked particularly well, or who had raised an important point, was chosen 'to go in the goldfish bowl', which entailed modelling to the whole class. Most of the activity in the RT sessions was oral. The children did have exercise

books which were used in some sessions, but they were primarily used for notes and reminders. For example, notes were made of questions the children wanted to ask and records were kept of whose turn it was to be the teacher. The exercise books were not used for formal writing activities; Spörer, Brunstein and Kieschke (2009) had used written worksheets for a paired RT condition and found that this led to a completing the task orientation, rather than lively discussion.

3.3.2. Details of the Reciprocal Teaching plus visualisation procedure

For the RTV group, RT took place as outlined above, but in addition visualisation was taught for half the lesson for the first five weeks. Thus the RTV group were taught five strategies, these being the four strategies of RT plus visualisation. Following the programme set out by Bell (1991) children were asked to use four categories and twelve terms to facilitate their visualisation. The categories/terms were: what (number, size, shape, and colour) where (background, perspective) when (time) and how (movement, mood, sound, smell and texture). The children proved adept at using the terms and showed no difficulty in understanding them despite the difficult vocabulary involved.

Initially the RTV group practised using the categories to describe a picture in front of them. Then they practised visualising a personal image (e.g. a family pet) followed by visualising a familiar noun (e.g. a boat or a clown). At each step, the emphasis was on providing a detailed description using the terms. From working with a single word children moved on to single sentence visualisation. When visualisation became more automatic children were asked about the main idea of a paragraph and their interpretation of what they had read. Although Bell (1991) recommends that children use squares of blank paper to visualise in their heads, the children also used the squares to draw pictures of what they could see. An example is given in Figure 3. The children requested to do this activity, but they did it as homework, to avoid taking away time from RT instruction.

Figure 3. Drawings of a child's visualisations from Study 1.

The mystery of how salmon can find their way back to their home rivers is solved. The salmon navigate by sun and stars when travelling in the ocean.

When the salmon nears the general area of the river in which it was born, it uses its nose.

The salmon can remember the smell of the home river that it left as a baby.
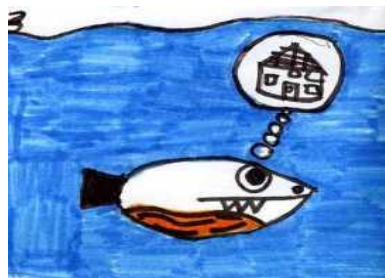
The NI group received an hour of literacy instruction per week over the same time period as the other two groups, following the normal teaching practice. This instruction was delivered by the three Year 5 teachers, involving practice exercises using a comprehension text-book and reading a novel for discussion as a whole class. The three teachers took it in turns to teach these lessons, with the NI group therefore also receiving instruction from a teacher different to their class teacher. The NI group were not taught by the researcher to ensure that they were not exposed to any of the RT or visualisation techniques.

3.3.3. Waiting control

It was planned at the outset that the NI group would be a waiting control group, so that if any benefits were found for RT or RTV, then the normal instruction group would receive that intervention in the autumn term. However, the school re-organised the classes over the course of the year which meant that it was no longer possible to fit the intervention into the timetable.

3.4. Results

The results for the NARA reading test for comprehension, accuracy and reading rate are presented first for the three groups. These are followed by the results of the strategy use questionnaire. The NARA results were analysed by means of mixed design ANOVAs, with group (RT, RTV, and NI) as the between variable and time (pre- and post- instruction, Time 1 and Time 2) as the repeated measure. Preliminary inspection of the data showed there were no outliers and the data were normally distributed. Such inspections were carried out for all the tests in the present research. Similarly, an alpha value of .05 was used to establish statistical significance for all the tests in this thesis. A summary of the post-intervention standardised scores on the NARA for the three groups is given in Table 9.

Table 9. Mean standardised scores from the NARA following intervention for the three groups in Study 1 (standard deviations are in parentheses).

	RT	RTV	NI
N	16	17	17
Comprehension	111.25	110.59	101.21
	(9.07)	(7.59)	(11.16)
Accuracy	105.06	107.88	98.94
	(10.14)	(8.18)	(10.79)
Rate	106.38	108.12	109.53
	(10.84)	(10.64)	(10.65)

3.4.1. Comprehension scores

Figure 4 shows the mean standardised NARA scores pre- and post- instruction for reading comprehension for the three groups. The two factor ANOVA showed no significant effect of group, but a significant effect of time ($F(1,47)=21.55$, $p<.001$) and importantly, a significant group by time interaction ($F(2,47)=12.61$, $p<.001$). Comparisons using t-tests revealed that there were no significant differences between the groups at Time 1 (NI vs. RT $t(31)=.89$, $p=.38$, $r=.16$, NI vs. RTV $t(32)=.32$, $p=.75$, $r=.06$, RT vs. RTV $t(31)=.60$ $p=.55$, $r=.11$) but at Time 2 there was a significant difference between the NI group and the RT group ($t(31)=-2.75$, $p=.01$, $r=.44$) and the NI group and the RTV group ($t(32)=-2.78$, $p=.009$, r = .44). The difference between the two RT groups was not significant ($t(31)=.23$, $p=.82$, $r=.04$)[9]. Paired t-tests by group showed a significant increase in scores for both the intervention groups from Time 1 to Time 2 (RT $t(16)=4.36$, $p=.001$, $r=.75$, RTV $t(16)=4.25$, $p=.001$, $r=.73$). For the NI group the difference was not significant ($t(16)=1.64$, $p=.12$), however there was an effect size of .38, with the scores being *lower* at Time 2 for this group.

It should be noted however, that the NARA standardised scores have a ceiling of 12 years 8 months, and the scores of some of the children were at ceiling, both at Time 1 and Time 2. In the NI group, 5 children were at ceiling at Time 1, and 6 at Time 2. In the RT group

---

[9] The analysis was repeated using ANCOVA, with the pre-instruction scores as a covariate, but the results were the same.

scores of 3 children were at ceiling at Time 1 and 10 at Time 2. In the RTV group 4 children were at ceiling at Time 1 and 9 at Time 2. For this reason the analyses were repeated using the NARA raw scores. The effect of group was not significant, but there was a significant effect of time (F(1,47)=39.72, p<.001) and a significant interaction of time and group (F(2,47)=14.19, p<.001). Comparisons using t-tests revealed that there were no significant differences between the groups at Time 1 (NI vs. RT t(31)=1.08, p=.29, r=.19, NI vs. RTV t(32)=.26, p=.79, r=.04, RT vs. RTV t(31)=.88, p=.39, r=.16), but at Time 2 there were significant differences between the intervention groups and the control (NI vs. RT t(31)=2.75, p=.01, r=.44, NI vs. RTV t(32)=2.9, p=.007, r=.46), whilst there was no significant difference between the two intervention groups (RT vs. RTV t(31)=.13, p=.90, r=.02). As in the analysis with the standardised scores, paired t-tests by group showed a significant increase in scores for both of the intervention groups from Time 1 to Time 2, (RT t(15)=5.79, p<.001, r= .83, RTV t(16)=4.82, p<001, r=.77). For the NI group there was no significant difference between Time 1 and Time 2 (NI t(16)=.69, p=.50, r=.17), although the scores were again lower at Time 2.

Figure 4. NARA standardised scores for reading comprehension pre- and post-instruction for the three groups in Study 1.
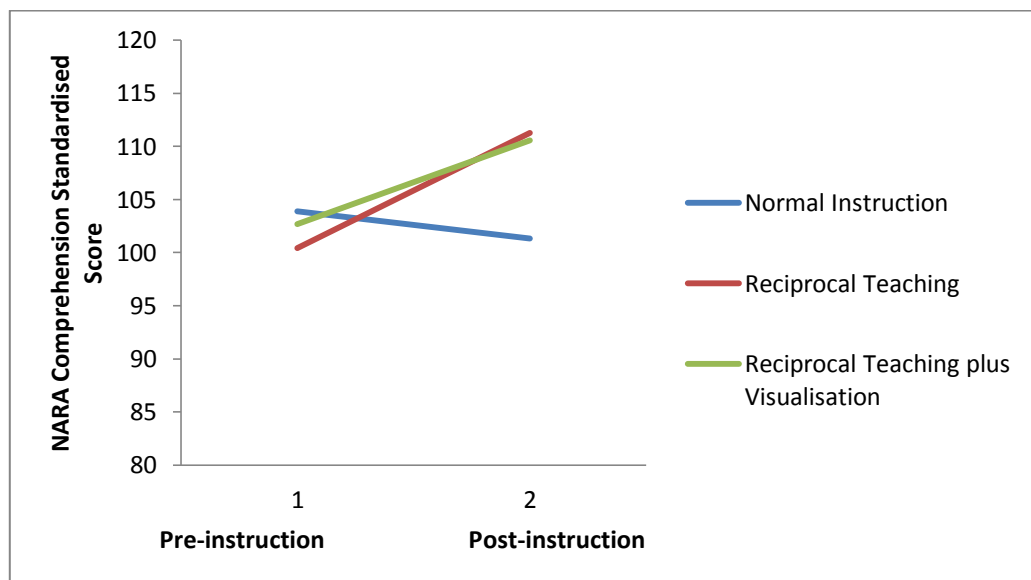
3.4.2. Ratio gain

As was discussed in the literature review, there is another way to measure the effectiveness of an intervention. In the reports of which literacy interventions are effective in UK schools

(Brooks, 2002, 2007, and 2013) Brooks used ratio gain as a means of determining the effectiveness of instruction methods. As outlined in the previous chapter, ratio gain is the difference in reading or spelling age in months from pre-test to post-test divided by the time between pre- and post-test in months. According to Brooks (2002) a gain of 1.4 or more represents a gain that is more than standard progress, and is therefore educationally significant, whilst a ratio gain of 2 or more represents not just satisfactory, but good extra progress. In 2007 and 2013, Brooks refined this scale, such that a gain of between 2 and 3 was considered to be modest, a gain of between 3 and 4 was considered to be substantial and a gain of 4 or above was considered to be remarkable. The ratio gain for the comprehension results on the NARA for the RT group in this study was 5.6, whilst the ratio gain for the RTV group was 4.16. Thus, the improvement for the RT group was higher than for the RTV group, but the difference was not significant ($t(31)=.96$, $p=.34$, $r=.17$) There was a ratio loss for the NI group of .95.[10]

3.4.3. Accuracy

Mean standardised NARA scores for accuracy pre- and post-instruction are shown in Figure 5. The scores for reading accuracy were analysed in the same way as for the comprehension scores.   Analysis with the mixed design ANOVA revealed that there was no significant effect of group, but there was a significant effect of time ($F(1, 47)=11.83$, $p=.001$), and a significant group by time interaction ($F(2,47)=11.41$, $p<.001$). Post-hoc analyses revealed that at Time 1 there were no significant differences between the three groups (NI vs. RT $t(31)=.22$ $p.83$, $r=.04$, NI vs. RTV $t(32)=-.38$ $p=.71$, $r=.07$, RT vs. RTV $t(31)=-.60$, $p=.55$, r=.10). At Time 2 the difference between the NI group and the RT group was not significant
 ($t(31)=-1.68$, $p=.10$ $r = .29$). However, the difference between the NI and RTV group was significant ($t(32)=-2.72$, $p=.01$, $r=.46$). Again, the difference between the two RT groups was not significant ($t(31)=.88$, $p=.39$, $r=.15$)[11]. Paired t-tests by group showed a significant increase in accuracy for both the intervention groups from Time 1 to Time 2
 (RT $t(16)=3.93$, $p=.001$, $r=.70$, RTV $t(16)=3.24$, $p=.005$, $r=.63$). For the NI group there was a significant decrease in accuracy ($t(16)=2.72$, $p=.015$, $r=.63$).

The results for accuracy were not affected by ceiling scores to the same extent as the comprehension scores were, but six children did achieve ceiling scores at Time 1,

---

[10] After removing the children who reached ceiling at Time 1, the ratio gain for the RT group was 7.04, and for the RTV group it was 5.60. For the NI group there was a ratio loss of 1.15.
[11] The analysis was repeated using ANCOVA, with the pre-instruction scores as a covariate, but the results were the same.

(3 children in the NI group, one in the RT group and two children in the RTV group), so the analyses were repeated with the raw scores. The ANOVA revealed that there was a significant effect of time ($F(1,47)=30.09$, $p<.001$) and a significant group by time interaction ($F(2,47)=13.09$, $p<.001$). Post-hoc independent sample t-tests showed there were no significant differences between the groups at Time 1
 (NI vs. RT $t(31)=.47$, $p=.64$, NI vs. RTV $t(32)=.48$, $p=.64$, RT vs. RTV $t(31)=1.01$, $p=.32$). At Time 2, the difference between the NI and the RT group was not significant, but had a small to medium effect size ($t(31)=1.65$, $p=.10$, $r=.28$). The difference between the NI group and the RTV group was significant ($t(32)=2.97$. $p=.006$, $r=.46$) with a medium to large effect size. The difference between the RT groups was not significant ($t(31)=1.26$, $p=.22$, $r=.22$). Paired sample t-tests showed that there was a decrease in the raw scores for accuracy at Time 2 for the NI group ($t (16) =1.34$, $p=.20$, $r=.32$), which although not significant, had a medium effect size. The increases in accuracy for the intervention groups were significant (RT $t(15)=5.28$, $p<.001$, $r=.81$, RTV $t(16)=4.33$, $p=.001$, $r=.73$).

Figure 5. NARA standardised scores for accuracy pre- and post-instruction for the three groups in Study 1.
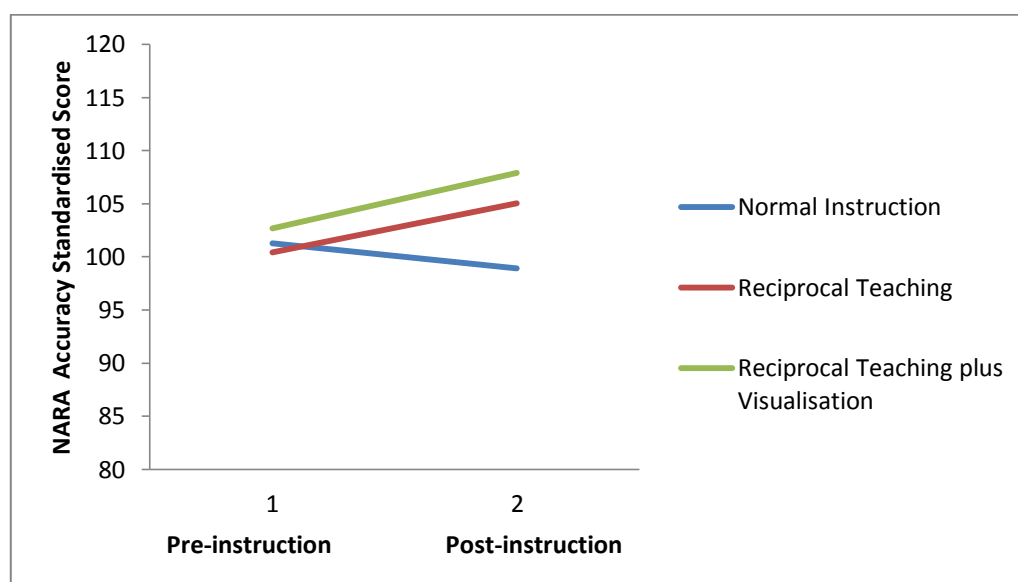
3.4.4. Reading rate

Mean standardised scores for reading rate pre- and post- instruction are shown in Figure 6. Analysis with the mixed design ANOVA revealed that the main effects of group and time were not significant, and neither was the interaction ($F(2,47)=2.26$, $p=.12$). However, as Figure 6 indicates, the reading rate in the two RT groups was slightly slower after instruction

than before, whilst that of the normal instruction group was slightly faster. As the a priori prediction was that the children in the RT groups would slow down in reading rate as they monitored their comprehension more closely, paired t-tests were conducted, which showed a significant decrease in reading rate for the RTV group, with a large effect size ($t(16)=2.77$, $p=.03$, $r=.51$). The decrease in reading rate for the RT group was not significant, but there was a medium effect size, indicating practical significance ($t(15)=1.42$, $p=.18$, $r=.34$). For the NI group, the slight *increase* in reading rate was not statistically significant ($t(16)=-.58$, $p=.57$, $r=.14$).

Again, a number of children achieved ceiling scores at Time 1 (8 children in the NI group, 8 children in the RT group and 9 children in the RTV group). The analyses were repeated using the raw score of words per minute. The ANOVA revealed that the main effects of group and time were not significant, but there was a significant group by time interaction ($F(2,47)=4.58$, $p=.015$). Independent t-tests at Time 1 showed no significant differences in reading rate between the three groups (NI vs. RT $t(31)=.09$, $p=.93$, $r=.05$, NI vs. RTV $t(32)=1.41$, $p=.17$, $r=.24$, RT vs. RTV $t(31)=1.33$, $p=.20$, $r=.23$). The differences at Time 2 were not significant either, (NI vs. RT $t(31)=1.37$, $p=.18$, $r=.24$, NI vs. RTV $t(32)=.97$, $p=.34$, $r=.17$, RT vs. RTV $t(31)=.60$, $p=.56$, $r=.11$). However, the paired sample *t*-tests showed that the NI group were significantly faster at Time 2 than at Time 1, with a large effect size($t(16)=2.63$, $p=.018$, $r=.55$), whilst the difference for the two intervention groups was not significant(RT $t(15)=.58$, $p=.57$, $r=.15$, RTV $t(16)=1.79$, $p=.09$, $r=.41$).

Figure 6. NARA standardised scores for reading rate pre- and post-instruction for the three groups in Study 1.
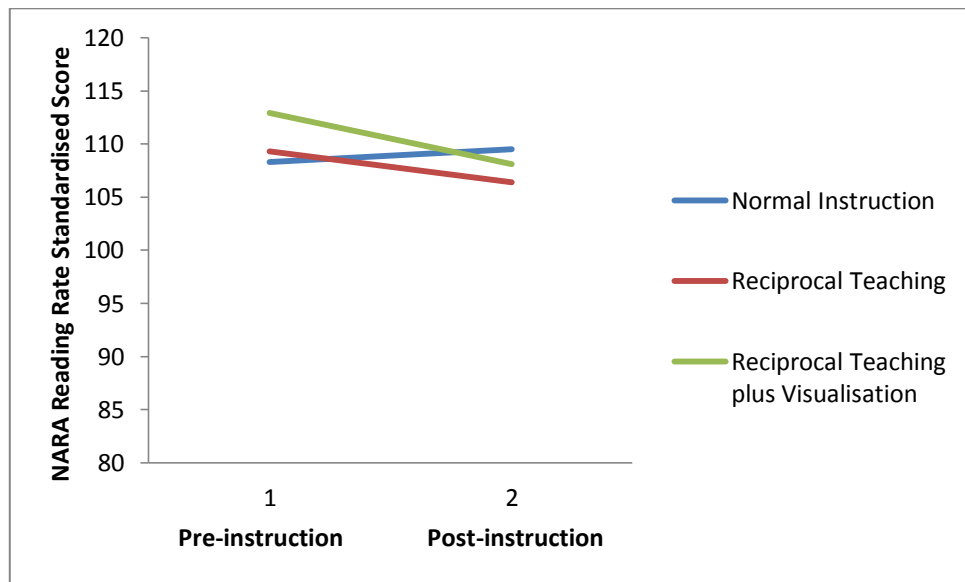
To summarise, the results from the NARA revealed that the two RT groups made significant gains in comprehension and accuracy scores. These gains were even more pronounced when raw scores were analysed since there were ceiling effects for the standardised scores. Finally, reading rate was slightly slower in the RT groups following the instruction period but this was not observed for the normal instruction group, who were reading significantly faster at Time 2 in terms of words per minute raw scores.

3.4.5. The strategy use questionnaire

The number of strategies reported pre-test for the three groups were: NI *M*=3.01, *SD*=.57, RT *M*=2.75, *SD*=.73, and RTV *M*=2.92, *SD*=.80, and for post-intervention the scores were: NI *M*=2.82, *SD*=.62, RT *M*=3.23, *SD*=.71, and RTV *M*=3.23, *SD*=.82. A two way mixed ANOVA with group as the between factor and time as the repeated factor was used to analyse the results. The effect of group was not significant ($F(2,44)$=.18, *p*=.84), but there was a significant effect of time ($F(1.44)$=7.14, *p*=.01) and a significant group by time interaction ($F(2,44)$=5.30, *p*=.009). Comparisons using t-tests showed there were no significant differences between the groups at Time 1 (NI vs. RT ($t(31)$=1.13, *p*=.27. *r*=.20, NI vs. RTV $t(30)$=.38, *p*=.70, *r*=.07, RT vs. RTV $t(31)$=.62, *p*=.54, *r*=.11), but at Time 2 the differences between the NI group and the intervention groups, although still not significant statistically, were more pronounced, with larger effect sizes (NI vs. RT $t(30)$=1.73, *p*=.09, *r*=.30, NI vs. RTV

t(32)=1.65, *p*=.11, *r*=.28). There was no significant difference between the intervention groups (RT vs. RTV *t*(30)=.01, *p*=.99, *r*=.002). Paired t-tests by group showed that the pre-post-test reduction in scores by the NI group was not significant (*t*(15)=1.24, *p*=.24) but there was a medium effect size (*r*=.30). For the intervention groups there was an increase in reported strategy use, which was significant for the RT group (*t*(14)=3.19, *p*=.007, *r*=.65) and which approached significance for the RTV group (*t*(15)=2.06, *p*=.06, *r*=.47).

3.4.5.6. The evaluation questionnaire

An evaluation questionnaire was used at Time 2 (after the intervention finished) to gather information about what the children thought they had learned and whether they had enjoyed the instruction. The questionnaire utilised a combination of Likert scales and open-ended questions. Group statistics for the responses are given in Figures 7 and 8

Figure 7. Evaluation questionnaire responses for whether children learned anything from the lessons for Time 2 in Study 1.
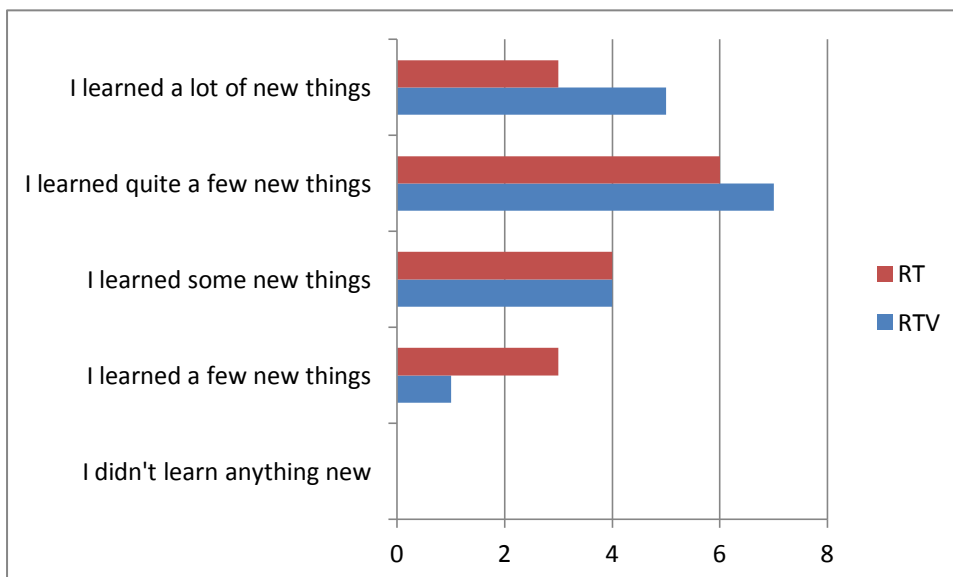
Figure 8. Evaluation questionnaire responses for whether children enjoyed the lessons for Time 2 in Study 1.
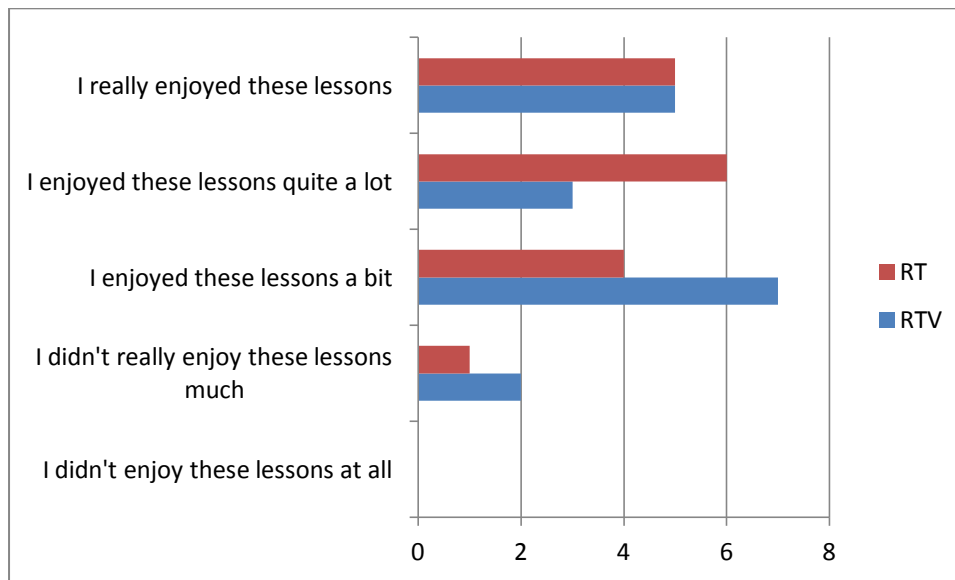
Across the two intervention groups, some 90% of the children reported that the instruction had helped them to understand more about reading. In answer to the open-ended questions about what they had learned, two children in the RT group mentioned the names of characters, and six children mentioned particular strategies, with clarifying being mentioned by four and predicting, summarising and questioning by two of these. More general comments in the what did you learn section pointed to the instruction being "fun" and "interesting" and that children had learned "how to read better", and one child said it had "helped me understand better, because before there was lots of stuff I didn't understand". The responses from the RTV group were very similar, with five children mentioning particular strategies- four mentioned clarifying, two summarising, and one named all four. One child said "I have learnt not just to read the book but to use the strategies". An additional four children talked about visualising or making pictures in their heads. One child had learned "how much more interesting reading can be" and another "that reading can be fun". One child said they had "learned to read bigger books" and another felt that they had learned to read better because they had slowed down.

In terms of answers to the open-ended questions about enjoyment, eight children in the RTV group and 6 children in the RT group mentioned that they had enjoyed the book they read. In the RT group, one child enjoyed "everything", one child liked being the teacher, two children liked working in a group and two children enjoyed "listening to what other

people have to say" and "listening to other people's thoughts". In the RTV group, one child again said they liked "everything", one enjoyed being the teacher and another enjoyed "being different roles". The RT group did not make any comments about things they did not enjoy, but in the RTV group there were three children who made comments; one child did not like summarising, one did not like visualising and one child did not like going in the goldfish bowl.

3.5. Discussion

Study 1 set out to investigate the effectiveness of RT as an intervention for improving reading comprehension in a whole class setting. Three Year 5 classes took part; two were taught for an hour a week by the researcher, and one class acted as a control group. One class received traditional RT, with instruction and practise in the four strategies of predicting, clarifying, questioning and summarising, whilst the second class received the RT plus instruction in the use of an additional strategy - visualisation. Pre- and post- instruction assessments of reading comprehension indicated that RT was significantly more effective than the usual practice of comprehension teaching, and that this may have been achieved by increasing strategy use and the active processing of text, as reflected in increased strategy use and slower reading rates in the RT groups at the end of the study (there being a large effect size for the pre- post-test decrease in reading rate for the RTV group and a medium effect for the RT group)[12]. Both intervention groups showed a significant increase in accuracy, both in terms of the standardised scores and raw scores. The hypothesis that visualisation would increase the effectiveness of RT was not supported, since there was no difference in comprehension scores for the two RT groups following intervention.

With regard to using RT in a whole class setting, although the class sizes in the school in the present study were small, the researcher conducted the instruction without the assistance of any support staff. Group work during the teaching sessions involved four groups. In a larger class it would be possible to have more groups, and with the support of a teaching assistant, eight groups would have the same amount of supervision as was achieved in the present study. No additional resources were needed, to carry out the instruction, beyond copies of the novel used.

---

[12] There were problems with ceiling scores, but when the children who had ceiling scores were omitted from the analysis the NI group were found to be reading significantly faster at Time 2, whilst there was no significant difference for the intervention groups.

3.5.1. Comprehension measures

Previous studies have often shown an improvement in scores on experimenter-designed comprehension tests, but not always on standardised tests. The results of the present study did reveal an improvement on a standardised test (the NARA). This may be because of the way RT was implemented in this study, or it may have to do with the way comprehension was assessed. Here, RT involved mainly spoken interactions, with little written work taking place. Assessments were only administered at the outset of the instruction and at the end of the study. Several of the previous RT studies have involved written comprehension passages attempted on a daily basis as part of the assessment process (Brady, 1990; Dermody, 1988; Lysynchuk, Pressley, & Vye, 1990; Palincsar & Brown, 1984; Palincsar, et al., 1987). This allows practise of an aspect of the assessment that was not a feature of the present study.

An alternative explanation for the difference in results between this and certain previous studies could be that the standardised test used in this study, the NARA, involves questions which are presented and answered orally. In previous studies the most frequently used test, according to Rosenshine and Meister's review, was the Gates-MacGinitie (Brady, 1990; Labercane & Battle, 1987; Lysynchuk, et al., 1990; Taylor & Frye, 1992). This test comprises multiple-choice questions, which avoids confounding assessment of reading comprehension with writing ability (Jenkins, et al., 2004), but does still require answering written questions as opposed to questions presented verbally by the tester. Rosenshine and Meister did compare the passages used in the Gates-MacGinitie test with the passages used in experimenter-designed passages, in an effort to understand why improvement was observed with the latter but not the former. They concluded that the experimenter-designed passages were more like the passages used for teaching, whereas the standardised test passages came from a wider range of material. Additionally, experimenter-designed test passages were organised in a repetitive, topic sentence and supporting detail format, and required less background knowledge and less searching of the text to answer than the standardised test passages. All these features were found in the teaching passages. In the present study, the text studied was narrative in form, in common with the NARA passages. Thus, use of an assessment which relates most closely to the way RT is delivered in the study may be the best means for measuring improvement.

Research has also shown how different the various measures of reading can be, and thus how difficult it is to compare the results of interventions when different assessments have been used (Cain, 1999; Cutting & Scarborough, 2006). Rimrodt, Lightman, Roberts,

Denckla and Cutting (2005) found that when three different measures of comprehension were used, only about 25% of the sample of children in their study, identified as having a comprehension deficit, were identified as such by all the tests. About 50% of the children were identified by only one of the tests. Some tests rely heavily on decoding rather than comprehension. The NARA has an additional comprehension component beyond that explained by word reading skill alone, since unknown words are supplied by the tester.

3.5.2. Metacognitive strategies

Improving children's metacognitive strategy use has been shown to be an effective method of improving children's reading (NICHD, 2000) and the importance of its role in comprehension instruction has been widely acknowledged (Baker, 2002; Trabasso & Bouchard, 2002). However, the issue of ascertaining the metacognitive strategies being used by children (or indeed adults) is not straightforward. We can measure the input (the text) and the result (comprehension can be measured, albeit imperfectly, through the use of various assessments) but understanding what happens in the middle is fraught with difficulty. It has been argued that to understand the effect that RT has on metacognition is vital to understanding how it works, and how it can be most efficiently implemented (Demmrich, 2005).

Both treatment groups in the present study reported an increase in strategy use over the instruction period, with the increase being significant for the RT group, and approaching significance for the RTV group. The lower rate of increase shown by the RTV group may be because the visualisation instruction took away teaching time from instruction or practice in the RT strategies. However, if additional time had been allocated to teach visualisation then comparing the types of instruction would have been problematic. One alternative would be to teach visualisation as well as each of the four RT strategies to separate groups of children. This would enable us to begin to untangle which of the strategies is most effective. This was a direction identified by Baker (2002), who felt it important "because multiple-strategies instruction is challenging to implement and labour-intensive on the part of teachers and students alike" (p.76). However, it may be the combination of strategies that is effective rather than one in particular. Indeed, much of the research on individual strategies would seem to confirm this. Intervention with multiple strategies has been shown to be more successful than with single strategy interventions. For example, the National Reading Panel report (National Institute of Child Health and Human Development, 2000) found that question generation alone had a positive effect on

111

standardised tests in only 3 out of 13 studies, whilst RT, although having a larger effect size for results from the experimenter-designed tests (.88) still had a significant effect, with an effect size of .32, in the remaining nine studies using standardised tests. The National Reading Panel concluded that the evidence supports the use of combinations of strategies. Furthermore, it may not be the strategies themselves which are important, rather the active engagement with the text that they foster. This was suggested by Chan and Cole (1986) when their study of three experimental conditions (question-generating, underlining and explaining interesting words and questioning and underlining combined) all produced similar improvements. Brady (1990) put forward a similar view, and in a personal communication cited in Rosenshine and Meister (1994 p. 510) Brady suggests that the continual emphasis on strategies "forced students to move beyond a belief that decoding the words was a sufficient response to the request that they "read" a passage". More recently Cain (2010) reiterated the view that training in specific skills may be effective as it encourages a more general engagement with the process of constructing meaning. This was indeed made clear in an early assessment by the programme's originators (Palincsar & Brown, 1986):

> In RT the acquisition of the strategies is not the ultimate goal of instruction. The strategies are but a means to an end, they provide the vehicle for teaching students to read for meaning and to monitor their reading to ensure that they understand. (p. 776).

What could be more important than trying to work out which strategies are most effective, may be to approach the question from the other direction. That is, if a particular combination of strategies has been shown to be effective, can we work out what the effects are on children's reading processes? Siegler (2006) maintains that "the only way to find out how children learn is to study them closely while they are learning" (p.469). A potentially informative line of research could be to examine children's reading in more detail before, during, and after a RT instruction period and over a more extended time scale than that of the present study. This was the objective of Study 2.

3.5.3. Assessment of strategy use

The use of self-report questionnaires to measure learning strategies is common ((Van Hout-Wolters, 2009) (translation provided by the author) but they have been criticised for various reasons. Firstly, they may be measuring the learner's perception of their learning activities, instead of the activities themselves (Perry & Winne, 2006). Secondly, learners can mention

learning activities that did not take place and which they would not have thought of without their presentation. This may be because they think such an answer is socially desirable (Veenman & Spaans, 2005), or because they do not understand what is being asked. Thirdly, it is not clear to whom the learner is comparing him/herself, when answering questions about the frequency with which they use a particular strategy (Veenman, Prins, & Verheij, 2003). Finally, particularly with young children, there is the possibility that they do not understand what they are being asked to do, or that they do not understand a particular question. Certainly in Study 1, where the adapted MARSI questionnaire was used, there were several instances of children reporting that they frequently used a strategy which had been included as a foil. For example, two children reported that they frequently used the strategy of counting the number of verbs in a passage before they started reading. This would cast doubt on the reliability of the answers given in such questionnaires.

Veenman and Spaans (2005) identified three measures of metacognition – prospective, concurrent and retrospective. The MARSI is a prospective measure, in that it asks children about their reading in general. To find out about how strategies are used in practice, a concurrent measure, such as the think-aloud protocol, could be effective. As the name implies, think-aloud measures require the reader to pause at given intervals and state what it is they have been thinking about whilst reading. Although think-alouds have been used to study problem-solving and adult reading processes, as reported by Pressley and Hilden (2004), fewer studies have used this method to look at reading in the early stages of development. Think-alouds are a part of the modelling process in RT, but they have not been used to assess how children might alter the way they process text in the course of a period of RT instruction. Since research has shown that the best way to measure metacognition is through multiple measures (Veenman & Spaans, 2005) then the information provided by the think-alouds could be combined with a semi-structured interview to provide a more complete picture of what the child is doing to comprehend text.

A second alternative concurrent measure is reading rate. In the present study the results indicated a slowing of reading rate over the instruction period in the RT groups, which, whilst only being statistically significant for the RTV group, also had a medium effect size for the RT group. This may have been due to the newly introduced monitoring strategy taking up limited processing resources. The possibility that this may occur during RT instruction was suggested by Collins & Smith (1982) and by Palincsar nearly 25 years later (2006) but it has not as yet been thoroughly investigated.

113

Whilst the reading rate of the intervention groups appeared to slow down, in the normal instruction group reading rate increased significantly when the raw scores were analysed. This may be related to the appearance of a slight decline in comprehension and accuracy in the normal instruction group. Without the focus on active engagement and processing of text that the RT groups received, the normal instruction group continued to read quickly. Indeed their speed of reading may possibly have come at the expense of accuracy and understanding. The NARA test does not penalise readers for self-correction, and so self-corrections were not recorded, but self-correcting errors will take up more reading time and show up as a decline in reading rate. This possibility was followed up in the next study by examining for a change in the rate of self-corrections over the course of RT instruction

3.5.4. The role of motivation in increasing reading comprehension scores

Research has shown that children with a positive attitude towards reading will be motivated to read more (Baker & Wigfield, 1999) and to achieve at a higher level (Broussard & Garrison, 2004). Internal motivation specifically, has been shown to correlate with reading comprehension and strategy use (Pintrich, 2000). The evaluation questionnaires completed post-instruction indicated that the RT had a motivational effect. As we have seen in the results section, some 90% of children reported that the lessons had helped them to improve their understanding of what they read to some extent. Comments that reflected enjoyment could be grouped into those reflecting the use of particular RT strategies, such as "I enjoyed making predictions about what would happen next" or "I enjoyed being the clarifier", and those which reflected the method of RT, such as "I enjoyed being the teacher" and "I enjoyed reading the text and listening to everybody's thoughts". There were also comments which reflected enjoyment more generally, such as "I have learned how interesting books can be", and "I learned reading can be fun". Additionally there were comments about what had been learned, such as "I have learned to predict and clarify", and "I have learnt to summarise". Comments were also made about reading rate, for example, "I have learnt to slow down so that I understand more." The comment about reading bigger books relates to the research mentioned in the introduction (Stanovich, 1991 as quoted by Cain and Oakhill 2011) whereby weaker readers only get to read easier books. This child obviously appreciated the chance to read "a bigger book" than those to which they were accustomed.

In terms of the response to visualisation, many children reported making pictures in their minds as they read (53% of the children reported prior to instruction, that they "always

or almost always" did so, the same percentage as at post-test) but we cannot know if they actually do this, for example, as they have been told it is a good strategy. Think-aloud protocols would identify children who spontaneously use this strategy, and follow-up interviews would enable a closer examination of how children employ imagery as they read. Both of these measures were included in Study 2.

3.5.5. Limitations of the research

One limitation of this study is that the NARA assessments at both Time 1 and Time 2 were carried out by the researcher, who was not blind to the grouping of the participating children. This was a result of resource and teaching constraints. In light of this it was important to ensure that the results were consistent with those obtained with other assessments. A comparison of the NARA results at Time 2 with practice SATs reading assessments, carried out by the class teachers, revealed no significant differences between the two sets of scores (NARA M=107.62 SD=10.46. SATs M=106.76 SD=9.98, $t(49)=1.22$, $p>.05$).

A further limitation of the present study is that the outcome may have been subject to a Hawthorne effect (Landsberger, 1958; Mayo, 1933, 1949). However, if this is taken to mean that the improvement in comprehension in the RT groups may have been attributable to the presence of a novel teacher and the extra attention paid to these groups, then this can be countered. Firstly, the researcher was not unknown to the groups, being already present in the school as a special needs teacher and having taught as a class teacher in the school previously. All three groups had been taught by the researcher as a supply teacher on various occasions over the preceding two years. Therefore, the researcher was not an unfamiliar adult for the children. In terms of extra attention, the two Year 5 teachers of the RT groups taught the normal instruction group for three literacy lessons each over the period of the instruction, so that this group also received literacy teaching that was out of the normal routine.

3.5.6. Follow-up

Following the advice given by Rosenshine and Meister (1994) it was planned to include follow-up assessments in Year 6 to see if the improvements in reading comprehension scores and strategy use shown by the RT groups had been maintained. However, this was not possible as the school did not give permission for the testing to be carried out.

3.5.7. Conclusion

The results of Study 1 indicated that RT instruction resulted in a significant improvement in reading comprehension in this whole-class instruction study. Visualisation did not appear to add to the effectiveness of the RT instruction, in terms of increase in reading comprehension scores, over the timescale of the present study at least

Instruction in comprehension strategies has been shown to have a positive effect on comprehension, in a time when such strategies are not being widely taught. Although the present study was small in scale and replication was necessary, the results indicated that RT can be effective in a UK classroom when used for the whole class taking part in collaborative learning in small groups. It has not been shown before that this approach can be used effectively for whole classes.

3.5.7.1. Implications for the next study

The results of the first study indicated that further research was needed to explore the potential benefits of visualisation, whilst a qualitative observational study might enable us to see more clearly what is happening when children are receiving RT instruction. Sensitive measures might include think-alouds to find out more about on-line processing, and interviews to find out more about strategy knowledge and use. The second study therefore investigated these claims in more detail by examining strategy use through think-alouds and interviews. Results of Study 1 had indicated that RT was effective, but we still needed to understand more about why and how it works. It was considered important to reflect on how comprehension could be measured. To measure what has been understood after reading is to be interested in the product of the reading process, and it is this which has so often been investigated in reading tests (Rapp, Van den Broek, McMaster, Kendeou, & Espin, 2007). However, reading is a process and to try to understand what is happening during reading we need to look towards alternative measures which aim to uncover the concurrent processes. Duke and Carlisle (2011) see this as a "fundamental challenge" facing comprehension assessment today (p.219). The second study addressed this by looking at the processes involved as children took part in a RT programme. In an article about effective interventions (Duff & Clarke, 2011) claimed that future research "should aim to clarify the factors that influence response to intervention" (p.7). This was the aim in Studies 2 and 3, to study the reading processes themselves by using the richer data provided by think-alouds.

The think-aloud procedure has been primarily used for adult readers (Pressley & Afflerbach, 1995) but fewer studies have examined the developing nature of reading in emerging readers (Pressley & Hilden, 2004). The second study provided data on how young children think as they read and its use in this way has not been widely researched (Pressley & Hilden, 2004). Caldwell and Leslie (2010) have looked at assessing comprehension through think-alouds with middle school students (aged 12 to 14) but their use with younger children has not been explored.

It was suggested that reading rate may have decreased for the RT groups in Study 1, possibly due to an increase in self-corrections during reading. In the second study, reading rate and self-corrections were recorded, so that this hypothesis could be examined further. The second study also investigated any lasting effects of RT by including a follow-up assessment one year after the intervention concluded, and addressed the possible criticism that a researcher (albeit a researcher who was also a teacher within the school at the time) rather than a teacher had administered the instruction (Duff & Clarke, 2011).

Chapter 4: Study 2 - Investigating strategy use during Reciprocal Teaching with children aged 7 to 8

"It made me realise what the real thing is about reading. You need to read between the lines."

4.1. Research aims

The aim of the second study was to investigate changes which might take place in reading processes. Think-alouds can be used to assess how children may alter the way they process text over the course of RT instruction. Given Siegler's (2006) exhortation that the only way to find out how children learn is to study them closely while they are learning, the think-alouds might help us to understand what makes RT effective if we examined children's reading in more detail before, during and after an RT instruction period and over an extended time scale. The intervention was implemented in a whole class situation, with children aged 7 to 8, in the UK. A supplementary aim was to see whether the addition of visualisation might increase the effectiveness of traditional RT. The study also investigated the delivery of RT by a class teacher rather than the researcher, and included follow-up assessments to see if any improvements in reading comprehension were maintained.

4.2. Methodology

The study involved examination of changes in strategy use/reading processes as a result of taking part in RT intervention with a group of Year 3 children using think-alouds and reading interviews. It took place in the academic year 2009-2010, with a follow-up assessment at the end of the academic year in 2011. The children's word reading ability and reading comprehension were assessed using standardised measures. Think-alouds and a strategy use interview were conducted with each child once before the RT intervention began and then at two time points during the intervention: ten weeks into the intervention and then after a further ten weeks of training plus visualisation. In order to examine whether any change in reading strategies/processes might be maintained over time the follow-up assessment was administered one academic year after the training finished. Four children were selected as case studies to enable a more detailed examination of any changes.

4.2.1. Participants

The participants were a Year 3 class (aged 7 to 8 years) from the same school as the children in Study 1 – a home counties UK primary school where the researcher worked as a special needs teacher. As noted in chapter 3, the school is an independent junior school, which follows the UK National Curriculum. Owing to its charitable status, the school has a more socially diverse intake than the average independent school. The school is non-selective in an area of highly selective schools with the result that the intake is also more diverse in terms of ability. The class was chosen as the class teacher had been in Year 5 at the time of Study 1, and had already tried some of the techniques involved when she moved to Year 3 the following year. She was keen to implement RT in a more systematic way and to examine the effects of strategy use.

Owing to an error in admissions during a change of staff some years previously, the Year 3 group was abnormally small. There were 24 children in all, divided equally between two adjoining classrooms, with two teachers and a shared teaching assistant. The classes had been divided to be of equal ability, as far as possible. CAT scores were not available for this cohort, but there was no reason to believe that the class differed in ability levels to other classes. Both classes received the instruction together, as one group, but the Head Teacher only gave permission for one class to undergo testing.

4.2.2. Ethical issues

As the study involved participants under the age of 18, a CRB check and ethical approval for the study were required and obtained. Guidelines from the British Psychological Society were followed and ethical approval was obtained from the Department of Psychology and Human Development, Institute of Education, London. Letters describing the study were sent out to the parents/carers of all children in the class and the opportunity was given for any parent/carer to opt out of the study. A copy of the letter can be found in Appendix F. Permission was not obtained from the parents of one child, so she was not tested, but she did receive the same RT instruction as the rest of the class, as it formed part of the curriculum. Data was anonymised, and all computer records referred to numbers. The key to the names was kept in a locked drawer in the researcher's office.

4.2.3. How the intervention was introduced

The researcher was introduced by the teacher a couple of weeks after the start of the autumn term, so that the class had had a chance to settle into their new routines. The children were familiar with the teacher as they had seen her in the school the previous year as she worked one-to-one with children with Special Needs. However, she had not taught any of the children in the class. The researcher explained that she was a scientist from the University in London and that scientists were interested in learning about things. In this instance she was interested in what was happening inside children's heads when they were reading. As it is not possible to lift off the top of children's heads and look inside, she was going to need their help. She explained that she was going to be taking children out of their lessons one at a time to do some reading and to give them a chance to tell her what was happening when they were reading. When she had seen everyone two or three times they would be working on a project about reading with their class teacher for the rest of the school year and she would be there to watch what was going on and to ask questions about what the children had learned at the end.

4.2.4. Pre-instruction measures

As in Study 1, assessments of reading ability and of reading comprehension were administered to the children prior to instruction. The tests of reading employed were the York Assessment of Reading for Comprehension (YARC Snowling et al., 2009) and the Test of Word Reading Efficiency (TOWRE Torgesen, et al., 1999). The YARC assesses reading comprehension whilst the TOWRE assesses decoding and word reading scores. Thus, standardised scores for word reading, phonemic decoding, and total word reading efficiency were obtained at Time 1, as well as reading comprehension , accuracy and reading rate. The data are presented in Table 10. The scores show that the children as a class had good word reading scores, but that there was wide variation. Even the poorest reader however, fell within the average range according to the standardisation norms. These results are very similar to those obtained from the children in Study 1.

4.2.4.1. The qualitative measures

Before the instruction commenced, but after the standardised assessments had been completed, the qualitative measures introduced for this study, were carried out. The Qualitative Reading Inventory (QRI Leslie & Caldwell, 1995) and the strategy interview (MPIR

Keene & Goudvis, 1995) were used to obtain think-aloud and strategy use data respectively. These measures are described in full in the measures section which follows.

Table 10. Mean standardised scores for the YARC and TOWRE at Time 1 for Study 2 (standard deviations are in parentheses).

	Mean
YARC reading comprehension	101.30
	(7.80)
YARC accuracy	104.90
	(8.87)
YARC reading rate	106.70
	(11.35)
TOWRE sight word reading	110.50
	(8.90)
TOWRE phonemic decoding	105.60
	(9.12)
TOWRE total word reading efficiency	109.50
	(10.55)

4.2.5. Post-instruction measures

The two standardised reading tests were administered before the intervention (Time 1), after ten weeks of RT instruction (Time 2), after an additional ten weeks of instruction of RT plus visualisation (Time 3) and at a one-year follow-up (Time 4). At the same time points, The Qualitative Reading Inventory (QRI) and the strategy interview (MPIR) were used to obtain the think-aloud and strategy use data.

As in Study 1, a questionnaire (see Appendix D) was used immediately after the intervention finished (Time 3) to find out what children thought about the intervention -

whether they had had enjoyed RT and whether they had learned anything from it. A questionnaire was also given to the class teacher to enable her to evaluate the intervention.

4.2.6. Materials

4.2.6.1. The Test of Word Reading Efficiency (TOWRE)

The TOWRE was used, as in Study 1, to assess reading accuracy and reading fluency for words and non-words. Children are given 45 seconds to read as many items as possible from a list of words and then a list of nonwords. Standardised scores were obtained for the sight word reading and phonemic decoding measures.

4.2.6.2. The York Assessment of Reading for Comprehension (YARC)

The YARC was not available at the time that Study 1 was carried out. It was published when Study 2 was being planned. The format of the YARC is similar to that of the NARA which was used in Study 1; however, it is considered to incorporate significant improvements over previous assessments of reading comprehension (Ricketts, n.d.). It assesses each child using a fiction and a non-fiction text, and it provides a Single Word Reading Test (SWRT) to enable a starting point to be determined reflecting the child's word reading ability, rather than relying on their chronological age. The NARA has been criticised for its limited content validity (Hurry & Doctor, 2007), in that 65% of the comprehension questions test literal comprehension and the remaining questions only address simple inference. The YARC was designed to include both literal and inferential questions, but the inferential questions address knowledge-based inferences and emotional inferences and the later passages sample more abstract inference types. Furthermore, it was designed to overcome the issue associated with the NARA that children can often answer comprehension questions on the basis of background knowledge.

4.2.6.3. The Qualitative Reading Inventory (QRI)

The QRI is an individually administered reading inventory, for assessing the reading ability of students from pre-primer (aged 4-5 years) through to high school (aged 14-16 years). It was used in the present study to obtain the think-aloud data. It is designed to give information about an individual's reading level and to provide diagnostic information. In addition to providing an estimate of the reading level for each child, it provides a tool for observing the reading behaviour of a child in a relatively natural context that approximates a real reading situation. It is an instrument that has been used by literacy researchers to evaluate progress,

and several studies have suggested that the passages are sensitive to immediate and long-term change (Alison & Paris, 2003; Catts & Adolf, 2006; Johnson-Glenberg, 2000).

Informal Reading Inventories such as the QRI are designed to evaluate different aspects of reading performance. They consist of graded word lists and reading passages, which are read either aloud or silently, according to the type of assessment required, and then oral responses are given to assess comprehension and recall. They are useful in determining reading levels and they also have a diagnostic function; through miscue analysis it is possible to see how a beginning reader is approaching decoding, whilst think-alouds enable an assessment of the range of strategies a more advanced reader is using. In a review of Informal Reading Inventories published since 2002, Nilsson (2008) evaluated eight such instruments, including the QRI-4 (Leslie & Caldwell, 2006), an earlier edition of the assessment used in the present research. As narrative and expository texts are structured in different ways, Nilsson considers it to be important that the inventory contains examples of both, and that the expository texts be structured in the way that they would be in a science text book, rather than in the same way as a narrative. The QRI is singled out as best fulfilling these criteria. It also includes comprehension questions which are suitable for expository text and these assess literal and inferential comprehension. The QRI also includes a measure of prior knowledge of concepts in the passages. In terms of reliability, the QRI is the only one of the eight reviewed assessments to include information about the degree of generalisability across alternate forms. The QRI was thus chosen over other Informal Reading Inventories for its validity and reliability, as well as the prior knowledge component.

Previous research on the thought processes involved in reading has largely focused on what happens when comprehension breaks down, and passages have been utilised which contain deliberate inconsistencies, or ambiguities (Winograd & Johnston, 1982). Little research has focused on what happens during normal reading (Hacker, 2004). The QRI passages reflect 'normal reading', and therefore allow for examination of the processes which take place when reading the kind of texts children normally encounter.

The QRI contains several passages for each level, making it possible to test children on a different passage even if they have not progressed to the next level. Indeed, as the passage levels equate to grade levels, then it would be expected that children tested at three time points in the same academic year would need to read more than one passage at the same level. Passages are provided of both expository and narrative text. Only the

passages of expository text were used in the present study as they were less culturally specific, and more suited to use in the UK.

Once a passage has been selected, the QRI begins with the tester asking the child several concept questions that determine their previous knowledge for the topic of that passage. A lack of knowledge may explain difficulties in comprehension, and data from a pilot study for the QRI showed that prior knowledge predicted passage comprehension more frequently than a general measure of reading achievement, with the correlations stronger and increasing from level three onwards (QRI-5 pp 455-456).

The authors of the QRI also provide guidance on miscue analysis, which can be carried out at the same time. While the pupil reads aloud the tester records errors (any deviation from the printed text) which can be analysed and used to determine if the pupil is primarily focused on reading words or deriving meaning from the text.

4.2.4.4. The Strategy Interview

The strategy interview was adapted from the Major Point Interview for Readers (MPIR) developed by Keene (1995). As the aim of the intervention was to make children more active readers by improving their strategy use, an assessment of strategy use was required. In Study 1 a self-report questionnaire had been used, but as discussed in chapter 3, this was not entirely satisfactory. Firstly, some children appeared to just tick 'yes' for everything, and it was not clear if they did indeed use the strategies they claimed to use. Secondly, a questionnaire is an off-line measure, and therefore not as reliable as an online measure, since it relies on the child remembering what they have done, which could lead to forgetting and/or elaboration. Some researchers have measured strategy use in a more targeted way, and tried to make the assessment more online by asking children to demonstrate their strategy use. For example in a study of explicit teaching in strategy instruction in fifth grade (age 10 to 11 years) in Norway (Andreassen & Braten, 2011) children were given a title and a text. First, the children were asked to predict the content of the passage from the title, then after reading the passage they were asked to clarify two words, compose two questions and then write a summary. However, as the authors acknowledge in the discussion section of their article, this is more a measure of strategy competence than of awareness or self-initiation. The children were tested on their procedural knowledge only. For this study, an online measure of strategy use, the think-aloud, was supplemented by a strategy interview, which although off-line was targeted at a particular passage rather than being about reading

behaviour in general. It was thought that the two measures together would give a fuller picture of which strategies a child was using spontaneously, and how well they could use them. Together they provided opportunities to test declarative, procedural and conditional knowledge.

The MPIR was developed as a formal and consistent way of assessing children's use of comprehension strategies. It consists of the following sections:

1. Using schema

2. Inferring/predicting

3. Asking questions

4. Determining what is important

5. Monitoring comprehension

6. Visualising

7. Synthesising

Some adaptations made were to ensure the suitability of the questions for expository rather than narrative text, but the MPIR was designed to be used for either genre. A copy of the adapted form of the MPIR is included in the appendix (Appendix G). A scoring rubric is also included for use with the MPIR, which was designed to allow teachers to quantify a child's growth in the use of each strategy. The scoring rubric is given in the appendix (Appendix H).

4.2.4.5. The evaluation questionnaires

The evaluation questionnaire for the children was the same as that used in Study 1 and can be found at Appendix D. The questionnaire was administered at Time 3, immediately at the end of the RT intervention. The teacher's questionnaire, administered at the same time, consisted of seven questions, asking about any expectations at the beginning of the year, what went well in the intervention, and what was more challenging. The teacher was also asked whether they would use reciprocal teaching again, if they would do anything differently another time and if they had learnt anything about teaching reading comprehension from its implementation.

4.3. Procedure

The details of administration for the assessments are described below. All the assessments were carried out in a quiet alcove, outside the classroom, usually used by teaching assistants when listening to children read. A recent report on research in psychology (Funder et al., 2013) has highlighted the importance of providing the verbatim wording of all task instructions to aid further replication attempts. This is particularly important where measures have been modified, as in the case of the Qualitative Reading Inventory in the present study. An outline of the administration of the measures is given in each section below, with additional detail provided in the appendices.

4.3.1. Word reading and phonemic decoding - TOWRE

The TOWRE assesses reading accuracy and fluency for words and pronounceable non-words. The test was administered according to the instructions in the manual. The TOWRE provides parallel forms of each sub-test, which are equivalent in difficulty. Form A was used at Time 1, Form B at Time 2, Form A at Time 3 and Form B at Time 4.

4.3.2. Reading comprehension - YARC

The format of the YARC is very similar to that of the NARA in that each child is required to read two passages aloud, and to answer eight open-ended questions after each passage (if they achieve a pre-determined level of accuracy). As in the NARA, if the child cannot read a word or makes a mistake, the correct word is supplied by the tester.

The testing session was tape-recorded and each child was informed at the beginning that they were being taped but that they were just to pretend the recorder was not there. The recorder was a digital one and very unobtrusive. Each child was first administered the Single Word Reading Test (SWRT) in accordance with the YARC manual instructions. The YARC provides parallel forms to facilitate re-testing, Form A was used at Time 1, Form B at Time 2, Form A at Time 3 and Form B at Time 4.

4.3.3. Strategy use – the think-aloud and strategy interview

In the next section an explanation is given of how the passage level was chosen for the think-aloud and strategy interview, and how they were administered.

Determining the starting point for the QRI passage

The level of passage to be read for the think-aloud and strategy interview was determined by word lists, which are supplied with the QRI. The wording for the administration of the word list is given in Appendix I.

Think-aloud procedure

After the word-lists had been administered a suitable passage for the think-aloud and subsequent strategy interview was selected, and presented on a different day. During the reading of the passage children were asked to stop at specified points and tell the researcher what they were thinking about as they were reading. The think-aloud forms part of the QRI, but only from Grade 6 onwards. The QRI procedure was extended to the lower levels to be used with younger children in this instance. The wording for the administration of the think-aloud is given Appendix J. The stop points were indicated on the text by pencil lines and generally occurred at the end of a paragraph. Three stop points were used for the shorter passages (197 words and below) and four stop points for the longer passages at levels 4 and up. If the child continued to read past a stop point they were asked to stop. At each stop point the question, "Can you tell me what you were thinking?" was posed. After the child's initial comments they were prompted with "Anything else?" Thinking time was limited to two seconds however, to ensure the child was reporting what they had been thinking and not elaborating.

If the child's first response was "No", or "I'm not thinking of anything", then the child was prompted once with "What do you think it's about?" in order to encourage a response, in case a child was anxious or reluctant to say what they were thinking. As the think-aloud was to be coded and counts of comments were important, it was necessary to use the prompts exactly as prepared for each child.

Previous studies using think-alouds (e.g. Schellings, Aarnoutse, & van Leeuwe, 2006) have used extensive practice passages before the think-aloud procedure was carried out, but it was considered that practising by modelling would suggest strategies the child might use or seem to be recommending a certain approach, whilst this study aimed to find out more about which strategies the children were applying without prompting. Indeed, previous researchers (e.g., Afflerbach & Johnston, 1984) have pointed out that a pervading issue in determining strategy use is the question of whether strategy measurement is of strategies that are induced or naturally occurring. By not modelling strategy use, it was hoped that the strategies revealed in the think-alouds would be naturally occurring.

Miscue Analysis

During the reading of the QRI passage an oral reading miscue analysis is undertaken, with the researcher noting any departures from the printed text, e.g. whole-word substitutions (such as "tried" for *trade*) non-word substitutions (e.g., "trad" for *trade*); and omissions and insertions of words. The researcher also notes any self-corrections made. According to the authors, the inclusion of miscue analysis provides information about word identification strategies, and suggests whether the child is focused primarily on reading out the words or whether they are concerned with deriving meaning from the text. When these observations are related to the comprehension score, it can indicate whether the child's focus is at the word level or a higher level of understanding.

Comprehension questions

After the passage has been read, the QRI provides questions to be asked which assess explicit and implicit comprehension. The child is allowed to look back at the passage to answer these questions. There are four or five explicit questions and zero, two, four or five implicit questions depending on the level of the passage. The QRI authors advise caution in using the questions to assess a child's ability to answer the two different types (explicit and implicit) since there only a few questions for each category, but the responses do provide a general measure of comprehension.

The strategy interview procedure

After the comprehension questions the researcher continued with the interview. The child was asked about using background knowledge, inference, questioning, determining importance, monitoring comprehension, visualising, synthesising and retelling. The interview questions were used verbatim as far as possible, but as the important thing was to find out about the strategies the child had used in reading the previous passage, clarification was sought when answers were unclear. The researcher was careful not to lead the child's responses in any way, but as the intention was to analyse the responses, rather than rely on a count of the comments, (as in the think-aloud) it was not so important to use only specified prompts. In practice, the questions about synthesising proved to be particularly challenging, and several children failed to make any response. As some children became concerned at their inability to answer some of these questions the rest of the section was omitted. For this reason the section on synthesising was not included in the results. Only the first question about retelling was used, and that was combined with the determining importance section.

4.3.1. Treatment of assessment results

This section explains how each of the assessments used in Study 2 was marked or coded. This is particularly important for the assessments which are not standardised, and which require further explanation.

4.3.1.2. TOWRE

The TOWRE was scored in accordance with the instructions in the manual. As the test had been tape-recorded it was possible to check the accuracy of scoring after the testing session. It was not considered necessary to have these tests scored by a second researcher as the measure consists of a straightforward count of the number of items read in 45 seconds.

4.3.1.3. YARC

The reading tests were marked according to the instructions in the manual, which gives examples of correct and incorrect answers for the comprehension questions. As the testing sessions were recorded it was possible for the tests to be second marked to ensure reliability. Of the 39 test tapes, 10 were chosen at random and second marked by an experienced teacher, qualified in Special Needs teaching and practised in administering reading tests. The second marker was blind to the participant and time point of testing. Agreement between the scores of the first and second marker was 94%. Those answers which received different scores were discussed and the remaining tests were re-examined in the light of the discussions.

4.3.1.4. Think–aloud

The recordings of the QRI passage and think-aloud comments were transcribed. The transcriptions were made of the entire process for each child, and were made verbatim. The use of punctuation was as close as possible to speech presentation and consistent with what is typically acceptable in written text (Mergenthaler & Stinson, 1992). Transcripts included elisions, mispronunciations, slang, non-grammatical errors and nonverbal sounds (McLellan, MacQueen, & Neidig, 2003). Whilst recognising that transcripts "are not the rock-bottom data…..but are artificial constructions from an oral to written form of communication" (Kvale, 1996) every effort was made to record the children's responses as accurately as possible.

Concept questions

Prior knowledge was assessed through the use of concept questions, as supplied with the QRI. These ranged from three to five questions, depending on the level. Each response was marked out of three. For three points the answer needed to be a precise definition, or an answer specifically related to the passage content. Two points were awarded for an example of the concept, but which did not appear in the text, and one point was given for general or personal associations (examples of answers for all these categories are given in the QRI, chapter 7, pp 48 – 51). The QRI authors have found that children who score at least 55% of the points possible on the concept task score above 70% on comprehension questions on the related passage.

Miscue Analysis

The recordings of the QRI passage being read aloud were listened to and the miscues that the researcher had noted on a copy of the text during the testing session were checked. For a more detailed description of how this was recorded and carried out, see Appendix K.

Think-aloud responses

The think-aloud comments were analysed using a coding schedule. To enable comparison with other studies of comprehension strategy use it was considered advisable to use a previously existing schedule. The schedule chosen was that of Schellings, Aarnoutse and van Leeuwe (2006) as it was developed for children in the third-grade (aged 8 to 9) which was very close to the age group in this study (aged 7 to 8). It was developed on the basis of theories of reading comprehension which have been discussed in the introduction (Kintsch, 1988; Pressley, 2000; Van Dijk & Kintsch, 1983) which see reading as an active and strategic process.

The schedule was amended after inspection of some of the think-aloud data, as it became clear to the researcher that some of the categories required dividing, as they did not distinguish between the different levels of some strategies. For example, the original category of visualising (B9) was divided into visualising that drew on the words in the text (B9) and visualising which went beyond the text (B9a). Additionally, since Schellings and colleagues used texts with deliberate inconsistencies, some categories were omitted (e.g. B11, reacting to a question in the text) as they were not applicable to the texts used in this study. The categories used and examples of think-aloud responses for each are given in Appendix L. Initially, the procedure outlined by Schellings and colleagues whereby the think-aloud comments were analysed in terms of units comprising a more or less complete idea,

was adopted. However, recoding at a later date showed this led to inconsistent scoring and the procedure of Trabasso and Magliano (Trabasso & Magliano, 1996) was adopted instead whereby the coding was conducted at the clause level. A clause contained a predicate (a main verb) and expressed an event, activity, or state.

The transcripts were anonymised and the coding was carried out by the researcher. Then 25% of the transcripts (chosen by a random number generator) were marked by a postgraduate student trained in the use of the coding schedule. Inter-rater agreement was 94%. The remaining six per cent were resolved by discussion, and the rest of the transcripts were re-examined in the light of the discussion.

4.3.1.5. The strategy interview

The strategy interviews were transcribed in the same way as for the think-alouds. The transcripts were anonymised and any references to date or time point were removed. Three years after the first interviews and seventeen months after the final interviews at Time 4, the transcripts were scored by the researcher, who was blind to the child or time point. The MPIR scoring rubric (which can be found in Appendix H) was used to score the strategy interview. This rubric allocates a score of between 1 and 5 for each of the seven sections (using schema, predicting, questioning, determining importance, monitoring comprehension, visualising, and synthesising). The authors give an overall criterion to bear in mind: When a reader can go beyond explaining his or her thinking and begins to articulate how using a strategy helps him or her comprehend better, then the response should be scored at least at a level 4.

The scores for determining importance and the retelling question from the synthesising section were averaged. This gave a score for skills which relate to summarising, which was one of the strategies taught in the RT intervention. The scores from the interviews thus related to four of the five strategies taught[13], plus the use of background knowledge (which was not taught explicitly, but which did form part of the class discussions) and monitoring comprehension, which was an ongoing theme behind each strategy. As previously noted, the scores for synthesising were not included. The maximum score was therefore 30. Examples of the comments for each level within each category are given in the appendix (Appendix M).

---

[13] Although the strategy of clarifying was not covered separately, it was encompassed in monitoring comprehension.

4.3.2. Details of the Reciprocal Teaching intervention

This section will explain how the class teacher implemented RT, and the plans and materials she used. Previous studies have been criticised for not including enough detail concerning how the teaching was carried out (Pressley, et al., 2006; Rosenshine & Meister, 1994). This section, and the materials included in the appendix to which the section makes reference, will give details of what happened in the class over the course of the two periods of instruction – ten hours of RT (between September 2009 and January 2010) followed by ten hours of RT plus visualisation (between February 2010 and May 2010).

In this study, the RT instruction was delivered by the class teacher rather than by the researcher as in Study 1. The reason for this was to ascertain whether it was possible for a class teacher to undertake RT in a whole class situation and to do so with the minimum of support. For an intervention to be practicable and acceptable it should not require extensive training. A criticism of the Transactional Strategies Instruction programme is that teachers need so much time for initial and ongoing training (4 half-day in-service training and ongoing support sessions, (El-Dinary & Schuder, 1993)). The model for RT in this thesis would be the usual one-day in-service training, supported by reading and learning materials.

The class teacher had a B.Ed. degree and 20 years of experience in primary school teaching across the year groups from Years 2 to 6. She had spent the last year teaching Year 3, and the previous three years teaching Year 5, and it was in Year 5 that she had come across RT when the researcher taught her class as part of Study 1. In the following year the teacher had read a book about RT (Oczkus, 2003) and introduced her new Year 3 class to the four characters and tried some of the techniques, and she was now keen to implement the programme in full. In preparing the programme, the teacher drew on her knowledge of RT from hearing about the research results from Study 1 in Year 5, in a twilight presentation by the researcher after Study 1, her reading of Reciprocal *Teaching at Work* (Oczkus, 2003), the unpublished manual compiled by Palincsar and colleagues (1989), and her experience with trying some of the techniques the previous year. The researcher had previously provided the teacher with these materials, and at the start of the programme she was also introduced to the guidelines provided by Rosenshine and Meister (1994) which she was asked to follow. The planning of the intervention was carried out by the class teacher, and the researcher took the role of an observer in the classroom. The programme and the instruction were the teacher's responsibility, and it was explained to her that it was important for the researcher to take the role of observer, to see if it was possible to implement the instruction effectively

in the way a teacher might go about changing her classroom practice in the usual way, i.e. by means of a short in-service session, followed by reading and the use of supporting materials. The class teacher prepared the lesson plans in accordance with the school's existing requirements. Examples of these lesson plans are given in the appendix (Appendix N).

The class teacher decided to introduce the four RT characters using a PowerPoint presentation, and she dressed up as each of the four characters to make them more memorable. The same names were used as in Study 1, viz. Mystic Mike (a weatherman) for predicting; Clarifying Clara (a scientist) for clarifying ; Dennis the Detective for questioning and Summarising Susie, a reporter, for summarising. The teacher referred to the characters as "The Fab. Four". After the initial lessons, which involved work in pairs, the children were divided into four mixed-ability groups of five and one group of four. Children from both classes were in each group. There was a teaching assistant present at each lesson, and the researcher sat in the classroom as an observer.

4.3.2.1. Teaching materials

The characters and their strategies were displayed in the classroom and an example can be found in Figure 9. Each child made a poster of their own incorporating the characters and these were also displayed (Figure 10 gives an example). Each child was given two bookmarks which illustrated the characters and gave guidelines as to how they worked (Appendix O) as had been done in Study 1. Each group had a notebook in which the teacher or the teaching assistant made notes about who was the teacher and who had which character role in each lesson. They also made brief notes about each child's contribution. The groups also had a folder each for keeping any written work.

As in Study 1, a class novel was used; in this instance it was *The Owl who was Afraid of the Dark* (Tomlinson, 2002), which had been used as a text in Year 3 for the past 2 years. Each child had their own copy, and as in previous years, the text was used as a stimulus for drama, writing and artwork, as well as being used in the RT lessons. However, as the custom had been to cover this book in its entirety in the autumn term, the class teacher utilised a variety of other texts as the year progressed, which included a poem, a history text and short chapter books which were available in sets (e.g. *Scratch and Sniff* (Ryan, 2006) and *Buffalo Bert* (Morgan, 2004). She also used a series of short texts which she had used in the past to help teach children about inferencing. A reference was not available for these texts as the teacher could not recall where they came from; examples are therefore included in the appendix (see Appendix P).

Figure 9. A display of the Reciprocal Teaching characters and an explanation of their roles, from the classroom in Study 2.

Figure 10. A section of the display of children's posters designed about the four characters, from the classroom in Study 2.

4.3.2.2.. Lesson format

A timetable of the lessons delivered during the instruction period is given in the (Appendix Q). As in Study 1, the RT instruction was planned to take place in a ten-week block of one hour lessons once a week. However, as there was more freedom in the time-tabling in Study 2, the four introductory sessions took place twice a week, to give a more intensive period of instruction at the beginning, in common with the programme as it was originally devised. Subsequently, there were occasions when the class teacher was unable to make an hour available, and 30 minute lessons were used instead, but the total instruction time amounted to 10 hours. The lessons were usually on a Tuesday morning at 9.30am. The second block of RT plus visualisation also amounted to 10 hours, once a week, with four lessons having to be of 30 minutes. The timetabling of lessons was the same as in the first block of ten weeks.

From the lesson plans for the first two weeks of the instruction (Appendix N) it can be seen that in the first lesson, the teacher introduced the idea of what makes a good reader and what a good reader does to understand what they read. She talked about the four strategies – predicting, clarifying, questioning and summarising, and dressed up as the four characters, Mystic Mike (with a crystal ball), Clarifying Clara (with a test-tube and white coat), Dennis the Detective (with a magnifying glass) and Summarising Susie (with a microphone and a notepad). In the second lesson she talked about predicting and questioning in more detail and modelled their use with the class reader. Children also worked on these two strategies in pairs, and conducted a plenary session. For lessons 3 and 4 the following week (each lesson being 30 minutes) she covered clarifying and then summarising in the same way.

After the initial sessions, which introduced the strategies, the teacher held a brief session where she asked the class to develop some of their own rules for group work. These set out the guidelines for co-operative working and ensured that each child would be listened to and that they would help each other by being positive. The rules were:

1. Work together

2. Listen to what other people have to say

3. Let *everyone* have a say

4. Think before you speak.

These rules were displayed in the classroom so that the teacher could direct children's attention to them if necessary.

Each lesson began with the teacher assigning the role of 'teacher' within each group, a role that rotated from week to week. Initially the classroom teacher provided scaffolding for the children, in the form of written prompts. For example, for the history text (the sixth lesson) each group was given a sheet asking them what the front and back cover of the book told them, what was the subtitle, and what did that tell them, and then what did the pictures tell them the book might be about. They were then asked to write at least three questions about Norman castles, which should have the answers in the text, and what else did they still wonder about Norman castles. The teacher had the role of making sure everyone contributed and was encouraged to give feedback - always beginning with something positive. The teacher also had to be responsible for talking about the group's

work at the plenary, which took place in the last 15 minutes of the lesson. The group work lasted 45 minutes, with the classroom teacher and the classroom assistant rotating between the groups to make notes and help the groups with their discussions. The classroom teacher was also involved in keeping order when groups became too noisy or spent too much time off task.

As the children became more familiar with the strategies and how to implement them, the scaffolding was gradually withdrawn, until after 13 hours of instruction the groups were no longer given written prompts and each child was assigned a particular strategy for each lesson. For the last 7 hours, each lesson began with the teacher handing out the notebook for each group. In the notebooks, she had written the name of the child who was to take the role of the teacher, and assigned roles within the group (the role of Predicting Pete was not assigned for these sessions, since it was not an appropriate strategy to use for the texts used). The text for that lesson would be given out and each group was told to begin with reading the text silently to themselves. The teacher in each group then read the text aloud. The child assigned the role of Clarifying Clara would talk about any unknown words and attempt to work out their meaning, with the help of the teacher and other children in the group where necessary. Then the teacher would ask Dennis the Detective to pose some questions to the group, which they would talk about, and then Summarising Susie was called upon to make a summary of what had been read. The children did make some rough notes to help themselves remember the words to clarify, questions to ask, or a summary of what happened, since they knew there would be a plenary session at the end of each lesson where the teacher in each group would be asked to contribute. However, the emphasis was on oral contributions and there was not any expectation that this written work would be seen by the classroom teacher.

The classroom teacher and teaching assistant again circulated between the groups, but the classroom teacher was always careful to ensure that when a child with a reading difficulty was assigned the role of teacher that she was present to help with the initial read-through of the text. The final 15 minutes were spent in a plenary, where the classroom teacher would ask the child playing the role of teacher in each group to talk about how well the discussions had proceeded, how well the group had worked and then to talk briefly about a particular strategy.

After the first block of 10 hours of RT, the Time 2 assessments were carried out. When these had been completed, the second period of instruction began with the introduction of a fifth character, Visualising Vincent. The class teacher dressed up again, as she had done when introducing the original characters and the children were given another bookmark with this strategy. The visualising was introduced in the same way as in Study 1. The teacher had access to Bell (1991) and she used the structure words and the movement from familiar single words to phrases to short texts outlined in that book (and covered in more detail in chapter 3 of this thesis). These activities took place over four 30 minute lessons. After this introduction the RT lessons carried on in the same way as for the first block of instruction, with the children in each group taking it in turns to be Visualising Vincent, as they did for the other four characters.

4.4. Results

The aim of Study 2 was to investigate in detail, using qualitative measures, how RT affected strategy use over time in those children receiving instruction. Although examining in detail the qualitative data obtained through think-alouds and interviews was the primary focus of this study, some quantitative measures were used, and these are examined first. This data consisted of scores in the YARC and TOWRE reading tests as well as quantitative aspects of the data from the think-aloud protocols and strategy interviews, both before and after the intervention with RT and then again after RTV. There was also a one-year follow-up during which time the class received normal instruction. There was therefore no control group in the present study, but analysis of the YARC reading test results, to be presented first, gives a measure of progress, whilst calculation of the ratio gain allows for an examination of the gain in relation to normal rates. Following the presentation of the TOWRE word reading and the YARC reading comprehension test results, the quantitative aspects of the think-aloud protocols and the strategy interviews are covered, followed by the longitudinal data for four children selected as case studies.

4.4.1. The TOWRE word reading scores

The TOWRE word reading efficiency standardised scores were obtained prior to instruction (Time 1), after 10 weeks of instruction with RT (Time 2), after an additional 10 weeks of instruction with RTV (Time 3) and at a one-year follow-up (Time 4). Preliminary inspection of the data showed there were no outliers and the data were normally distributed. A repeated

measures ANOVA was used to analyse the scores with time as the within-subjects variable. There was no violation of the assumption of sphericity. The effect of time was not significant (F(3,24)=2.08, p =.13).

4.4.2. The YARC reading comprehension scores

YARC reading comprehension standardised scores were obtained at the same four time points as for the TOWRE. One child was initially given a comprehension score from one reading passage, following the instructions from the first edition of the test (Snowling, et al., 2009). However, the manual for the second edition of the test (Snowling et al., 2011) explains that reliability is improved if pairs of passages are scored; with the implication being that a comprehension score obtained from one passage is unreliable, it was decided to exclude data from that child. This child also had a high absence record, which made excluding her data important for an additional reason. Scores were therefore obtained for 10 children at Times 1 to 3 inclusive and for 9 children at Time 4, as 1 child left the school at the end of Year 3. A summary of these data is given in Table 11. There were no outliers, and although there was some negative skew, showing more children were better readers than might be expected in a normal sample, there was not a great enough departure from normality to exclude the use of parametric tests (T1 D(10)=.24, p=.12, T2 D(10)=.25, p=.07, T3 D(10)=.24, p=.12, T4 D(9)=.24, p=.14).

Table 11. Mean standardised scores for YARC reading comprehension at four time points in Study 2 (standard deviations are in parentheses).

	N	Mean
Time 1 (pre-instruction)	10	101.30
		(7.80)
Time 2 (after 10 weeks RT)	10	107.00
		(7.33)
Time 3 (after 10 weeks RTV)	10	110.90
		(8.02)
Time 4 (1 year follow-up)	9	106.56
		(7.06)

A repeated measures ANOVA was used to analyse the scores at the four time points. There was no violation of sphericity. The results revealed there was a main effect of time (F(3,24)=5.13, p=.007). Paired- sample t-tests were used to compare the scores at the four time points. The difference between Time 1 and Time 2 approached significance, whilst the large effect size showed that the increase is of practical importance (t(9)=2.08, p=.07, r=.57). The difference between Time 1 and Time 3 was significant (t(9)=3.26, p=.01, r=.74) whilst that between Time 1 to Time 4 was again close to significance, with a medium effect size showing practical worth (t(8)=1.94, p=.09, r=.32). The difference between Time 2 and Time 3 was significant (t(9)=2.94, p=.02, r=.49) as was the decrease in scores between Time 3 and Time 4 (t(8)=2.27, p=.05, r=.63). The difference between Time 2 and Time 4 was not significant (t(8)=.20, p=.85, r=.22).

As in Study 1 the reading comprehension scores were also analysed using ratio gain. As outlined previously, this involves taking the reading age equivalent score in months post-intervention, minus the reading age in months pre-instruction, and dividing by the time elapsed in months. Chronological ages and reading age equivalent scores in the YARC for the children in Study 2 are given in Table 12, where the considerable difference between chronological age and age equivalent scores following RT instruction can be seen. As noted in chapter 3, Brooks (2002) maintains that a ratio gain of 1.4 or more is educationally significant, whilst gains of 4 or above are "remarkable" (Brooks, 2007, 2013). A ratio gain of this latter magnitude was found from Time 1 to Time 2 (ratio gain=4.78) and from Time 1 to Time 3 (ratio gain=4.01). From Time 1 to Time 4 (ratio gain=1.77) there was an educationally significant modest gain and from Time 2 to Time 3 (which was after the introduction of visualisation as an additional strategy) the gain was 'substantial' at 3.75 (Brooks 2013).

Table 12. Mean chronological ages and reading ages from the YARC, at each of the four time points for Study 2 (standard deviations are in parentheses).

	Time 1	Time 2	Time 3	Time 4
	M	M	M	M
N	10	10	10	9
Chronological age in months	88.61	92.10	96.10	108.10
	(2.38)	(2.38)	(2.38)	(2.38)
Reading age in months	91.19	108.20	123.20	126.33
	(13.11)	(14.87)	(20.02)	(17.23)

Note. Time 1 (pre-instruction), Time 2 (after 10 weeks RT), Time 3 (after 10 weeks RTV), and Time 4 (1 year follow-up).

4.4.3. YARC reading accuracy scores

A summary of the standardised scores for accuracy at the four time points is given in Table 13. Data was not included for the child for whom a comprehension score was not obtained. Preliminary data analysis showed that there were no outliers and no significant departures from normality at any of the four time points (T1 $D(10)=.15$, $p=.20$, T2 $D(10)=.18$, $p=.20$, T3 $D(10)=.23$, $p=.15$, T4 $D(9)=.21$, $p=.20$). There was no violation of sphericity and a repeated measures ANOVA showed no significant effect of time ($F(3)=2.81$, $p=.06$).

Table 13. Mean standardised scores for accuracy from the YARC at each of the four time points for Study 2 (standard deviations are in parentheses)

	Time 1	Time 2	Time 3	Time 4
	M	M	M	M
N	10	10	10	9
YARC accuracy	104.90	102.20	106.70	99.56
	(8.87)	(6.88)	(10.29)	(10.60)

Note. Time 1 (pre-instruction), Time 2 (after 10 weeks RT), Time 3 (after 10 weeks RTV), and Time 4 (1 year follow-up).

4.4.4. YARC reading rate scores

A summary of the standardised scores for reading rate at the four time point is given in Table 14. Data were not included for the child for whom a comprehension score was not obtained.

Table 14. Mean standardised scores for reading rate on the YARC reading test at four time points in Study 2 (standard deviations are in parentheses)

	Time 1	Time 2	Time 3	Time 4
	M	M	M	M
N	10	10	10	9
YARC reading rate	106.20	105.30	105.10	103.00
	(11.35)	(9.21)	(10.97)	(11.33)

Note. Time 1 (pre-instruction), Time 2 (after 10 weeks RT), Time 3 (after 10 weeks RTV), and Time 4 (1 year follow-up).

Preliminary data analysis showed there were no outliers and no significant departures from normality at any of the four time points. However, Mauchly's test indicated that the assumption of sphericity had not been met ($\chi^2(5)$=.10, p=.008) and therefore Greenhouse-Geisser estimates of sphericity were used to adjust the degrees of freedom. The results showed that there was no significant effect of time ($F(1.74)$=.26, p=.74). In terms of ratio gain, the difference between pre-instruction reading rate and that at follow-up (ratio gain=0.79) shows that the children made less gain than would be expected under normal development.

4.4.5. Self-corrections on the YARC reading test passages

The number of self-corrections that were made by each child during text reading in the YARC was also recorded. Although the YARC does not make any provision for self-corrections in the accuracy figures, and the QRI authors warn of the difficulties inherent in counting self-corrections, it was felt that such a measure would help establish the degree to which children were monitoring their comprehension. Self-corrections were those occasions where a word was initially read incorrectly (or omitted) and then corrected without prompting. Only whole word responses were counted as self-corrections, since in the case of parts of words it is difficult to distinguish between a correction and a hesitation. If a phrase was repeated and this included a previously omitted word, that was also counted as a self-

correction. There was one outlier at Time 4, but as there was little effect on the 5% trimmed mean the data was retained. Although there was some skewness in the data, tests of normality confirmed that the conditions for a parametric test were met (Time 1 $D(10)=.33$, $p=.19$, Time 2 $D(10)=.25$, $p=.25$, Time 3 $D(10)=.21$, $p=.20$, Time 4 $D(9)=.21$, $p=.20$).

At Time 1 all of the children made uncorrected errors and often phrases were read which made no sense at all, e.g. "Their eggs are *planned* with register spots" (instead of *pale* with reddish spots) and instead of *What was that Mum?* , one child read: "Was as that Mum?" However, the children did make more self-corrections as the instruction proceeded (T1 $M=2.10$, $SD=2.18$, T2 $M=5.10$, $SD=2.90$, T3 $M=5.80$, $SD=2.13$, T4 $M=4.67$, $SD=1.94$). A repeated measures ANOVA revealed that there was a significant effect of time ($F(3,24)=6.43$, $p=.002$). Post-hoc paired sample t-tests showed that there were significant differences between the number of self-corrections made at Time 2 and at Time 3, when compared with Time 1 ($t(9)6=.13$, $p=<.001$, $r=.91$, $t(9)=4.47$, $p=.002$, $r=.85$, respectively). The difference between the number of corrections made at Time 4 compared to Time 1 was not significant, but the large effect size shows it may be important in practical terms ($t(8)=1.93$, $p=.09$, $r=.56$). The differences at the other time points were non-significant. Thus, there was a statistically significant increase in the number of self-corrections from Time 1 to Time 2, after ten weeks of RT. From Time 2 to Time 3, after a further 10 weeks of RTV instruction there was no significant change in the number of corrections, although there were still significantly more made than at Time 1. After a period of a year without any RT instruction, the number of self-corrections fell, but not by a significant amount ($t(8)=.83$, $p=.43$) and when compared to Time 1 the effect size suggests that there was still a substantial increase in level of self-correction.

4.4.6. Spontaneous comments made during the YARC testing

As was discussed in chapter 1, Garner (1992) noted that children receiving RT would eventually make spontaneous interjections about the text whilst the teacher was reading. It was also suggested that looking at the way children make comments, or question their understanding during reading, might be a way to assess the change into more active readers. This was investigated in this study by looking at the think-alouds in the QRI passages, but it was also possible to look at children's comments made during the reading of the YARC passages. These passages are presented very differently to the QRI. During the QRI the children are asked what they are thinking, but during administration of the YARC, far from being asked to comment, they are discouraged from doing so by the instructions. The

children are told that they are being timed, and although the administrator goes on to tell them to "just read normally", the instructions combined with the use of a stop watch must be a disincentive to stop reading and make a comment.

As might be expected given the formality of the test and the wording of the instructions, at Time 1 (pre-instruction) there were no spontaneous comments from any of the children. At Time 2 (after 10 weeks RT instruction) 7 children out of the 10 made a total of 15 comments. Six comments queried either the meaning of a word (e.g. *"Foaming*, what does that mean?") or expressed confusion over a mispronunciation when decoding had failed (e.g. "I don't know that word" or "That doesn't make sense"). One child identified with the text by saying that he had a cat and that it too hunted birds, whilst two others made it clear that they had made the necessary inferences to understand the main point of the story; another child after reading the final sentence about a little girl getting lost in the fog said: "That was Chris" (the brother with whom she had been fighting earlier in the story) and one child, after reading that Mum's handbag had been found in the toilets, said "She must of (*sic*) gone to the toilets and left it there" and when Mum apologises at the end for blaming Dad, the same child said "So it was her fault not his".

At Time 3, nine children made 18 comments (the same children as at Time 2 plus two additional children). Again the majority of the comments related to individual words with which the children were unfamiliar (e.g. "'brandishing', I don't know that word") but one child also commented about a phrase she did not understand ("That doesn't make sense") and which she re-read. Three children reacted to the burglar who had broken into a policeman's house by laughing and saying it was funny. Comments were also made which connected the children to the text (e.g. "I went camping last year" and "I've seen a lizard in a zoo") and one child thought that 35 eggs was a lot for a lizard to lay at once.

At Time 4, the number of comments declined. Only three children said anything spontaneously; two queried words they could decode but did not understand (e.g. "'ensuring', I don't know what that means") and one child said "I think drones are workers" when trying to clarify what drones are in the passage about bees The pattern of spontaneous comments suggest that some children at least were becoming more active in their reading as they increased the number of spontaneous comments, made under test conditions and without prompting, over the period of instruction. It was also noted that one child did not make any comments, at any time point, and that was child 2.3. Her resistance to the RT instruction is explored further in section 4.6.6.

144

4.4.7. The Qualitative Reading Inventory

In this section the results from the comprehension questions, the think-aloud and the strategy interview are presented. As outlined in the previous section, this data was collected from reading aloud a graded passage of expository text taken from the QRI. The think-aloud procedure required the child to stop at pre-determined points in the passage and say what they were thinking. After the passage was read, the children were asked comprehension questions. Then a strategy interview was conducted, based on the Major Point Interview for Readers (Keene & Goudvis, 1995). The passage reading, think-aloud and strategy interview were conducted, as for the other assessments, at four time points for each child. Time 1 was pre-instruction (September 2009) Time 2 was after 10 weeks of RT (February 2010) Time 3 was after a further 10 weeks of RT plus visualisation (June 2010) and Time 4 was a follow-up after an academic year without RT instruction (June 2011). The answers to the comprehension questions will be considered first, followed by the think-aloud responses and then the strategy interview.

4.3.7.1. The comprehension questions

The QRI provides 6 or 8 questions at the end of each passage, according to the level. These questions are equally divided between what the authors call explicit and implicit questions. As the number of questions for each category is limited (3 or 4 for each according to the length of the passage) the authors warn against drawing conclusions about the variation between the scores for each category. However, it does provide an additional measure of comprehension, which adds validity, so the results are considered here. If RT improves children's inferencing skills, then it would be useful to look at the sub-types of questions in the comprehension tests used. It is not possible to make any comparisons of question sub-types on the YARC, since they are not consistent across passages and forms.

Explicit questions
The mean and standard deviation for the percentage of questions of both types are given in Table 15. Preliminary data analysis showed that there was some skewness in the data with possible ceiling effects at Times 3 and 4 (Time 1 $D(10)=.23$, $p=.13$, Time 2 $D(10)=.24$, $p=.11$, Time 3 $D(10)=.30$, $p=.01$, Time 4 $D(9)=.26$, $p=.07$) so a non-parametric test was considered to be suitable.

Table 15. Mean percentage of questions answered correctly for the QRI at four time points in Study 2 (standard deviations are in parentheses).

	Time 1	Time 2	Time 3	Time 4
	M	M	M	M
N	10	10	10	9
Explicit questions	65.00	72.50	82.50	81.11
	(24.15)	(24.87)	(20.58)	(20.73)
Implicit questions	41.67	57.50	75.00	69.44
	(11.11)	(28.99)	(20.41)	(20.83)

Note. Time 1 (pre-instruction), Time 2 (after 10 weeks RT), Time 3 (after 10 weeks RTV), and Time 4 (1 year follow-up).

Wilcoxon signed-rank paired sample tests showed that the differences in the percentage of correct responses to the explicit questions at the four time points were not statistically significant, although there were small to medium, and even medium to large, effect sizes (Time 1 to Time 2 z=-0.68, p=.50, r=.15, Time 1 to Time 3 z=-1.44, p=.15, r=.32, Time 1 to Time 4 z=-1.44, p=.15, r=.33, Time 2 to Time 3 z=-1.08, p=.28, r=.24, Time 2 to Time 4 z=-1.84, p=.07, r=.42, Time 3 to Time 4 z=-1.34, p=.18, r=.31).

Implicit questions

Preliminary data analysis showed the presence of outliers at Time 4 and departures from normality for the mean of correct responses to the implicit questions at Time 3 (Time 1 D(10)=.23, p=.13, Time 2 D(10)=.24, p=.11, Time 3 D(10)=.30, p=.01, Time 4 D(9)=.26, p=.07) so a non-parametric test was considered to be suitable. Wilcoxon signed-ranks paired sample tests showed there were differences in the number of correct answers to the implicit questions at all time points except Time 3 to Time 4 (z=-0.38, p=.71, r=.09). At Time 1 to Time 3 (z=-2.54, p=.01, r=.57) and Time 1 to Time 4 (z=-2.20, p=.03, r=.50) the differences were significant and showed a large effect size. At Time 1 to Time 2 (z=-1.19, p=.23, r=.27) and Time 2 to Time 3 (z=-1.82, p=.07, r=.41) the differences were not statistically significant, but there were medium to large effect sizes. At Time 2 to Time 4 there was no significant difference, but there was a small to medium effect size (z=-0.88, p=.38, r=.20).

4.3.7.2. The think-aloud data

The data from the think-alouds constituted two independent variables: the number of comments per think-aloud opportunity and the number of strategies used. The number of comments will be discussed first, followed by the number of strategies.

Think-aloud: number of comments

After coding the think-alouds, according to the schedule (Appendix L) the number of comments per think-aloud opportunity was obtained for each child. The longer passages (over 200 words, from Level 2 onwards on the QRI) contained 4 stop points for thinking aloud, as opposed to the 3 stop points for Level 1, Primer and Pre-Primer passages. Consequently, for the analysis, an average was obtained for the number of comments per think-aloud opportunity, rather than a count of the totals. A summary of the data at the four time points is shown in Table 16.

Table 16. Mean for the average number of comments made per think-aloud opportunity for four time points in Study 2 (standard deviations in parentheses)

	Time 1	Time 2	Time 3	Time 4
	M	M	M	M
Number of comments	1.60	2.39	2.98	2.53
	(0.93)	(0.74)	(0.73)	(0.63)

Note. Time 1 (pre-instruction), Time 2 (after 10 weeks RT), Time 3 (after 10 weeks RTV), and Time 4 (1 year follow-up).

The scores at all four time points had a normal distribution, as confirmed by histograms and an examination of the skewness and kurtosis figures. The assumptions for analysis with parametric statistical tests were therefore met. A repeated measures ANOVA was used to analyse the scores. There was no violation of the assumption of sphericity. The results revealed a main effect of time ($F(3,24)=4.96$, $p=.008$). Post-hoc t-tests showed there was a significant difference in the number of comments generated per think-aloud opportunity between Time 1 and Time 2 ($t(9)=2.20$, $p=.05$, $r=.59$) Time 1 and Time 3 ($t(9)=4.07$, $p=.003$, $r=.8$) and Time 1 and Time 4 ($t(8)=3.68$, $p=.006$, $r=.79$). There was no

significant difference in the average number of comments between Time 2 and Time 3, Time 2 and Time 4, and Time 3 and Time 4.

Think-aloud: Number of strategies reported

Following coding, the number of strategies reported in each think-aloud as a whole was analysed. Inspection of the data prior to analysis revealed the presence of an outlier at Time 3, but as there was little effect on the mean (M= 4.20, 5% Trimmed Mean= 4.28) that data point was retained. A summary of the data is given in Table 17.

Table 17. Mean number of strategies reported during the think-aloud for four time points in Study 2 (standard deviations in parentheses).

	Time 1	Time 2	Time 3	Time 4
	M	M	M	M
Number of strategies	2.70	3.90	4.20	4.22
	(0.93)	(0.94)	(0.92)	(1.64)

Note. Time 1 (pre-instruction), Time 2 (after 10 weeks RT), Time 3 (after 10 weeks RTV), and Time 4 (1 year follow-up).

Preliminary tests showed that the distribution of scores at Time 3 was not normal (D (10) = 0.30, p = .01) so a non-parametric test, the Wilcoxon signed-rank test, was used. The analyses revealed a significant difference in the number of strategies used between Time 1 and 3 (T = 4.50, z = 2.15, p = .03, r = .48) but none of the other comparisons were significant.

4.4.8. The Strategy Interview

Strategy interviews (based on the MPIR, Keene & Goudvis, 1995) at the four time points were scored according to the rubric provided (to be found in Appendix H). For the strategy interview the data comprised scores for four of the strategies explicitly taught (predicting, questioning, summarising and visualising) plus the use of schema and comprehension monitoring, which were implicit in the programme. A total score was obtained, with the maximum score possible being 30. A summary of the data is given Table 18.

Table 18. Mean strategy interview score at the four time points in Study 2 (standard deviations in parentheses)

	Time 1	Time 2	Time 3	Time 4
	M	M	M	M
Strategy interview score	11.30	14.20	17.20	17.89
	(2.92)	(3.01)	(3.71)	(2.37)

Note. Time 1 (pre-instruction), Time 2 (after 10 weeks RT), Time 3 (after 10 weeks RTV), and Time 4 (1 year follow-up).

A repeated measures ANOVA was used to analyse the scores. There was no violation of the assumption of sphericity. A significant main effect of time was found (F(3,24)=13.34, p<.001). Post-hoc t-tests revealed that there were significant differences at Time 1 to Time 2 (t(9)=2.21, p=.05, r=.59) from Time 1 to Time 3 (t(9)=10.07, p<.001, r= .96) from Time 1 to Time 4 (t(8)=7.15, p<.001, r=.93) and from Time 2 to Time 4 (t(8)=3.32, p=.01, r=.76). The difference between Time 2 and Time 3 approached significance (t(9)=2.03, p=.07, r=.56) whilst the difference between Time 3 and Time 4 was not significant.

4.4.9. Children's evaluation questionnaire

As it was suggested in Study 1 that RT might improve children's reading through increasing their motivation, a questionnaire was used at Time 3 (after the intervention finished) to gather information about what the children thought they had learned and whether they had enjoyed the instruction. The questionnaire utilised a combination of Likert scales and open-ended questions. Group statistics for the responses are given in Figures 11 and 12. In Study 1, 90% of the children (N= 33) who received RT reported that the lessons had helped them to improve their understanding of what they had read; in this study the figure was 100% (N=10). When asked how much they had learned, all of the children said they thought they had learned at least a few new things. When asked about enjoyment, only one child said they did not enjoy the lessons much, none said they did not enjoy them at all and four said they enjoyed them quite a lot or really enjoyed them.

The open-ended questions were also about learning and enjoyment. The children said they had learned about using strategies (one mentioned "The Fab. Four", another

mentioned all the characters by name, and three mentioned a specific strategy), they had learnt to "read better" (two responses) and to let people speak (one response). Two did not give an answer. When asked about what they had enjoyed, two said they liked the reading, one said they enjoyed it because it helped you to learn, one enjoyed being the teacher, two enjoyed using the high-lighters, two enjoyed drawing pictures , and one enjoyed reading the books, using the high-lighters **and** drawing the pictures. One child did not make a comment. When asked if there was anything they had not enjoyed, one said no, two said they did not like someone in their group, one said there was not enough time, one said there were not enough non-fiction books to read, one said that sometimes the books were boring and one said they did not enjoy thinking. Three children did not make any comment.

Figure 11. Evaluation questionnaire responses for whether children learned anything at Time 3 in Study 2.
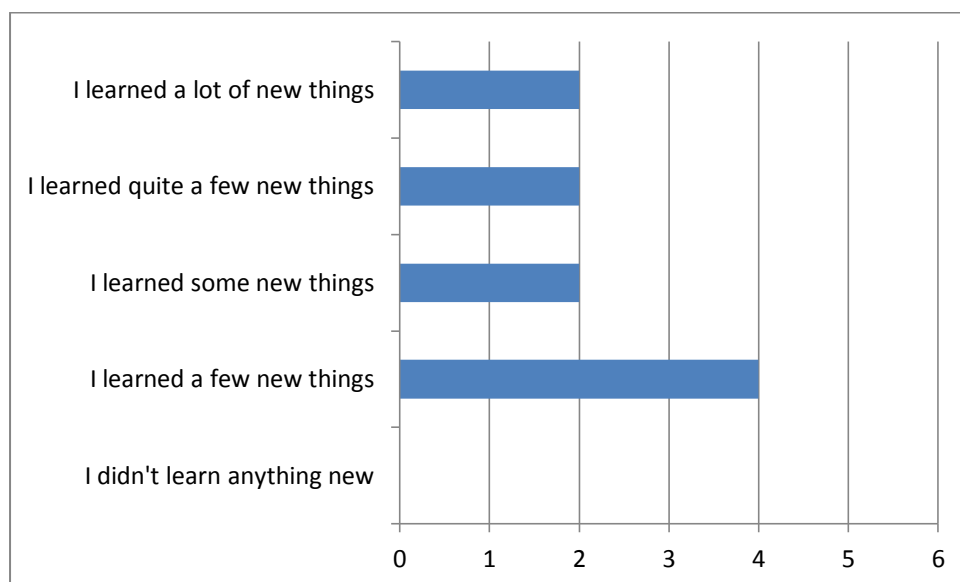
Figure 12. Evaluation questionnaire responses for whether children enjoyed the lessons at Time 3 in Study 2

4.4.10. The class teacher's response to the intervention

At the end of the year, the class teacher was invited to put down her thoughts about the intervention on paper. A discussion of her comments follows. The excerpts are given in the order in which she wrote them and her comments are given in full. The teacher's words provide several interesting points. Firstly, it is clear that her intentions were aligned with the aims of the researcher:

> When I started teaching reciprocal reading I was hoping to increase the children's understanding of what they had read. Most of the children in Year 3 can read a text reasonably accurately, but often cannot tell you what has just happened or infer information. I wanted to give them the tools to extract meaning from the text, to realise their own mistakes and to engage more fully with what they had read.

The teacher had observed that the children were good word readers, but they did not always understand what they had read, nor did they make inferences - they did not make a situation model. The teacher wanted them to be able to do this by monitoring their own understanding and to engage with what they were reading; in other words, to be active readers.

Secondly, the class teacher thought the teaching of the strategies went well:

> The teaching of the 5 strategies, predicting, questioning, clarifying and summarising and visualising went well and in group reading the children were clear which character had to do which, and with guidance could deploy those strategies. The prompt cards were very useful for this.

> The children enjoyed discussing the texts and reading it altogether gave the children who struggled with decoding a chance to show that they did actually understand the content of what they had read.

From a practical point of view, there is an acknowledgment that the props were useful, but perhaps more importantly, the teacher has expressed the opinion that the intervention provided a chance for poor decoders to participate in the discussion of a text they would not otherwise have the opportunity to read as it would be considered too difficult for them. Thus RT, through the use of heterogeneous groups, gives access to more difficult texts which is denied when Guided Reading groups are more traditionally constituted.

> The group work was not easy to begin with, not least for the teacher:

> The children and I found the group work challenging to begin with. The strategy suggests the teacher hands over the running of the group to a different child each session and the teacher's role is largely observational. As a control freak I found this very difficult.

> The class teacher is being very honest here, but she is expressing the often unequal relationship of teacher and taught (Jackson, 1967). Classroom talk has been, and in many cases still is, dominated by teacher talk (Durkin, 1978; Myhill, 2006; Parker & Hurry, 2007). Even with the introduction of the National Literacy Strategy , the recitation script continues to rule (Alexander, 2004), and teachers spend the majority of their time either explaining or using highly structured question and answer sequences, with most of the questions being of a low cognitive level designed to funnel pupils' responses towards a required answer (Smith, Hardman, Wall, & Mroz, 2004). In contrast to this, RT aims to give children time to talk, and to make a contribution; as Garner noted and as we saw in the introduction: "Information flow in the RT transcripts was clearly *not* unidirectional from texts to readers or from teachers to readers; much information came from readers" (Garner, 1992 p. 228).

> However, it was not only the teacher who found this transition difficult initially:

....most of the children who were leaders found it very difficult to recap all the strategies and help those in their groups. That said, as the weeks went on they did become more independent and the more able children were able to lead.

The children did need support, but RT acknowledges this in the model of transfer of control. It is not expected that the children will immediately be able to utilise the strategies and run the groups unaided. The teacher needs to release control gradually.

A further difficulty seemed to arise from the particular form of the intervention which this teacher chose to use, a form taken from Oczkus (2003) as previously noted, whereby the teacher assigns a particular role to each child in the group:

Next time I do this I think I might focus on one or two strategies per session as some of the less able children found it difficult to be the only one using their strategy in their group, particularly if the group leader was not able to help them.

This echoes a comment in one of the note books by the teacher in the spring term, that some of the children found it difficult to stick to their own strategy. Perhaps the original format whereby all children in the group have an opportunity to discuss each strategy is a better model. Finally, the teacher observed:

I found that RT actually broke down the skills needed to help children make more sense of what they were reading. It was particularly interesting to find how difficult the children found it to ask questions of the text, suggesting that even if they could pick out information they weren't really thinking about its meaning. It was interesting that the second time I taught this[14] the children were much better at this, although collectively they were not such an able group.

I also learnt that initially children understand far less of what they have read than you think they do, so the techniques to help clarify the meanings of words were particularly useful, albeit a bit of an eye opener the first time we did it.

The teacher found that questioning was the most difficult strategy for the children to learn and that clarifying was not only useful, but served to highlight that she often assumed children knew a word when they did not. The latter becomes apparent when looking at the words children highlighted in the text as words they needed to clarify. They were often surprising. As for questioning, it is not surprising that children find this difficult as it

---

[14] The teacher had used some of the Reciprocal Teaching techniques with her class the previous year.

something they are so rarely asked to do. As we have seen, in most classrooms the teacher asks the questions.

The next section will look at the four case studies which were selected to examine any changes taking place in strategy use, during and after the intervention, in different types of comprehenders.

4.5. Case studies

The use of qualitative measures at each of the four time points provided data for each child in the form of a miscue analysis and a think-aloud protocol from the QRI, and an interview on strategy use (MPIR). The QRI also provided an assessment of background knowledge and a small number of comprehension questions, both explicit and implicit. The presentation of this data is discussed in more detail after an explanation of which children were chosen for closer analysis, and why they were chosen.

4.5.1. The selection of cases for individual presentation

As the aim in Study 2 was to look in detail at the effect of RT and RTV instruction, then the case study data is as important as group data. Figure 13 shows YARC comprehension standardised scores for each of the ten participating children in Study 2 at the four time points. The black line in the figure depicts the mean for the group. In light of the high degree of individual variation illustrated by the figure, four children were selected for individual case study presentation: pupils 2.3, 2.4, 2.9 and 2.10. Since research into RT has largely involved poor comprehenders it was decided to choose one child who was identified by the class teacher as a good reader, pupil 2.10, as well as one who was identified by the class teacher as a poor reader (in relative terms) pupil 2.4. It can be seen that although the trend to improve in reading comprehension across time was observed for most children, there were notable individual differences, viz. pupil 2.3 and pupil 2.9. Pupil 2.3 was the only child to show a decrease in comprehension score between Time 1 and Time 3, and pupil 2.9 showed a much greater improvement from Time 1 to Time 2 than the other children.

Figure 13. YARC comprehension standardised scores at four time points for all pupils in Study 2.
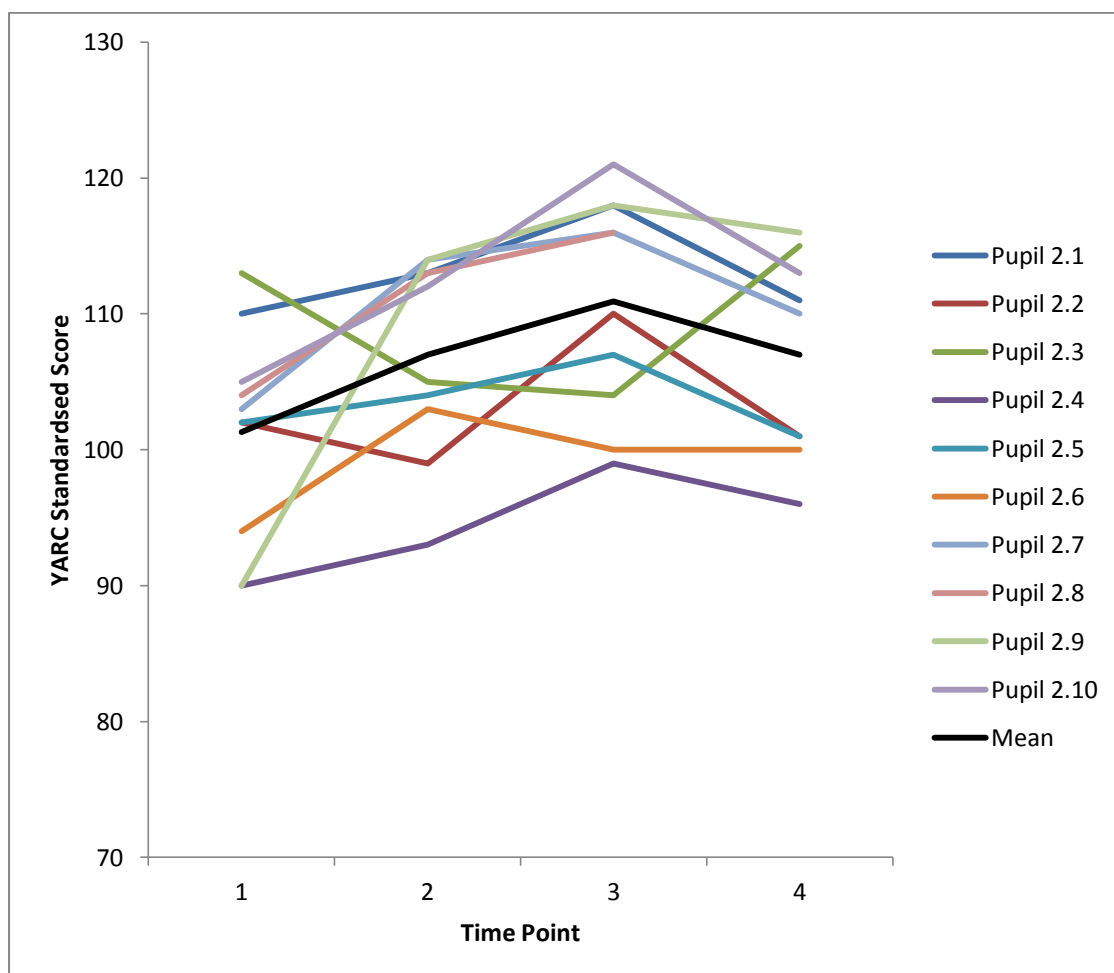
4.5.2. Presentation of the case studies

These four cases will be presented in the following way: data are provided showing the pupil's scores for word reading, the number of strategies reported in the think-aloud and the strategy interview scores for each of the four time points. Then a figure shows the standardised reading comprehension scores and selected comments from the think-alouds and strategy interviews. The table that follows gives the QRI passage used and its level, along with scores for familiarity, accuracy, acceptability and comprehension (as assessed by explicit and implicit questions from the QRI). This is followed by a commentary on how the child read the passage aloud and the types of errors they made. For reasons of space, the details of the think-aloud responses and strategy interview are provided in the appendices,

but an overall assessment of the child's ability to use appropriate strategies and to read actively is given here.

4.5.3. Pupil 2.10 (A.): a good reader

A. was identified by the class teacher as the best reader in the class. A. was an enthusiastic reader of non-fiction books and he was always the first to get his book out for silent reading and the last to put it away. Figures 14 and 15 show his word reading and comprehension standardised scores over the duration of the study, along with plots of his strategy use from the think-alouds and from the interview, along with selected comments. Interestingly, his word reading scores were stronger than his reading comprehension score, which, although above average, was not the strongest in the class. The teacher may have been lead to believe his reading was better than it was as a consequence of A.'s proficient word reading skills. However, with a reading comprehension age of seven months in advance of his chronological age, he was a good reader.

Figure 14. (a) TOWRE standardised score for total word reading efficiency (b) number of strategies reported in the think-aloud and (c) strategy interview score, for pupil 2.10 in Study 2.

(a)

(b)

(c)

Note. Time 1 (pre-instruction), Time 2 (after 10 weeks RT), Time 3 (after 10 weeks RTV), and Time 4 (1 year follow-up).

Figure 15. Standardised scores for YARC comprehension together with think-aloud and strategy interview comments at four time points for Pupil 2.10 (A.).

Table 19 shows the QRI passages read at each time point and the corresponding QRI scores. There is a steady progression in terms of acceptability and comprehension from Time 1 to Time 3, which equates to the increase in comprehension scores on the YARC over the same period. There is an equivalent slight decline from Time 3 to Time 4, the follow-up period.

Table 19. The QRI passages and scores obtained at the four time points for Pupil 2.10 (A.).

QRI	Time 1	Time 2	Time 3	Time 4
Passage	Level 3: Cats, Lions and Tigers	Level 4: Saudi Arabia	Level 6: Earthquakes	Level 7: Life Cycles of Stars
Familiarity	50%	42%	66%	58%
Miscue Analysis Accuracy Acceptability	95% 97%	96% 98%	97% 99%	97% 99%
Comprehension: Explicit Implicit	75% 25%	75% 100%	100% 100%	80% 100%

Note. Time 1 (pre-instruction), Time 2 (after 10 weeks RT), Time 3 (after 10 weeks RTV), and Time 4 (1 year follow-up).

4.5.3.1. Miscue analysis for pupil 2.10 (A.)

Time 1

A. read aloud at speed (102 WPM); conjunctions and other small words were often misread, or omitted, for example, he read: "When kittens are first born they drink milk **of** their mothers" (instead of *from*). However, although grammatically incorrect, the overall meaning was maintained.

Time 2

A. struggled with the proper nouns in this text, but he self-corrected when a sentence did not make sense – for example, he read: "They don't worry **what**………..they don't worry **that** it will rain." He continued to make a number of small errors with word endings (he missed

159

the letter 's' to make *wave* and *spring* into the plurals which the text required). However, his reading was notably slower than at Time 1 (86 wpm).

Time 3

A. read accurately, but he also made several self-corrections, which showed he was reading for meaning. He showed a greater attention to the text, particularly the endings of words. He read faster that he did at Time 2 (108 wpm) which is the same speed as at Time 1. A. commented "Oh no" when he accidentally read the line above again at the beginning of a sentence near the end of the text. He immediately knew it did not make sense, showing a good awareness of comprehension monitoring.

Time 4

A. read very fast (127 wpm). He self-corrected two errors (*style* was changed to "cycle" and *nebula* was initially mispronounced). Both these corrections were evidence that A. was reading for meaning. The first was a predictive miscue, since the preceding word was *life*, so A. initially read "life-style". Other substitutions maintained the meaning (e.g., he read "a" for *the* and "becomes" for *comes*) but he did read "nuclear fusion **makes** place" instead of ***takes*** place", which did not make sense.

4.5.3.2. A Summary of think-aloud and interview responses for pupil 2.10 (A.)

Time 1

A. was aware of a number of strategies – predicting, clarifying, summarising, visualising, and using background knowledge. However, only summarising was evident in the think-aloud. Questioning was not a strategy of which A. was aware, and although he knew reading is for meaning, he was still more focused on the word level at times. For A.'s detailed responses at all time points see Appendix R.

Time 2

A. knew how to use a number of strategies. As his interview showed, he could explain, predict, clarify, question, summarise and visualise. In the think-aloud he showed that he could relate what he was reading to something outside the text. He made inferences in his think-aloud and in the question answers, and although they were not always correct, they did show an engagement with the text and an ability to read between the lines. A. also knew that the text was about more than decoding the words, since he found the words "easy" but the ideas "middle". A. did not include all the important points in his summary.

160

Time 3

A. only used one strategy in his think-aloud - questioning - but he used it very effectively. His questions were well thought out and showed he was thinking about the text and its implications (it also demonstrates the inherent danger in using only counts in a coding schedule). In the strategy interview A. could explain about the five strategies he had been taught, showing he had a range of strategies that he could call on when necessary.

Time 4

A. was reading more actively that he did at Time 1, showing he could use a range of strategies effectively. He predicted, summarised and questioned, but he seemed to have forgotten some of his comprehension monitoring strategies, which may account for the drop in his YARC comprehension score from Time 3 to Time 4. He was involved in what he read and knew that he was learning from this non-fiction text, but his level of tolerance for coherence had fallen, and he should be reminded what to do when understanding breaks down and not to be content with reading a sentence which does not make sense.

4.5.4. Pupil 2.3 (B.): the only pupil to show a decrease in reading comprehension score from Time 1 to Time 3

B. was an excellent word reader (TOWRE total word efficiency standardised score of 126 at Time 1). She read at an early age and considered herself to be "the best reader in the class". She was the first to become a 'free reader' and was proud of that fact. Figures 16 and 17 show her word reading and comprehension standardised scores over the duration of the study, along with plots of her strategy use from the think-aloud and from the interview, along with selected comments.

Figure 16. (a) TOWRE standardised score for total word reading efficiency (b number of strategies reported in the think-aloud and (c) strategy interview score, for pupil 2.3 in Study 2.

(a)

(b)

(c)

Figure 17. Standardised scores for YARC comprehension together with think-aloud and strategy interview comments at four time points for Pupil 2.3 (B.).

Table 20 shows the QRI passage levels and corresponding scores at the four time points. The YARC comprehension trend over time is mirrored by the scores on the miscue analysis. The comprehension scores on the QRI are higher then might be expected given the dip in the YARC scores, but the passages at times 2 and 3 had very high familiarity scores and may well have been answered from background knowledge. In this respect, the YARC offers a truer tests, since as we have seen, the developers tired to ensure that it would not be possible to answer the questions correctly without having read the passage.

Table 20. The QRI passages and scores obtained at the four time points for Pupil 2.3 (B.).

QRI	Time 1	Time 2	Time 3	Time 4
Passage	Level 4: Beavers	Level 5: The Octopus	Level 5: Laser Light	Level 6: The Lifeline of the Nile
Familiarity	33%	92%	75%	42%
Miscue Analysis Accuracy Acceptability	96% 99%	95% 98%	94% 98%	97% 99%
Comprehension: Explicit Implicit	75% 50%	75% 50%	75% 75%	75% 75%

Note. Time 1 (pre-instruction), Time 2 (after 10 weeks RT), Time 3 (after 10 weeks RTV), and Time 4 (1 year follow-up).

4.5.4.1. Miscue analysis for pupil 2.3 (B.)

Time 1

At Time 1 B. read quickly (110 wpm) but accurately. She corrected three meaning-change miscues, but she also corrected one miscue that did not change the meaning. The authors of the QRI indicate that this is a sign of a reader who is paying more attention to the individual words in the text, rather than the meaning. It did only occur once, but as it was a rare occurrence in the miscue analyses of the whole sample, it is worth noting here. She also missed out a whole sentence and part of another, which also indicated a lack of reading for

meaning. A further example of an uncorrected miscue was when, instead of reading: *Animals that feed on **trees lose** their food supply,* B. read: "Animals that feed on **loose trees** the food supply". This again, did not make sense. Therefore, although the total acceptability score was high, those mistakes she did make would have had a significant effect on her understanding.

Time 2

B. read quickly (98 wpm, although not quite as fast as at Time 1). Her eight meaning-change miscues were corrected four times (e.g. *swarms* was changed to "swims") but four errors were left uncorrected (she read "tell" for *tells* which was grammatically incorrect, "arm" for *arms*, and *unlike* was mispronounced). The gravest error occurred when B. lost her place, and reread some lines again without seeming to realise she was repeating herself. This repetition may have contributed to the lower reading rate, rather than it being as a result of reading more slowly overall. The passage as printed was:

> *The octopus bites into the crab with its strong beak. This sends poison into the crab's body. The octopus protects itself in three ways. First, when frightened, the octopus can push water away from its body in a powerful stream. This action pushes the octopus forward very rapidly. This allows it to escape.*

B. read:

> The octopus bites into the crab with its strong beak. This sends poison into the crab's body. The octopus protects itself in three ways. First when frightened the octopus can push with its beak. This sends poison into the crab's body. The octopus protects itself in three ways. First when frightened………

By letting her eye travel back to part of a sentence previously read, B. read two conflicting statements about how the octopus uses the water to make itself move rapidly; firstly she read that it did so by using its beak and then she read the correct way of pushing the water away with its body. She did not make any verbal attempt to resolve this conflict, and her answer to the comprehension question about how the octopus moves forward suggested she was unclear about how this happens – she answered that "it pushes the water back" - but having read that it does this by using its *beak,* and then its *body*, the exact mechanism would have been uncertain. Although the acceptability level was high at 98%, these uncorrected errors did show a lack of understanding in some instances and a lack of comprehension monitoring.

Time 3

B. read this passage at a much slower rate than at Times 1 and 2, (81 wpm). B. made 11 meaning-change miscues, but she corrected 5 of them. She missed out part of a sentence, but she backtracked and re-read so it made sense (in contrast to Time 2 when she just carried on reading). Three of the uncorrected changes concerned mispronunciations, where a non-word was substituted for a word (for *reattach* and *retina*) or *uses* where the emphasis was placed in the wrong place. These errors did show a certain lack of concern with meaning as no attempt was made to make sense of the sentence: *Doctors use laser beams to reattach the retina*.

Time 4

B. read at speed (119 wpm) but with a high level of accuracy. She made eight meaning-change errors, but self-corrected five of them. The other three were missed word endings, which, whilst not grammatically correct, did not dramatically alter the sense (e.g. "harvest" for *harvested*).

4.5.4.2. A Summary of think-aloud responses and strategy interview for pupil 2.3 (B.)

Time 1

B. was able to summarise parts of the text, but she found explaining strategy use difficult, and she used a limited range of strategies (just summarising and visualising were used correctly). She did not seek to clarify any unfamiliar words in the text, or question any of the ideas. The text was unfamiliar, but she did not seem to have acknowledged that she had learned anything from what she had read. B's detailed responses at all four time points are given in Appendix S.

Time 2

B. was only able to use a general connection from the text and an incorrect summary. She did not always monitor her comprehension, and although she had excellent prior knowledge she was unable to answer all the comprehension questions correctly. She did know about the four characters of RT, but she did not apply them to help her use the strategies effectively. She did use visualisation as a strategy, although again she did not appear to be utilising it to monitor her understanding.

Time 3

B. used a much greater range of strategies than at Time 1 and Time 2. She attempted to summarise what she had read and went beyond repeating sections of the text to paraphrase

166

in her own words. She attempted to connect her background knowledge with what she had learnt in the text and was much more aware of comprehension monitoring and why it is important. B. did not refer to the characters by name but she showed she was capable of applying the strategies they teach. Interestingly, despite all the work she and the class had done on visualising she maintained that she did not make any pictures and that she never made pictures when she read. This was not the case with the think-aloud, where she had spontaneously described a picture she had made of the octopus. Perhaps she was again showing her initial resistance to being shown what to do when reading.

Time 4

B. used a much wider variety of strategies - predicting, clarifying, questioning, summarising and visualising, were all in evidence. She was also much more able to explain how she used the strategies and how important it was to monitor her comprehension. The think-aloud responses were more detailed and, together with the interview comments, provided evidence of an active reader who questioned the text and her own understanding of it. B. appeared to have overcome her resistance and adopted many of the strategies taught. It is interesting to note that she was the only child to show an increase in comprehension score from Time 3 to Time 4. The instruction may have undermined her confidence in her reading ability initially and maybe she had gradually come to realise that to read well means to understand and not just read the words. Certainly her interview responses showed her awareness of this had grown over the instruction period and the follow-up.

4.5.5. Pupil 2.4 (C.) : a poor comprehender

C. enjoyed reading fiction, especially books about fairies. She was always keen to read to an adult. With a YARC comprehension standardised score of 90, C. had the lowest score out of the ten pupils (M=101.30, SD = 7.80). Her single word reading scores were below the mean for the group, but she did not have the worst TOWRE scores. Table 21 shows C.'s scores relative to those of the ten children in the study. Indeed her phonemic decoding skills were average for the group and above average for her age.

Table 21. Scores on the TOWRE for pupil 2.4 (C) in comparison with the children in Study 2.

Time 1	Pupil 2.4	Study 2 M	Study 2 SD	Study 2 range
Sight word reading	99	110.50	8.90	98-126
Phonemic decoding	105	105.60	9.12	91-118
Total word reading efficiency	102	109.50	10.55	93-126

C.'s reading comprehension score was below that to be expected given her word reading scores. Table 22 shows her scores relative to those of good and poor comprehenders, as defined by Cain (2010, p.150). C. can therefore be considered to have been a poor comprehender at Time 1.

Table 22. Typical characteristics of good and poor comprehenders (Cain, 2010) and pupil 2.4 (C.).

Time 1	Good comprehenders	Poor comprehenders	Pupil 2.4 (C.)
Chronological age	7;07	7;07	7;07
Word reading	7;09	7;09	7;08 (sight word)
(age equivalent score)			8;08 (phonemic decoding)
Reading comprehension	8;01	6;07	6;04
(age equivalent score)			

Note: Chronological age and age equivalent scores are given in years, months.

Figure 18 shows her word reading scores over the duration of the study, along with plots of her strategy use from the think-aloud and from the interview, whilst figure 19 shows her standardised reading comprehension score and selected comments from the think-alouds and strategy interviews.

Figure 18. (a) TOWRE standardised score for total word reading efficiency (b) number of strategies reported in the think-aloud and (c) strategy interview score, for pupil 2.4 in Study 2 at each of the four time points.
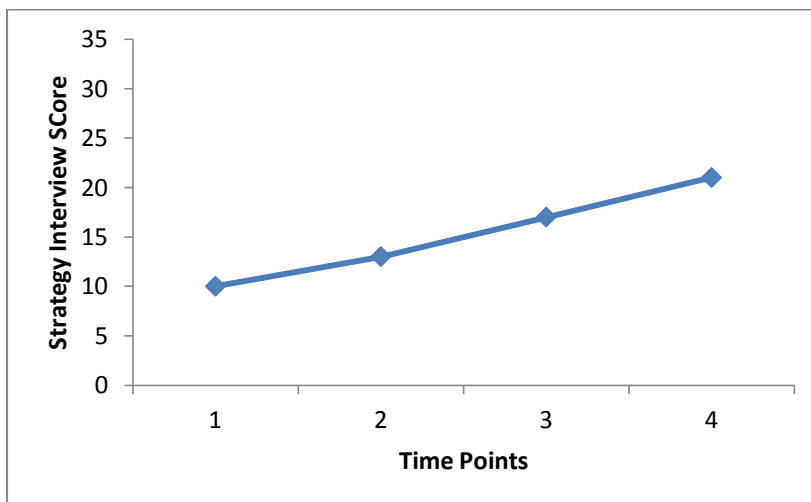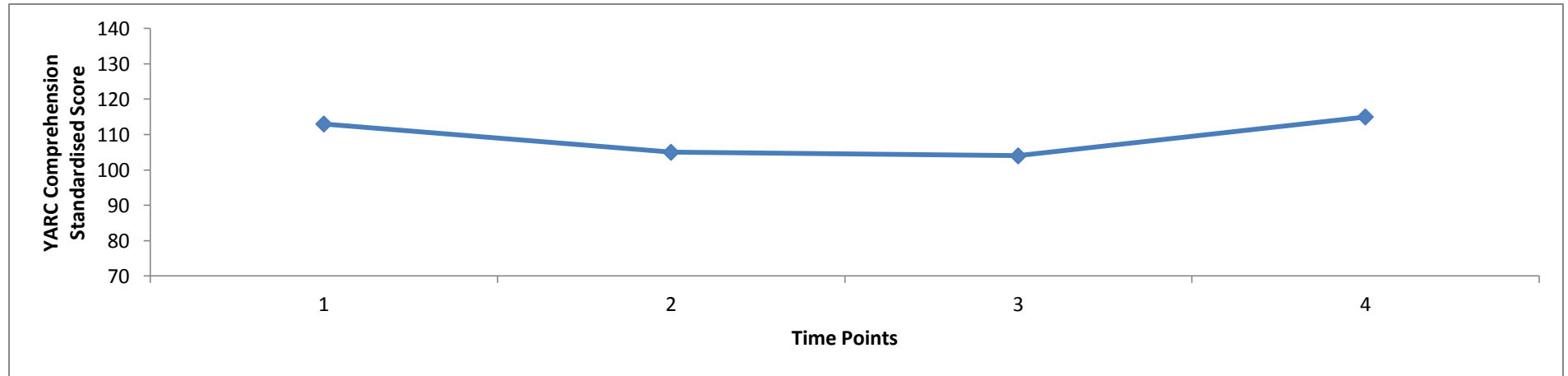
(a)
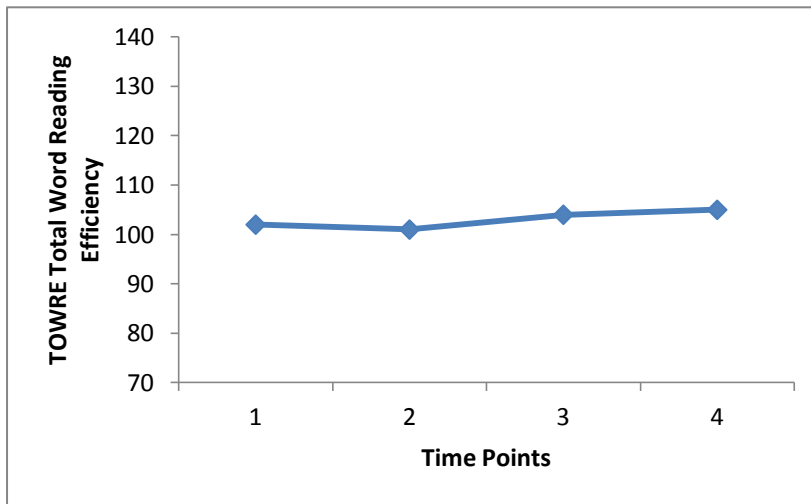
(b)
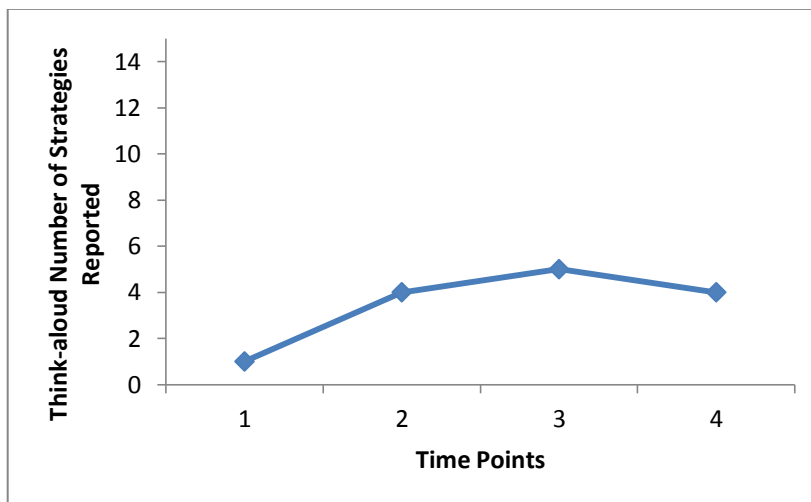
(c)

Figure 19. Standardised scores for YARC comprehension together with think-aloud and strategy interview comments at four time points for Pupil 2.4 (C.).

Whales and fish live in the water. *(Repetition of a part of the text)*

It's all about where people live. Like I live in a village. *(A summary of the text with a personal connection)*

The beaver is swimming under water to go inside its home. (*An inference from the text*)

A suburb might be a village or another name for it. (*An attempt at clarifying*)

Not sure. *(No response)*

I could use a dictionary…read on…find out (*Names strategies to work out the meaning of a word*)

The main problem was understanding the ideas. I could read it again. *(A move away from a word level emphasis, with a strategy to solve a problem)*

You could try and say what you think is right. *(Focus now on the word level)*

Table 23 shows the QRI passages read and the scores obtained. The comprehension scores reflect a similar pattern to that seen in the YARC, with an increase from Time 1 to Time 3, and a slight decline to Time 4, the one year follow-up.

Table 23. The QRI passages and scores obtained at the four time points for Pupil 2.4 (C.).

QRI	Time 1	Time 2	Time 3	Time 4
Passage	Level 2: Whales and Fish	Level 3: Where Do People Live?	Level 4: The Busy Beaver	Level 3: Where Do People Live?
Familiarity	22%	42%	42%	50%
Miscue Analysis Accuracy Acceptability	 97% 97%	 97% 98%	 93% 97%	 97% 98%
Comprehension: Explicit Implicit	 50% 25%	 50% 75%	 75% 75%	 50% 75%

Note. Time 1 (pre-instruction), Time 2 (after 10 weeks RT), Time 3 (after 10 weeks RTV), and Time 4 (1 year follow-up).

4.5.5.1. Miscue analysis for pupil 2.4 (C.)

Time 1

C. read at 74 wpm, which is in line with the speed expected for her word reading ability. She made six meaning-change miscues, but only corrected one of them. Some of the errors showed she was not reading for meaning; for example she read: "Mother whales give birth to **lĭve** whales", instead of: *Mother whales give birth to līve whales*)". She read "Whales and fish both **like** in the water" when she should have read *live in the water*. There was no difference between her word accuracy score and the total acceptability score, underlining the lack of corrected miscues.

Time 2

C. read more quickly than at Time 1 (93 wpm). Her good word reading skills meant she could read this passage with few errors, but this time she did self-correct four out of the nine mistakes she made, thus improving her acceptability score compared to her word accuracy.

She corrected "they also have many **more** theatres" to what was in the text - *they also have many movie theatres*. The error was a predictive miscue; C. was expecting the word *more* to come after *many* as 'many more' is a common expression. The QRI authors are of the opinion that such predictive miscues "suggest attention to meaning at least at the sentence level" (p. 66). Another corrected miscue was "**there** live in suburbs", to "**they** live in suburbs", which was again necessary to preserve meaning.

Time 3

C. found this passage more difficult to read than the passages at Times 1 and 2. This passage is at Grade 4 level, which is for a chronological age of 9 to 10, compared to C.'s age of 8 years 4 months. She read more slowly (64 wpm) and made 19 errors. This was still within what the QRI authors have calculated to be an instructional level (frustration is not realised until more than 30 errors). Of these 19 errors, 13 were meaning-change, and of these 5 were self-corrected. The first sentence of the text is: *Have you ever heard someone say 'busy as a beaver?'* C. initially read: "Have you ever heard **of** someone say", which was a predictive miscue. C. showed she was reading for meaning when she reread the sentence to make sense of the ungrammatical first attempt.

Of the errors that went uncorrected, four were minor word ending discrepancies (e.g. "changes"/*change* and "part"/*parts*) and only two represented a real challenge to the comprehension of the passage. Both of these were unfamiliar words. The first was *elsewhere*, at which she made an attempt, but then declared "I'm not quite sure what that says", which showed her awareness that the word was problematic. The second instance was a mispronunciation of the word *canals*. C. used a nonsense word, which showed a lack of monitoring at that point. However, although the passage as a whole was challenging for her at the word level, her increase of 4 % for acceptability over accuracy shows she made a good attempt to read for meaning.

Time 4

Although C.'s single word reading standardised scores suggested that she still had above average word reading skills (TOWRE sight word 102, phonemic decoding 104, total word reading efficiency 105) she was not able to read enough words correctly to repeat the level 4 passage she read the year previously, and level 3 was administered. Level 3 is appropriate for children aged between 8 and 9, and C. was 9 years and 4 months at the time of this assessment. Owing to an administrative error she was given the same text to read as at Time 2, instead of an alternative level 3 passage as was intended. However, as there were 16

months between Time 2 and Time 4, and as there was not any element of discussion about whether answers were correct or not, it was unlikely to have affected her responses.

C. made fewer errors this time, ten in all - nine of which were meaning-change miscues. Of these nine miscues, four were corrected (e.g. initial mispronunciations of *suburbs* and *offices*) which showed she was reading for meaning. However, five remained uncorrected, for example, C. read: *"**There** live in suburbs"* which did not make sense, and "the suburbs are **quite** than the city" instead of *quieter.* This may indicate that the level of her requirement for text coherence had fallen compared to Times 2 and 3. She read at 93 wpm which was the same speed as at Time 2.

4.5.5.2. A summary of think-aloud and interview responses for pupil 2.4 (C.)

Time 1

C.'s lack of response during the think-aloud despite extra prompting and her lack of knowledge about strategies during the interview suggested that she was not an active reader. She failed to reread sentences which did not make sense suggesting a low tolerance for meaning making. She did visualise, but her description was very brief and only added one word to those from the text (*blue*). Although a lack of response could be due to nervousness, C. did not appear nervous and read the QRI word lists with confidence. She did not appear withdrawn, just unable to answer questions about what she was thinking, or to have any suggestions to make about strategies other than visualising and an incorrect prediction. C.'s detailed responses at all time points are given in Appendix T.

Time 2

C. was reading far more actively. She corrected for meaning when reading aloud and drew on her personal experience to identify with the text. She was able to talk at much greater length about what she was thinking and showed that she could predict, clarify, question and summarise (to a limited extent) . She was able to visualise at Time 1 and continued to use this as a strategy but she had not developed it further.

Time 3

C. used all the RT strategies – predicting, clarifying, questioning, summarising and visualising. She was able to explain how these strategies helped with comprehension monitoring, and her use of visualisation had developed into moving images which could be used to repair comprehension. Her more active engagement with the text had resulted in a higher than

expected score for comprehension, which was a huge improvement from the Time 1 difference between word reading and comprehension skills.

Time 4

C. showed she can use predicting, questioning, summarising and visualising, but her use of clarifying only went as far as identifying an unfamiliar word without going on to resolve its meaning. She was more active as a reader than at Time 1, but she did not appear to be monitoring her comprehension as carefully as she was at Times 2 and 3. This was reflected in her standardised scores for comprehension obtained on the YARC, which showed an increase from Time 1 (90) to Time 2 (93) to Time 3 (99) but a slight decrease from Time 3 to Time 4 (96).

4.5.6. Pupil 2.9 (D.): a good responder- and also a poor comprehender

D. was a very articulate pupil, with a very good general knowledge. He excelled at oral work but he was more reluctant about reading and writing. Table 24 shows how D. compared with the poor comprehender profile presented by Cain (2010). In common with C. he also fitted the profile, with a year's difference between his word reading and comprehension skills.

Table 24. Typical characteristics of good and poor comprehenders (Cain, 2010) and pupil 2.9 (D.).

Time 1	Good comprehenders	Poor comprehenders	Pupil 2.9 (D.)
Chronological age	7;7	7;7	7;1
Word reading	7;9	7;9	8;4 (sight word)
(age equivalent score)			7;9 (phonemic decoding)
Reading comprehension	8;1	6;7	6;1
(age equivalent score)			

Note: Chronological age and age equivalent scores are given in years, months.

174

Figure 20 shows D.'s word reading scores over the duration of the study, along with plots of his strategy use from the think-aloud and from the interview, whilst figure 21 shows his standardised reading comprehension score and selected comments from the think-alouds and strategy interviews.

Figure 20. (a) TOWRE standardised score for total word reading efficiency (b) number of strategies reported in the think-aloud and (c) strategy interview score, for pupil 2.9 (D.) in Study 2.

(a)
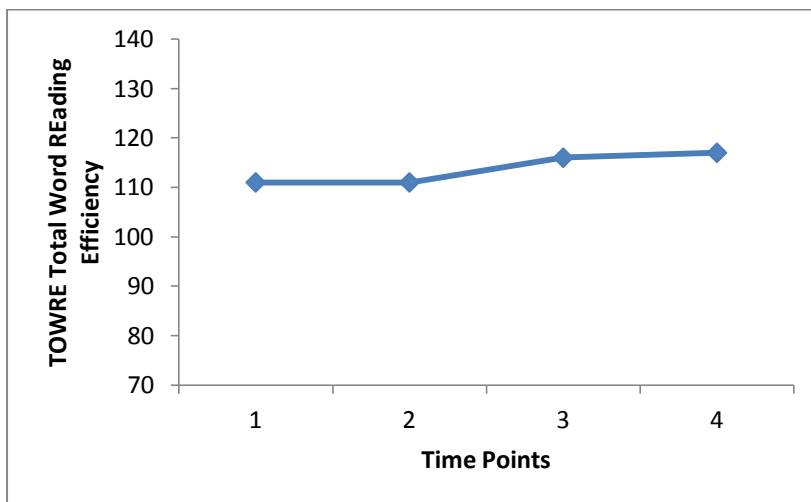
(b)
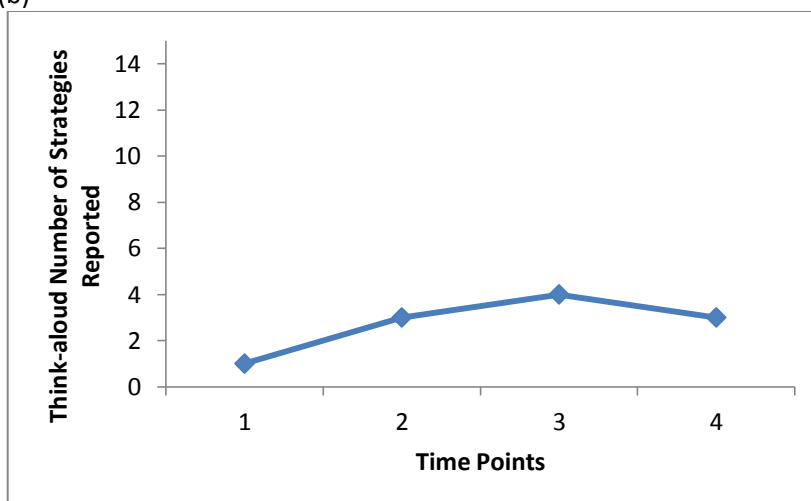
(c)
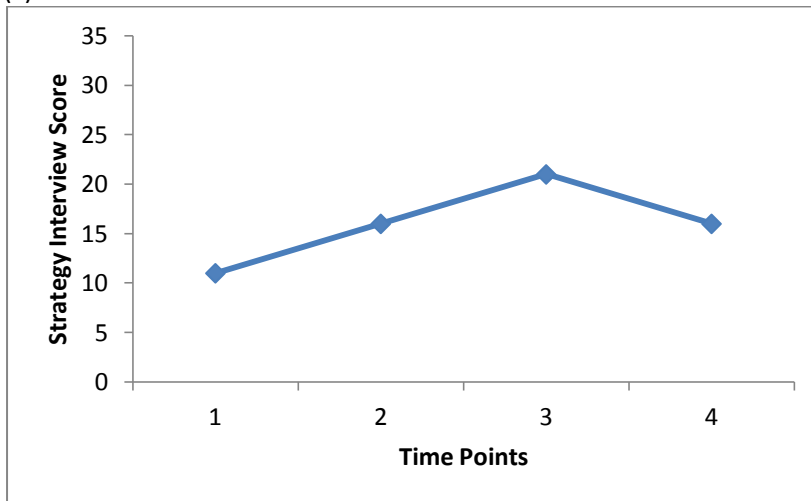
Note. Time 1 (pre-instruction), Time 2 (after 10 weeks RT), Time 3 (after 10 weeks RTV), and Time 4 (1 year follow-up).

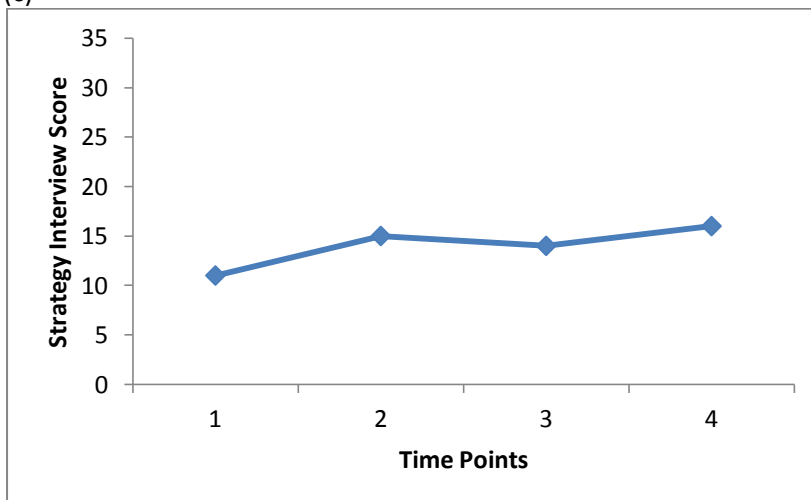*Figure 21.* Standardised scores for YARC comprehension together with think-aloud and strategy interview comments at four time points for Pupil 2.9 (D.).
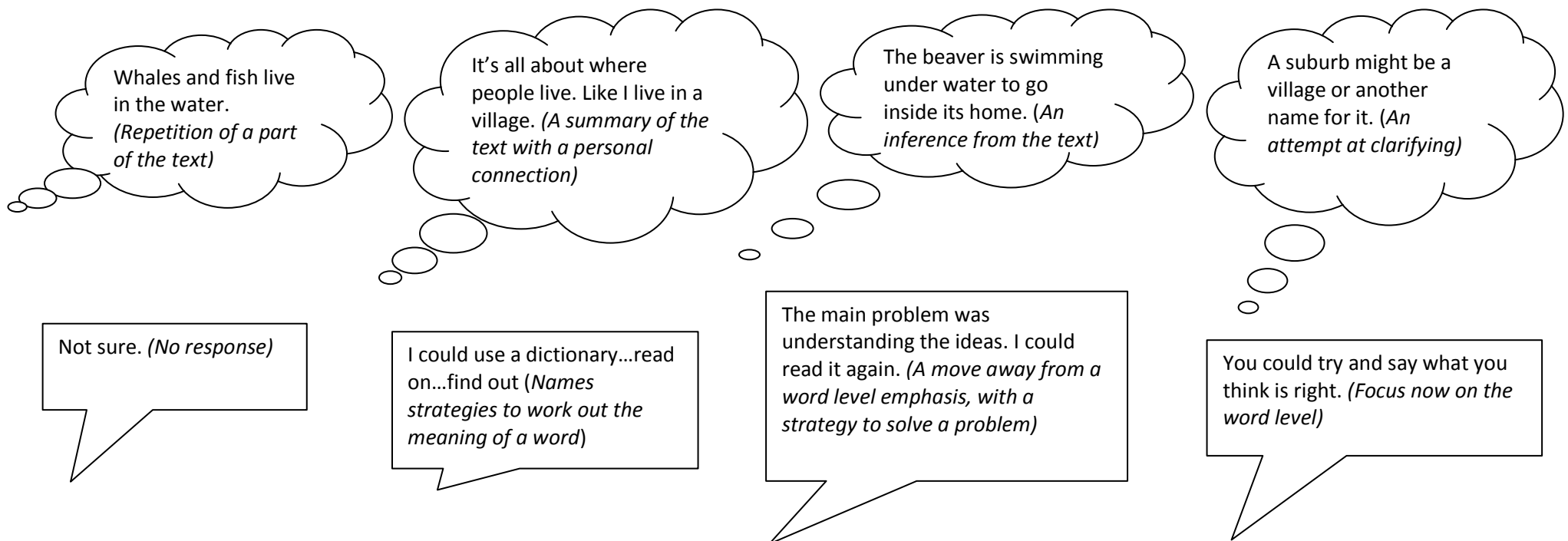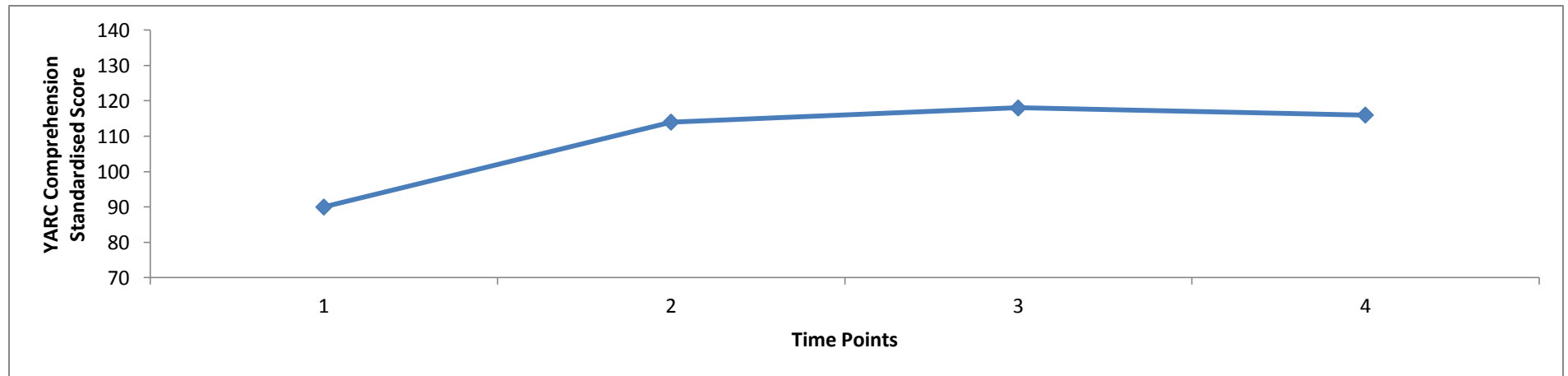

(Q. Did that make you think about anything?) No. *(No response)*

I thought about the squirrel 'cos we have some sort of birds and we put nuts out but the squirrels eat them. *(A comment on something in the text, linking it to a personal experience)*

It remembered me about my cat's tongue 'cos it's rough when he licks me. Now I know why. (*A comment on something in the text, linking it to a personal experience and acquiring new knowledge)*

It reminds me of the school lake and someone standing on the lily pads. (*Linking the known with the unknown by using imagery)*

If you don't know a word you can split it up. *(Decoding emphasis)*

You could make up your own pictures. (*A strategy beyond the word level*)

Get all the things happening in the story and combine them up into one important thing and make a picture of that. *(Using a strategy at the text level to remember what is important)*

Saying the words was the hard bit. *(Focus now on the word level)*

Table 25 shows the passages read for the QRI and the scores obtained. The comprehension scores on the QRI reflect those on the YARC, with an increase from Time 1 to Time 2, and a decline from Time 3 to Time 4. However, ceiling effects may mask any increase between Time 2 and Time 3. The same thing can be said for the total acceptability scores on the miscue analysis.

Table 25. The QRI passages and scores obtained at the four time points for Pupil 2.9 (D.).

QRI	Time 1	Time 2	Time 3	Time 4
Passage	Level 1: Air	Level 2: Seasons	Level 3: Cats: Lions and tigers in Your House	Level 4: Plant Structures for Survival
Familiarity	67%	42%	50%	75%
Miscue Analysis Accuracy Acceptability	95% 96%	98% 100%	96% 100%	97% 98%
Comprehension: Explicit Implicit	75% 25%	100% 100%	100% 100%	50% 75%

Note. Time 1 (pre-instruction), Time 2 (after 10 weeks RT), Time 3 (after 10 weeks RTV), and Time 4 (1 year follow-up).

4.5.6.1. Miscue analysis for pupil 2.9 (D.)

Time 1

At Time 1 D. read at 80 wpm, which is slightly faster than the expected rate at an instructional level; according to the QRI authors (p.70) the range at this level is between 37 and 77 words per minute. Therefore, his word reading was not so effortful that it could be hindering his comprehension. He made five meaning-change miscues and did not correct any of them. He read a sentence that did not make sense, reading: "Strong **things** can **ever** move a house' in place of the text: *Strong **winds** can **even** move a house.* Additionally he read "We can **weight** two balloons", instead of *weigh* despite having read the word correctly just four words before.

178

Time 2

D. made two meaning-change miscues, but he corrected them both. He was unsure of the pronunciation of the word *tulips* but he made several attempts until he was satisfied it was correct. He initially read: "We have less light that day than **or** any other day", but he corrected *or* to "on" when he realised it did not make sense. The other two miscues were very minor and did not affect the meaning (an additional "the" and "that" replacing *the*).

Time 3

D. made 10 miscues, but his speed of reading (90 wpm) again fell with the range expected for an instructional level of text (56 to 104 wpm) indicating that the text was not too difficult in terms of decoding. Additionally, up to 27 miscues are allowed before the text is at frustration level according to the QRI guidelines. Of these miscues, only four were meaning-change and of these, D. corrected three. The only meaning-change miscue that remained was when D. lost his place and missed out a phrase. He was conscious of having lost his place – he said *"Oh no"* when he omitted two words and it did not make sense. He then went back and re-read, again missing out several words; what he read did make sense, it was just not the complete sense required for the text (he did score 100% acceptability after rounding up however).

Time 4

D. made eight miscues, seven of which were meaning-change. He corrected two of them, but five remained uncorrected, meaning a much smaller gap between accuracy and acceptability and therefore a lower standard of coherence than at Times 2 and 3. Three of the miscues were mispronunciations, but two of these he did repeat several times, with a questioning inflection, showing his uncertainty. The other two uncorrected miscues were substitutions ("shall" for *shallow* and "observe" for *absorb*) which did not make sense.

4.5.6.2. A summary of think-aloud and interview responses for pupil 2.9. (D.)

Time 1

D. did not show any evidence of being an active reader. He had a low standard for coherence and did not combine his background knowledge with facts in the text that were inconsistent with his beliefs. He did make pictures in his head but they were general associations rather than comprehension fostering or monitoring. His focus was at the word level and his comprehension suffered as a result. D.'s detailed responses can be found in Appendix U.

Time 2

D. was reading much more actively that he did at Time 1. He re-read when something did not make sense and he clarified an unfamiliar word. He had learned from the text and combined what he knew before with new knowledge about the seasons. He answered all the comprehension questions correctly. D.'s scores at Time 1 and Time 2 on the TOWRE and the YARC score for accuracy remained the same, at 111 and 106 respectively. Two different sources therefore confirmed that D.'s word reading ability had not improved and could not therefore have been responsible for the improvement in comprehension score.

Time 3

D. continued to show he was more active in reading; he self-corrected the text and thought about meaning as he read (as shown by his reflections over the word *rough).* He had learned from the text too, shown not only by answering all the questions correctly, but by coming to understand the meaning of *cat family*, a concept with which he was not familiar before reading this passage. D. had shown he used visualisations at Time 1 and 2, so this was a strategy with which he was already familiar - however, his visualisations were more detailed and used new elements, such as describing things in his own words rather than the words of the text, and describing the background of the picture as well.

Time 4

D. was reading actively, but his use of visualisation and background knowledge had become more generalised since Time 3. If the think-aloud were being used by the class teacher as an assessment of strategy use, this would enable D. to be reminded about the best ways to use visualisation as a comprehension monitoring strategy, whilst not detracting from the rich visual imagery he obviously used instinctively as he read.

D.'s level of comprehension, although markedly better than at Time 1, was not as good as at Time 3. This same pattern is reflected in the standardised scores for comprehension on the YARC reading test, of 90 at Time1, 114 at Time 2, 118 at Time 3 and 116 at Time 4. This may because he was no longer questioning the text in the same way, and because he had become more concerned with the word level when decoding was a problem. D. could be reminded to focus on a higher level for coherence.

4.6. Discussion

The aim of this second study was to investigate the changes which take place in strategy use when RT is used in a whole class situation, with children aged 7 to 8, in the UK. A supplementary aim was to see whether the addition of visualisation might increase the effectiveness of traditional RT. The study also investigated the delivery of RT by a class teacher rather than the researcher, and included a follow-up assessment to look for the maintenance of any gains.

The design involved pre-test assessment, first intervention, post-test assessment, second intervention, post-test assessment, and a one year follow-up assessment. A standardised test of reading comprehension, complemented by think-alouds and strategy interviews were employed. The results revealed a significant improvement in reading comprehension scores across the period of instruction for the group of ten children for whom data were obtained. Reading comprehension standardised scores, the number of strategies used, the number of comments made and the strategy interview scores all showed a pattern of increase from Time 1 to Time 2 to Time 3 and then a decline to Time 4. The most significant increases were from Time 1 (pre-test) to Time 2 (after 10 weeks of instruction). These results may indicate that by becoming more active readers (i.e., making more comments and using more strategies, as reflected in the think-aloud data, and knowing more about strategies and how to use them, as reflected in the strategy interview data) the children improved their comprehension ability. These improvements in reading comprehension were not brought about by improvements in word reading ability, since there was no significant difference in total word reading efficiency as measured by the TOWRE at each of the four time points.

The differences in accuracy and reading rate assessed with the YARC were not significant, but the small sample size means it is difficult to detect small effects. Study 3 included a larger number of participants to address this issue. It would be expected that more active readers would make less errors in reading connected text as they would be using the context to help with any deficiencies in word reading ability. In Study 1 there was some indication that children slowed their reading rate, and it was considered this may have been as a result of making more self-corrections. The children in Study 2 did make more self-corrections but there was no indication of slower reading rate across time, although analysis with ratio gain did show that they increased in reading rate across time by less than would be expected for normal development. The effect may not have been statistically significant,

181

but as Brooks (2002, 2007, and 2013) maintains there may still be differences which are significant educationally. Conversely, there may also be times when changes which are statistically significant do not seem to have educational significance. For example, in a recent study (Clarke, et al., 2010) contrasting oral and text base approaches aimed at ameliorating comprehension difficulties in Year 4 children (aged 8 to 9) the results section reports that on the standardised vocabulary test used (WASIWechsler, 1999) the Oral Language group showed a significant gain compared with the control group at Time 3. Furthermore, it was reported that neither of the other two intervention groups (a text comprehension and a combined group) made reliable gains. However, since Clarke and her colleagues have also reported the raw scores for each time point it is possible to see that the difference in gain between the Oral Language and control group at Time 3 amounted to just 1.51 of a mark out of a possible 64, after an intervention lasting 30 hours. It would appear that statistical significance can be used to justify the effectiveness of an intervention when the raw scores show that the gains achieved are very small. Indeed, whilst arguing for the superiority of the Oral Language approach over the text based approach, the difference in the two interventions at Time 4 (after 20 weeks of 3 x 30 minute lessons) was just 1.13 of a mark out of 64 on the WASI.

4.6.1. Why was the biggest increase between Time 1 and Time 2?

Given that one of the aims of Study 2 was to determine whether RT could be improved by the addition of visualisation, it was expected that the biggest improvement would be observed between Time 2 and Time 3, after visualisation had been added to the instruction. However, it would appear that a number of children were already using visualisation as a strategy which may have limited the effect. Four children visualised during the think-alouds at either Time1, Time 2, or both, which was before the strategy was taught. However, five children used visualisation for the first time at either Time 3 or Time 4, suggesting the teaching was effective. There was also an increase in the mean score for visualisation in the strategy interview (Time 1 M=2.75, Time 2 M=2.62, Time 3 M=3.37, Time 4 M=3.12) again suggesting children were more aware of the strategy and better able to explain how and why to use it.

Since it has been argued that RT works by encouraging active reading through participation it may be that the greatest effect comes after the introduction of the idea of strategic reading and participation in activities promoting reading in that way. It is not so much the strategies themselves that are important, but the way children are encouraged to

read as engaged and active readers. Cain (2010) suggested this by saying that inference training may work by teaching children to focus on content (p.185) and by helping children evaluate their understanding. It is the contention here that RT achieves an improvement in comprehension monitoring and comprehension fostering in the same way.

4.6.2. Why the decline at the one year follow-up?

Between the end of the intervention in May of Year 3 and the follow-up testing in Year 4 twelve months later, the children did not receive any RT or specific strategy instruction. The comprehension instruction they received followed the normal pattern of the school, that is, by means of comprehension exercises from a published scheme completed as homework, and weekly Guided Reading sessions in ability groups. However, the amount of declarative and procedural knowledge retained was manifest in the reading interview scores, which had not declined. It seems to have been the conditional knowledge or maybe the incentive to use the strategies which had declined, which is not surprising perhaps given the fact that their use was no longer being encouraged.

4.6.3. Implementation of Reciprocal Teaching

An important part of this study was the intention that the class teacher would implement it in the same way that a teacher might after attending an in-service training session. The teacher would be introduced to the strategies and the teaching method, they would be provided with some materials, and then be expected to go away, do some reading for themselves, and then introduce the programme to their class making adaptations as they went along. This was very much the way the implementation proceeded in this study. The class teacher made more use of Oczkus (2003) than the original manual (Palincsar, et al., 1989) in that she allocated roles in each group session. However, through examining her plans and the observations of the researcher it is possible to say that she followed the criteria checklist for the quality of the description of RT from Rosenshine and Meister (1994) which was the intention. Thus, the students were instructed in a repertoire of strategies: the four strategies of predicting, clarifying, questioning, and summarising between Time 1 and Time 2, and these four strategies plus visualisation between Time 2 and Time 3. The teacher modelled each of the activities, and additionally, during the plenary sessions children modelled them to the whole class. Children were invited to make comments regarding the modelling and the passage, and this occurred both between the teacher and the class and between children in their groups. Students were provided with guided assistance, which

came from the teacher and the teaching assistant. The teacher and the teaching assistant also supported participation in dialogue about the passage, through feedback, praise, prompting, additional paraphrasing, coaching, hints and explanations, and encouraged children to initiate discussion and through their role as the 'teacher' to provide feedback in a similar manner.

During the year the classroom teacher gradually shifted from doing much of the work to the children taking over the major role. This was achieved through a transition from explicit teaching, to modelling, to group work where the role of the pupil teacher became that of the class teacher. Finally, during the dialogues, the teacher explained why, where and when these strategies might be applied. And where a particular text did not warrant the use of a particular strategy she did not allocate that role.

4.6.4. The challenge of group work

As we have seen from the teacher's evaluation, getting children to work in groups is not always easy. Keeping several groups on task simultaneously places demands on the teacher and the children. Previous studies of RT which involved the whole class have highlighted the same issue. Hacker and Tenent (2002) investigated teachers' implementation and practice of RT in their qualitative study in two elementary schools. They reported that all seven of the teachers in their second study encountered problems with getting their students engaged in group discourse, and that disruptions were regular occurrences. Some of the teachers made modifications to their implementation to try and prevent the amount of time off-task; for example, one teacher set homework on the passage the night before, which required the students to read the passage and compile five questions arising from it. Three teachers used whole-class instruction, with one saying "I don't particularly like leaving the kids 'in charge' every time" (p.706). One teacher even abandoned the RT altogether because her students had difficulties with the co-operation and collaboration needed for the RT groups to work. Their third study encountered similar problems, with one teacher noting that her students were often off-task, and another saying that their students did not participate fully. To counteract this, several teachers in the third study used a mix of whole-class instruction and RT groups work, with introductory and plenary whole-class work interspersed with work in groups. This was the method adopted by the teacher in Study 2, since she was following the format of the Literacy hour, which had been instigated in the school previously. Hacker and Tenent provide data to show that RT was effective in all three of their studies, but as they give aggregated figures we cannot distinguish which adaptations were more successful, and

184

whether the changes made to facilitate group work had any effect on comprehension gains. The authors do note however, that given the difficulties encountered in getting children to collaborate, then it may be necessary to provide training in basic discourse skills as a prerequisite to successful teaching. Mercer and Littleton (2007) contend that "there is not enough emphasis in educational policy and practice on the value of teaching children how to use language for learning" (p.2) and that by teaching children the language necessary to think together, they can not only work together effectively but they will learn better ways to think alone. In Study 2 in this thesis, although the class teacher mentioned the difficulties in getting children to work in groups independently initially, the children did achieve it with practice. It was therefore of interest to see how the management of small groups within a whole class environment was approached by the teacher in Study 3.

4.6.5. A discussion of the four case studies

The four case studies were selected from plots showing the difference in comprehension scores in the YARC across the study. They were selected to examine changes in strategy use for pupil 2.10 (A.) a good reader, pupil 2.3 (B), a poor responder, pupil 2.4 (C.), a poor comprehender and pupil 2.9 (D), a good responder. The changes in strategy use were shown to be linked to the improvements in comprehension scores, and that as children read more actively and were able to use and even explain their strategy use, their reading comprehension also improved. When comprehension ability declined (as in the case of pupil 2.3) it was suggested that this was because of a certain resistance to changing the way reading as an activity was perceived. This resistance and the theory behind a possible explanation are explored in the next section.

4.6.6. How does the learning of strategies take place?

RT, as we have seen, takes the teaching of Vygotsky to explain how the use of strategies in reading can be internalised by exposure to explicit teaching and modelling. However, since a large part of the programme is not just exposure to the strategies but the practice of the strategies by the children, and their participation in giving feedback to the teacher and their peers, I was drawn to the idea of the relationship between learning and the social situation in which it occurs (Lave & Wenger, 1991). Thus, during RT, children become part of a community of practice, and through 'legitimate peripheral participation' they gradually become a full participant. Lave and Wenger's account of cognitive apprenticeships also draws our attention to a member who chooses to resist the practices of the community, and

this seems to have been the case for B. She seemed to have firm ideas about reading before the instruction began. Her conception of reading as being based on proficiency in decoding and her measure of success being the early attainment of the status of a 'free reader' had served her well in the past and she was initially reluctant to change these ideas. There were many parallels with B.'s reluctance to think rather than read (a distinction she drew herself) and Lave and Wenger's characterisation of learners who find it difficult to move from their own preconceptions to participate in those of the group. Thus the teacher's notes about B. throughout the first 10 weeks of the instruction period say she reads quickly and fluently in her group, that she is very quick to correct any errors others make when reading aloud and she does not allow for any group discussion when she is the teacher. However, by the end of the second instruction period the teacher noted: "B. working hard and trying to encourage the others to join in". B. has gradually come to realise that by participating as a group the text can be read and understood more effectively, and that just reading it aloud herself may not be enough.

After observing this fit between what I had seen taking place in the RT sessions and the importance of participation as outlined by Lave and Wenger, I read an article which made the same connections (Davis, 2011). In this article Davis contends that although studies of comprehension strategy instruction have relied on the metaphor of internalisation in the past, a more appropriate metaphor might be that of participation. From this perspective, "learning to read strategically is a process of identity development in which students learn to become a particular kind of reader" (p. 102). Thus B. had learnt to become an active reader; but she had not merely learnt a set of strategies, she had adopted a strategic orientation as an important part of her reading identity (p.103). However, it took a long time to overcome her initial resistance owing to her own strong preconceptions and it would appear that she did not make these changes until the very end of the programme, and the results of this change were not seen until the follow-up at Time 4. Pupil 2.10 (A.) however, began with the mindset of a strategic reader, and was therefore more able to respond quickly to develop as an active reader, through participation in a programme teaching how to use strategies more effectively. He was able to add strategies to his existing repertoire as well as improve the way he uses those of which he was already aware. Pupil 2.4 (C.) and pupil 2.9 (D.) were both poor comprehenders at the outset, and were restricted in their strategy use. Their reading became more active and their comprehension improved over the period of instruction as they participated in reading strategically.

Davis (2011) concludes his article by saying that the participation metaphor is "a useful complement" (p.105) to the internalisation metaphor since it reminds teachers that social practice is important in strategy teaching, but that we also need to consider the strategies that children already have in place. The latter has been seen to be important in the light of case study 2.3 (B.) where a pre-existing conception of what makes a good reader conflicted with the aims of the instruction. Additionally, some children were seen to be using various strategies already at Time 1 – in particular visualisation – but even when visualisation was being used it could be improved upon, as we saw with case studies 2.10 and 2.9.

Davis also warns that strategies can sometimes be invoked for social uses. When participation in the method of the group becomes important, then members of that group may respond by saying they are using a strategy rather than actually using it because it is helping their comprehension. Davis says for this reason it is necessary to ask students about their individual strategy use. In this study, this was achieved by using the think-aloud method, in which the pupil does not have time to think of an answer because it is socially acceptable. If a think-aloud response is "I'm using Clarifying Clara", then that might fall into the category of a socially pleasing response, but when someone says "I don't know what fall means" and then tells you how they worked out the meaning, then they are demonstrating a strategy which they are using as an active reader, to improve their understanding of what they read. They are reporting what they are thinking rather than what they think you, or the group, would like to hear.

4.6.7. Differences between narrative and expository text

A previous think-aloud study (Kucan & Beck, 1996) examined the differences in fourth grade (aged 9 to 10) children's responses to narrative and expository text. They found that the narrative text elicited more recall and summarisation of important ideas, more inferencing, more predicting, and greater synthesis of incoming text information. In contrast, the expository text produced think-aloud responses which were "focused more on personal knowledge and experiences, providing commentary about or creating comparisons in response to details and more local text information" (p. 259). Thus it is perhaps to be expected that in these expository passages from the QRI, children made few, if any, predictions.

Differences in text genres were also detected by Zabruchy and Moore (1999) in their studies using texts with inbuilt inconsistencies. Rereading was more apparent in the

187

expository texts, which the researchers attributed to the greater importance of the ability to monitor comprehension and repair breakdowns in this kind of text. When reading narratives, since integration is easier, selective rereading may be less important (Janssen, Braaksma, & Rijlaarsdam, 2010). This may suggest that in these test passages it is more important that children reread passages that do not make sense and that when they fail to do so this is an indication of an inability to monitor comprehension successfully.

Roberts and Duke (2010) say that the question of whether instruction in reading comprehension can transfer to one genre to another is under-researched, but that the research that does exist suggests that transfer will not be complete. The texts used in the instruction programme in the present study were a combination of both, and the YARC uses a text of each in the assessment, but the think-aloud protocols only used expository text. The QRI texts used had to be carefully selected since they were taken from American literature and many contained ideas and themes unfamiliar to British school children. In order to have multiple texts at each level only the non-fiction texts were suitable. As Roberts and Duke point out, however, there are dangers in trying to teach genre-specific strategies. In the USA this has led to a focus on text features in expository text and the use of strategies has been neglected (Purcell-Gates, Duke, & Martineau, 2007). Indeed the authors contend that a difference in genre may even account for the fourth-grade slump (Hirsch 2003), in that young children are largely given narrative text to read and then in fourth grade (ages 9 to 10) they are suddenly presented with informational text, across a variety of disciplines, and it was assumed that since they could read (or rather decode) they would be able to understand what they read.

The use of texts

The inferencing texts used (included in Appendix P) may have helped the children to be more active in their reading, since there was much content that was not clear on an initial reading and the text had to be searched for clues. Yuill and Jocelyne (1988) showed that training children to look for clues in deliberately ambiguous texts brought about a significant increase in their comprehension of similar texts. The children in this study did seem to enjoy working out what was happening, and the texts provided good opportunities for devising questions, thus the same texts were used again in Study 3.

4.6.8. The role of vocabulary

As we saw in the Introduction, vocabulary has a role to play in comprehension. If a child does not understand the meaning of a word they will be hampered in their understanding, even if their decoding skills are good. A meta-analysis of training studies (Stahl & Fairbanks, 1986) showed that vocabulary instruction not only had a large effect on the comprehension of passages containing taught vocabulary (.95) as might be expected, but that there was also an effect on global measures of comprehension (.35). Beck and colleagues (1982) maintained that instruction in vocabulary affects comprehension in that it increases an awareness of vocabulary in general. Since improvements in vocabulary have not only been linked to an improvement in vocabulary (as the meta-analysis showed) but also to overall academic success (Baumann, Edwards, Boland, Olejnik, & Kame'enui, 2003) then any improvement in vocabulary knowledge is important. Since RT involves an awareness of vocabulary through the strategy of clarifying and since children are taught how to infer the meanings of words from context and previous knowledge, as well as asking someone or using a dictionary, it might be expected that their vocabulary knowledge will improve over the period of instruction. However, research on vocabulary acquisition in RT students is lacking (Mandel, 2007).

Vocabulary tests were used in the present study at Time 1, but owing to pressure of time these tests were only repeated for four children at Time 3. The tests used were the receptive and expressive vocabulary sub-tests of the Test of Word Knowledge (Wiig & Secord, 1992). There was an improvement in the standardised scores for receptive and expressive vocabulary from Time 1 (pre-instruction) to Time 3 (after 20 hours instruction) but paired sample t-tests showed that these differences were not statistically significant, likely due to the very small sample size (receptive vocabulary, Time 1 M=103.75$, SD$=7.5, Time 3 M=107.5$, SD$=1.7$,t$(3)=1.21$, p$=.31, expressive vocabulary, Time 1 M=112.00$, SD$=10.3, Time 3 M=115.00$, SD$=6.32, t(3)=.67$, p$=.55). However, given that an improvement in standardised scores was observed, this has practical implications. As we saw in chapter 2, Clarke et al. used a difference in raw score of 1.51 (out of a possible 64) to justify their claims of the superiority of the Oral Language approach. The effect of RT on vocabulary needs to be examined further.

4.6.9. Implications for Study 3

Davis (2011) contends that when learning is thought of in terms of participation, then the outcome of strategy instruction should not be just measured by reading tests but also by the

quality of strategic contributions during shared reading activities. The think-alouds produced during the QRI provide an opportunity to assess children's contributions, but the spontaneous comments made during the YARC provided a similar opportunity, which was examined further in Study 3.

The results of Study 2 indicated that RT can be effective in improving reading comprehension by encouraging readers to become more active, as suggested by the increases in strategy use and spontaneous comments. The case studies indicated that this can help poor comprehenders as well as good comprehenders, but one case study showed how a child who perceived herself as a good reader appeared to be resistant to change in reading, so in Study 3 additional evidence for this profile was explored.

Chapter 5: Study 3 - Investigating the changes in strategy use with children aged 7 to 8 in a different school

"You could feel your heart doing everything it wasn't doing before, 'cos your heart wasn't understanding it."

5.1. Research Aims

The aims of Study 3 were similar to those of Study 2. That is, the aims were to investigate whether RT could be an effective method of improving reading comprehension in a whole class situation, in the UK, for children in Year 3 (ages 7 to 8) and to investigate the changes in reading processes and motivation over the course of RT instruction, but with a larger participant sample than in Study 2. As in the previous studies a supplementary aim was to see whether the addition of visualisation might change strategy use.

The study also investigated the delivery of RT by a different class teacher from Study 2, and investigated the intervention's use in a different school setting. Finally, the study aimed to investigate whether RT and/or visualisation had any lasting effects on reading comprehension, strategy use and motivation.

5.2. Methodology

The third study, involving an examination of the changes in strategy use through think-alouds and reading interviews, took place in the academic year 2011-2012, with a follow-up at the end of the academic year in 2013. In common with Study 2, a test of word reading ability (TOWRE, Torgesen, et al., 1999), a reading comprehension test (YARC, Snowling, et al., 2011), think-alouds and a strategy use interview were conducted with each child, once before the intervention began and then at two time points during the RT intervention: ten weeks into the intervention and after a further ten weeks of training plus visualisation. In order to examine whether any changes in reading strategies/processes were maintained over time there was a fourth testing occasion, one academic year after the training finished. Again, as in Study 2, four children were selected as case studies to enable a more in-depth examination of any changes in strategy use. As in Study 2, an evaluation questionnaire for the children was used at Time 3, but in this study it was also used at Time 2. At Time 4 children were asked for their evaluation in a semi-structured interview. At Time 3, as in

Study 2, the teacher was asked to complete a questionnaire about her response to the intervention.

5.2.1. Participants

The participants were a Year 3 class (aged 7 to 8) in a different school to the first and second studies. This is a mainstream state school, for children aged 5 to 9. It is a voluntary aided school in the Home Counties, which was last inspected by OFSTED in 2009, when it achieved a Grade 1 ("Outstanding") status. It was familiar to the researcher as it had been her children's school some years previously, but it was not a school at which the researcher had ever taught, and the pupils were not known to her.

The class had 28 pupils (14 boys and 14 girls). Two boys had statements of special educational need; one pupil with Down syndrome and one with ADHD. Both these pupils had learning support assistants with them in the classroom. The pupil with Down syndrome took part in the lessons, but the pupil with ADHD did not. Neither of these two pupils took part in the assessments at the request of the Head-teacher. In the later stages of the intervention three children with decoding difficulties were withdrawn from some sessions to receive extra phonics teaching; as discussed below their data was not included in the analyses for this study. This left data from 23 children (10 boys and 13 girls).

5.2.2. Ethical issues

The study involved participants under the age of 18, so a CRB check and ethical approval for the study were required and obtained. Guidelines from the British Psychological Society were followed and ethical approval was obtained from the Department of Psychology and Human Development, Institute of Education, London. Letters describing the study were sent out to the parents/carers of all the children in the class and the opportunity was given for any parent/carer to opt out of the study (Appendix V). Data was anonymised, and all computer records referred to numbers. The key to names was kept in a locked drawer at the researcher's office.

5.2.3. How the intervention was introduced

As in Study 2, the researcher was introduced by the teacher a few weeks after the start of the autumn term, so that the class had had a chance to settle into their new routines. The children were not familiar with the researcher as she had not taught in that school before. The researcher explained that she was a scientist from the University of London and that

192

scientists were interested in learning about things. In this instance, she was interested in what was happening inside children's heads when they were reading. As it is not possible to lift off the top off children's heads and look inside, she was going to need their help. She explained that she was going to be taking children out of their lessons one at a time to do some reading and to give them a chance to tell her what was happening when they were reading. When she had seen everyone two or three times they would be working on a project about reading with their class teacher for the rest of the school year and she would be there to watch what was going on and to ask questions about what the children had learned at the end. Before the pre-testing took place, the researcher spent three days in the classroom, listening to children read in groups as part of the normal school routine, and assisting in lessons, so that the children became used to her presence.

5.2.4. Pre-instruction measures

Table 26 presents a summary of the standardised scores for word reading and reading comprehension obtained from the TOWRE and YARC tests before instruction commenced. For the TOWRE scores there was one outlier for each of the sub-tests and the combined score, but as these cases had little effect on the 5% trimmed mean they were retained.

The QRI and MPIR were also administered pre-instruction, as in Study 2, although a revised edition of the QRI was used, which had become available since the previous study (QRI-5Leslie & Caldwell, 2011).

Table 26. Mean standardised scores from the YARC and TOWRE at Time 1 for the children in Study 3 (standard deviations are in parentheses)

	Mean
YARC reading comprehension	106.91
	(7.62)
YARC accuracy	103.35
	(7.63)
YARC reading rate	107.87
	(8.33)
Sight word reading	110.39
	(11.70)
Phonemic decoding	107.78
	(13.69)
Total word reading efficiency	111.57
	(14.75)

5.2.5. Materials

The materials and procedure were the same as in Study 2. Each of these instruments was described, and an outline of their administration was provided, in the methodology section of chapter 4, and in Appendices I and J. The assessments were carried out in an alcove between the Year 3 and Year 2 classes, which was shared space between the semi open plan classrooms used by either class for small group activities such as painting. At times this was a noisy place with distractions from both classrooms, but it was all that was available.

5.2.5.1. The Qualitative Reading Inventory

The recording, transcribing and coding was carried out in the same way as in Study 2. The number of scripts involved was greater here, therefore the think-aloud transcripts were entered into NVivo (QSR International Pty Ltd, 2010) a qualitative analysis programme, with each file being identified by its first line. As NVivo collates sources alphabetically this had the effect of randomising the entries and removing any information which could identify the child or the time of the recording. The coding was done by the researcher, and then 25%

of the transcripts (chosen by a random number indicator) were marked by a postgraduate student trained in the use of the coding schedule. Inter-rater agreement was 92%. The remaining eight per cent was resolved by discussion, and the rest of the transcripts were re-examined in the light of the discussion. Coding reliability was further enhanced by the use of NVivo, since it is possible to collate all the clauses coded under a single node, making it easier to read through them all and ensure that they were all examples of that particular strategy.

5.2.5.2. The strategy interview

The strategy interviews were taped, transcribed and coded in the same way as in Study 2. Again, the scripts were marked some time after they were taped (with over a year elapsing between the assessments and coding) with the exception of the Time 4 interviews which were transcribed and coded shortly after their administration. The transcripts were scored by the researcher, who was blind to the child at all time points and blind to the time point (with the exception of recordings at Time 4).

5.2.6. Post-instruction measures

The YARC and TOWRE standardised tests were again used post-instruction and for a one year follow-up, as in Study 2. Similarly, the QRI and MPIR were repeated. As in the previous studies a questionnaire was used, this time at Time 2 (after 10 hours RT instruction) and at Time 3 (after 10 hours of RT plus visualisation). The questionnaire was very similar to that used in Study 1 and Study 2 (Appendix D), with the exception of the rewording of one item, and the addition of another. The question "Do you think these lessons have improved your reading?" was amended to "How do you think these lessons have improved your understanding of what you have read?" This change was made because the younger children in Study 2 (as opposed to the older children in Study 1) tended to use just the one word answer "yes" to the former question without offering any explanation. An additional question: "What do you do now when you are reading?" was asked at Time 3 to try to get children to expand on their answers, which had been rather brief at Time 2. The questionnaires were completed anonymously and the children were told that in order to be helpful to the research project they should be honest. The evaluation questionnaire was marked in the same way as in Study 2.

An additional measure in this study was the scores from Standardised Assessments Tasks (SATs). In common with other state schools, children in Year 3 and Year 4 sit optional

195

SATs test papers. In this school it is the practice to use these twice in Year 3; in September at the start of the academic year and in March, halfway through the year. These optional tasks are used for assessment purposes, and to make the children familiar with the process. In September 2011, the children in the Year 3 class who were going to be taking part in the Reading Project[15] took the reading test from the 2006 optional task – *The Hunt for the Secret Treasure*. In March 2012, the 2003 optional task, *Gifts from the Sea*, was used. In Year 4 the optional task from 2006, *Antarctic Adventures*, was used in May 2012. These time points coincided with Time 1 (pre-instruction) and Time 4 (the one year follow-up) whilst the second assessment in Year 3 was very close to the Time 2 assessments (testing occurred two weeks into the second period of instruction). These assessments were made by the class teacher, and so were independent of the researcher's assessments and provided another means of assessing progress in reading comprehension over the course of the instruction and at the one year follow-up.

5.3. Procedure

As in Study 2, the assessments were carried out pre-instruction (Time1, September/October 2011) and after ten hours of RT, the assessments were repeated (Time 2, February 2012). Then, ten hours of RTV instruction took place, followed by a third round of assessments (Time 3, May 2012). Finally, after a year of normal instruction, the assessments were repeated again (Time 4, May 2013).

In this study, the RT instruction was delivered by the class teacher. The class teacher was very experienced, with 17 years of experience in primary school teaching across the year groups from Year 1 to Year 4. At the time of the study, the teacher was in the process of completing a master's degree in the teaching of mathematics. She was not familiar with the RT programme. Over the preceding summer she was given the same materials as the teacher in Study 2. These comprised a book, *Reciprocal Teaching at Work* (Oczkus, 2003) and a copy of the unpublished manual compiled by Palincsar and colleagues (1989). She was also given a copy of an article about RT in the primary years (Pilonieta & Medina, 2009), which had not been available at the beginning of Study 2. After reading these materials the teacher met with the researcher before the beginning of the autumn term to discuss the implementation of the programme. At this meeting the teacher said she would be happy to deliver the intervention but that she would need lesson plans drawn up for her. It had been

---

[15] This was the term used by the class teacher when referring to this research.

hoped that the teacher would plan the lessons herself, as had been the case in Study 2. However, given the existing pressures on the teacher's time from her master's degree assignments, it was decided that to proceed with lesson plans/scripts written by the researcher would be the best compromise. Indeed, after Study 2 was completed, a criticism of previous studies was encountered (Duffy, Miller, Howerton, & Williams, 2010) in that not enough detail had been provided about what had actually been taught and how it had been taught (Cervetti, Pearson, Bravo, & Barber, 2006; Guthrie, McGough, Bennett, & Rice, 1996; Van Keer, 2004). As the teacher required more extensive support in this third study, the detailed lesson plans/scripts (see Appendix W for an example) would provide more detail about how the intervention had been taught and what the lessons had contained. Additionally, since a recent report on improving the dependability of research (Funder, et al., 2013) has emphasised the importance of replicability, the use of scripted lesson plans would make this possible.

5.3.1. Details of the Reciprocal teaching intervention

A timetable of the lessons delivered during the instruction period is given in the appendix (Appendix X). In view of the research on distributed practice (see Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006) for a review) it was decided to deliver the instruction over two 30 minute lessons per week rather than one 60 minute lesson. Distributing the teaching in such a way should optimise learning and recall. Although the theory behind this way of learning has been demonstrated since the late 1880s, educational practice has largely ignored the benefits (Dempster, 1988; Dempster, 1989; Pashler, Rohrer, Cepeda, & Carpenter, 2007). The lessons were time-tabled for Tuesday at 9.30 am and Thursday at 11.30am. The instruction was planned to be delivered in two ten week blocks. After Christmas, the class teacher had to make changes to the timetable which meant the Thursday lesson was no longer possible, and the instruction reverted to the one 60 minute lesson a week that had been used in the previous two studies. Thus, ten lessons were of 30 minutes, and the remaining 15 hours instruction was delivered in 60 minute lessons. In Study 2, although twenty 60 minute lessons were planned, four of those hours were taught in eight 30 minute periods owing to last minute changes of timetabling or other activities the teacher had to plan. There was therefore very little difference in terms of distributed practice between the two studies.

As in Study 2 the first few lessons were spent in introducing what it means to be a good reader and what a good reader does to understand what they read. The four strategies

of predicting, clarifying, questioning and summarising were introduced, and four children dressed up as the four characters (see Figure 22).

After the initial sessions, as in Study 2, the teacher drew up rules for working in groups which were initiated by the children and agreed upon by them. They were as follows:

1. Take it in turns.
2. Do give everyone a chance to speak.
3. Use quiet voices in groups.
4. Listen without speaking.
5. Work as a group to solve a puzzle.

These rules were then displayed in the classroom, and referred to when necessary.

After the introduction of the strategies, the children were divided into groups, comprising three groups of five and three groups of four. The teacher assigned children to the groups, with the aim being to make them heterogeneous in terms of reading ability. Each group was named after a Roald Dahl novel. One group had a learning support assistant whose role was to help the child with Down syndrome participate in the lessons. There was an additional teaching assistant who helped with the group work under the teacher's direction; the researcher was present in the classroom as an observer. It was decided to structure the group work more rigorously than in the previous two studies in the light of the teacher's feedback after Study 2, the work done by Pilonieta (2009) and the knowledge that this was a very lively class. The group work often involved working in pairs within the group initially, and the groups were given deadlines, for example, the teacher would say " In two minutes I expect each group to have a word to share that they had to clarify", or "You have five minutes to come up with one 'between the lines' question from the first paragraph." The groups had a 'teacher' as in the previous studies, whose role was to be in charge of the group, to encourage everyone to participate and to provide feedback, both within the group and then to the whole class in the plenary session. The role of teacher was rotated from lesson to lesson. The classroom teacher and the teaching assistant initially circulated from group to group as in Study 2, but as noise and time off task was a problem, a more structured approach to this was also implemented, ensuring each group had support for a set amount of time within a lesson and when the group was unsupported there were written tasks to complete to aid focus. Gant charts were prepared to facilitate this (two examples are given in Appendix Y). At the same time, a talking stick was introduced which the

'teacher' in each group used to ensure that that only one child at a time was permitted to talk.

As in Study 2, the classroom teacher provided modelling and scaffolding. The difference in this study was that written tasks formed a greater part of the instruction as they were necessary to keep the children on task as they seemed much less able to work independently on oral tasks. For this reason, children were not assigned individual character roles, but they worked in pairs within their group or as a whole group, on a particular strategy. Lessons therefore contained a mixture of paired, group and whole class discussion. As in Study 2, lessons contained plenary sessions when the groups would feed back to the class; these plenaries sometimes came during the lesson as well as at the end. Plenary sessions fulfilled an important role in sharing the children's understanding of the text and showing that they had valuable contributions to make.

Figure 22. Children dressed up as the four characters in Study 3.

Predicting Pete

Clarifying Clara
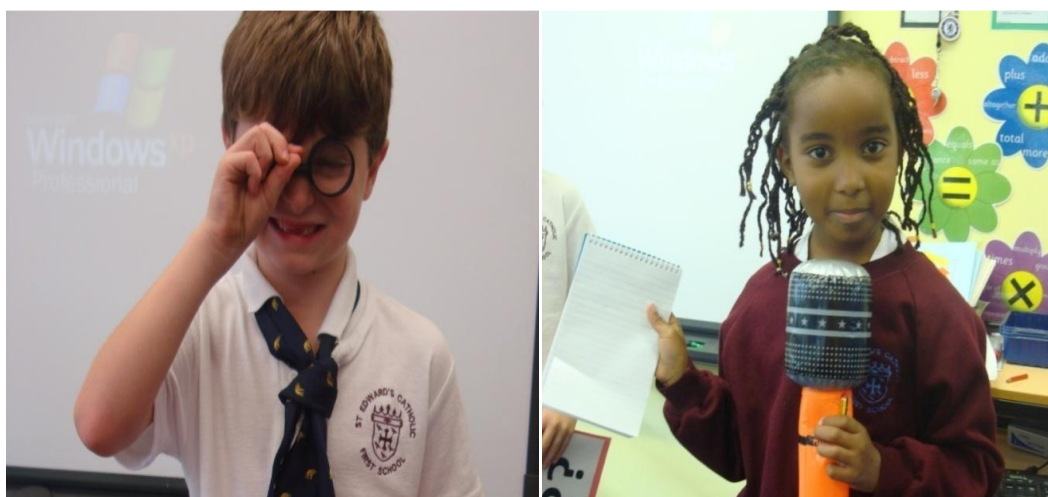
Dennis the Detective

Summarising Susie

5.3.2. Details of the Reciprocal Teaching plus visualisation procedure

Visualising was introduced in the same way as in Studies 1 and 2. The method used was from Bell (1991) with structure words and the movement from familiar single words to phrases to short texts outlined in that book (and covered in more detail in chapter 3 of this thesis). The teacher allowed the children to think of their own name for this character – the boys called him Vincent and the girls called him Vera. The new character was added to the fold-out card along with the other characters, and the structure words were included as a memory aid.

5.4. Results

The aim of Study 3 was to see how RT affected reading processes in those children receiving strategy instruction. The data consisted of scores in the TOWRE and YARC reading tests as well as think-aloud protocols , strategy interviews, and questionnaires, both before and after the intervention with RT and then after RT plus visualisation. There was also a one-year follow-up assessment. During the intervening year the class received normal instruction. The characteristics of the class at Time 1, in terms of word reading and comprehension, are presented first, followed by analyses of the YARC and TOWRE results, the think-aloud protocols and the strategy interviews at each time point, followed by the questionnaire data. Finally, as in Study 2, four case studies are presented to look more closely at the individual responses to instruction by different types of comprehender.

There was no control group in Study 3, but analysis of the YARC reading test results, to be presented after the group characteristics, shows that the children did make more than expected progress in relation to the norms provided in this standardised test. The assessments were marked and coded in the same way as for Study 2. The TOWRE was recorded and double-checked, but not double marked. Of the 92 YARC reading tests, twenty were chosen at random and second marked by an experienced teacher, qualified in Special Needs and practised in administering reading tests. Inter-rater agreement was 95%, and those answers which received different marks were discussed and the remaining tests were re-examined in the light of the discussions. The SATs tests were marked by the class teacher and the raw scores were converted to standardised scores[16].

The TOWRE test of word reading

The scores at all four time points had a normal distribution, as confirmed by histograms and examination of the skewness and kurtosis figures. The assumptions for analysis with parametric statistical tests were therefore met. Although there were differences in the children's word reading efficiency at the four time points (Time 1 M=111.59, SD=3.22, Time 2 M=110.71, SD=3.31, Time 3 M=114.18, SD=2.79, Time 4 M=111.36, SD=2.98) a repeated measures ANOVA revealed that these differences were not significant ($F(3,63)$ =2.29, p=.11).

---

[16] At Time 1, 11 children were too young to score, as the test was designed to be administered in the summer term. Standardised scores for these 11 children were obtained by extrapolating the data in the table of the manual.

The YARC test of reading comprehension

Preliminary inspection of the data revealed that three children did not score on the YARC at Time 1, as the second edition of the test requires the completion of two passages. These three children were only able to complete one passage, and their data were therefore excluded. The comprehension score for one child (child 3.17) at Time 1 was discovered to be incorrect owing to an error in the administration of her test (the score on the second passage meant she should have been administered a further passage) so it was excluded. Of the remaining children, the scores for three children at Time 3 and two children at Time 4 were shown to be outliers, but a comparison of the mean with the 5% trimmed mean showed that these outliers did not distort the average, and could therefore be included. A summary of the data is given in Table 27. The scores at all four time points had a normal distribution, as confirmed by histograms and examination of the skewness and kurtosis figures. The assumptions for analysis with parametric statistical tests were therefore met.

Table 27. Mean standardised scores for reading comprehension from the YARC at four time points in Study 3 (standard deviations are in parentheses)

	N	Mean
Time 1 (pre-instruction)	22	106.91
		(7.62)
Time 2 (after 10 weeks RT)	23	108.83
		(9.43)
Time 3 (after 10 weeks RTV	23	114.39
		(7.72)
Time 4 (1 Year follow-up)	22	112.10
		(8.53)

It was thought to be of interest to see if there was a gender difference in the response to RT instruction, so gender was included as a variable. The data were analysed using a two factor ANOVA. There was no violation of the assumption of sphericity. The effect

of time was significant ($F(3,57)=13.62$, $p<.001$), but there was no significant effect of gender ($F(1,19)=.543$, $p=.47$), and no significant interaction between time and gender ($F(3,57)=.542$, $p=.66$). Paired- sample t-tests showed that the significant differences in reading comprehension scores were from Time 1 to Time 3 ($t(21)=5.83$, $p=<.001$, $r=.79$), Time 1 to Time 4 ($t(20)=4.50$, $p<.001$, $r=.71$), Time 2 to Time 3 ($t(22)=3.65$, $p=.01$, $r=.61$) and Time 2 to Time 4 ($t(21)=2.38$, $p=.03$, $r=.47$). There was a significant difference between Time 3 and Time 4, with the one year follow-up scores being significantly lower ($t(21)=2.08$, $p=.05$, $r=.41$). The difference between Time 1 and Time 2 was not significant ($t(21)=1.68$, $p=.11$, $r=.12$). Plots of the individual children's responses to the intervention, in terms of YARC comprehension standardised scores are given in Figures 23 and 24. The data have been split by gender to avoid cluttering.

Figure 23. Standardised scores for comprehension on the YARC reading test at four time points for the boys in Study 3
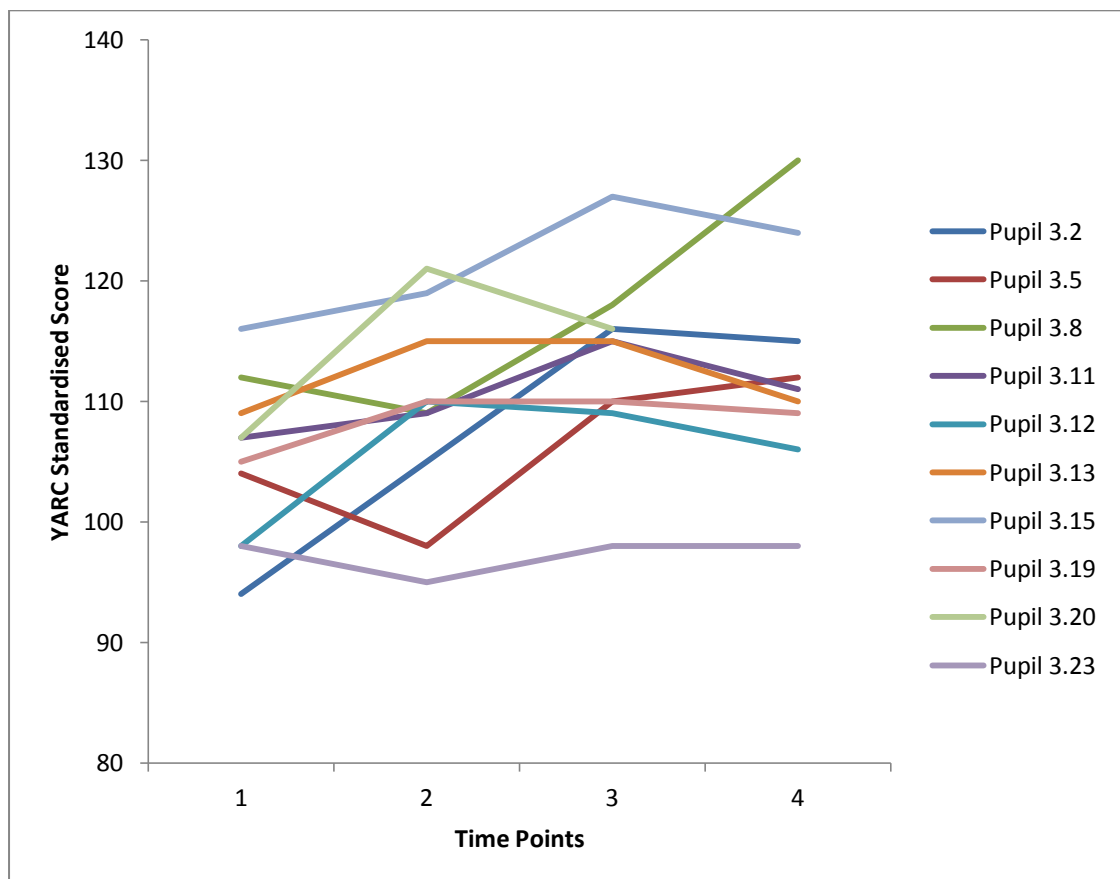
Figure 24. Standardised scores for comprehension on the YARC reading test at four time points for the girls in Study 3.
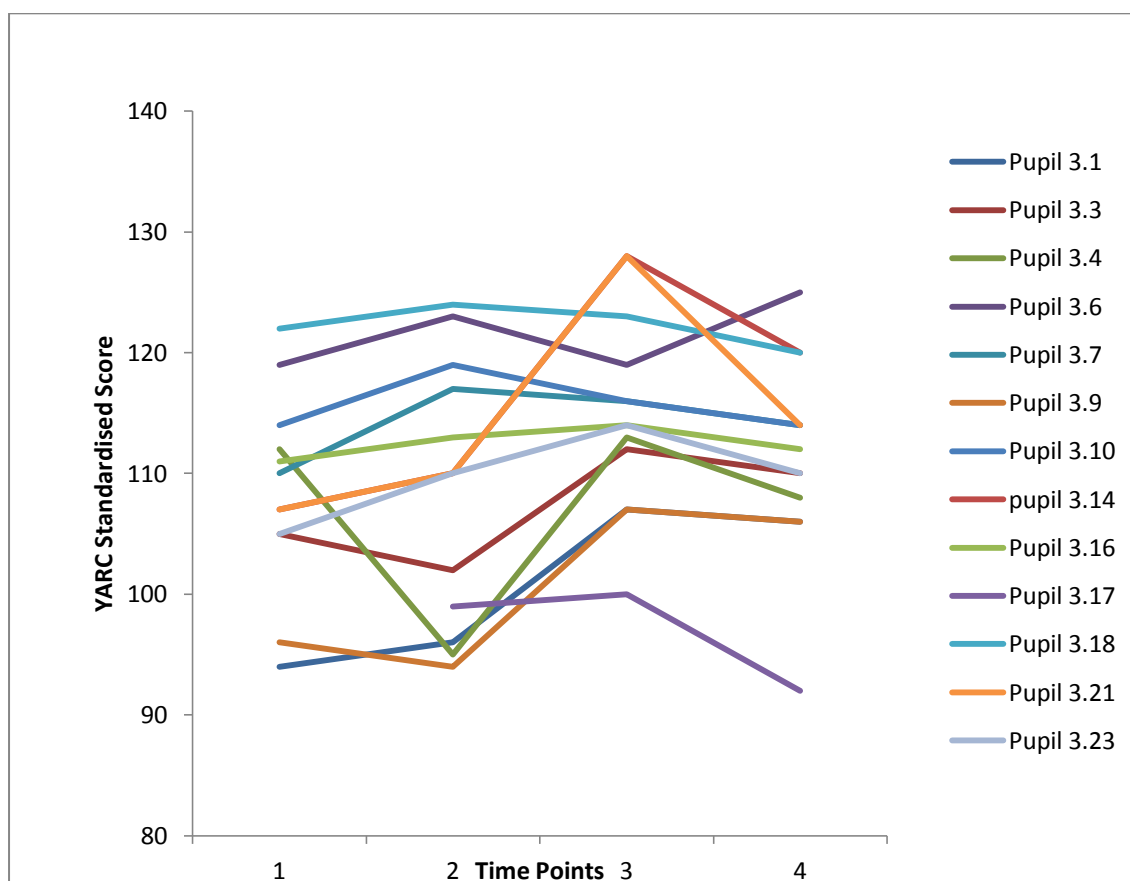
* Pupils 3.14 and 3.21 were born in the same week and have identical standardised scores at three time points; although they answered different questions correctly their total scores were the same.

** Pupil 3.4 shows an apparently large dip in comprehension at Time 2, but her score at this time point might be considered unreliable owing to home circumstances at the time of testing.

5.4.1. Comprehension ability: Ratio gain

Brooks (2002, 2007 and 2013) maintains that a ratio gain of 1.4 or more is educationally significant. Gains of this size, or considerably more, were found from Time 1 to Time 2 (ratio gain 2.55, "modest" according to Brooks) from Time 1 to Time 3 (ratio gain 3.31, "substantial" according to Brooks) and from Time 1 to Time 4 (ratio gain 1.63). The biggest gain was seen from Time 2 to Time 3 (ratio gain 4.07, "remarkable" according to Brooks) which was after the introduction of visualisation as an additional strategy. A summary of

chronological age and reading age equivalent scores from the YARC at the four time points in Study 3 is given in Table 28.

Table 28. Mean chronological age and reading age equivalent scores from the YARC reading test at four time points in Study 3.

	Time 1 (pre-instruction)		Time 2 (10 weeks RT)		Time 3 (10 weeks RTV)		Time 4 (1 year follow-up)	
	M	SD	M	SD	M	SD	M	SD
Chronological age in months	88.61	3.38	92.61	3.88	96.11	3.88	108.61	3.88
Reading age in months	101.83	15.00	112.04	22.15	128.30	15.73	134.36	14.49

Note. Time 1 (pre-instruction), Time 2 (after 10 weeks RT), Time 3 (after 10 weeks RTV), and Time 4 (1 year follow-up).

5.4.2. Accuracy

The YARC results for reading accuracy are summarised in Table 29. Tests of normality showed a significant departure from normality at Time 1 ($D(22)=.21$, $p=.014$) so the data were analysed using a non-parametric test. There was one outlier at Time 4, but as this had little effect on the 5% trimmed mean, and the results were little changed if that data point was excluded, it was retained. The Wilcoxon signed-rank test showed that there was a significant improvement in accuracy from Time 1 to Time 2 ($z=-2.88$, $p=.004$, $r=.43$) and from Time 1 to Time 3 ($z=-3.73$, $p<.001$, $r=.56$). The improvement in accuracy from Time 2 to Time 3 was not significant ($z=-1.31$, $p=.19$, $r=.19$). There was a significant decline in accuracy from Time 3 to Time 4 ($z=-3.81$, $p<.001$, $r=.57$) and from Time 2 to Time 4 ($z=-2.59$, $p=.01$, $r=.39$). There was no significant difference between the scores at Time 1 and Time 4 ($z=.47$, $p=.64$, $r=.07$).

Table 29. Mean standardised scores for accuracy from the YARC at each of the four time points in Study 3 (standard deviations are in parentheses).

	Time 1	Time 2	Time 3	Time 4
	M	M	M	M
N	22	23	23	22
YARC accuracy	103.41	107.00	108.74	102.45
	(7.80)	(8.33)	(8.83)	(8.22)

Note. Time 1 (pre-instruction), Time 2 (after 10 weeks RT), Time 3 (after 10 weeks RTV), and Time 4 (1 year follow-up).

5.4.3. Reading rate

The standardised scores for reading rate on the YARC are summarised in Table 30. The scores at all four time points had a normal distribution, as confirmed by histograms and examination of the skewness and kurtosis figures and there were no outliers. There was a violation of the assumption of sphericity (χ^2 (5) =0.18, p<.001) so the Greenhouse-Geisser correction was applied to the degrees of freedom. A repeated measures ANOVA showed there was no significant effect of time (F(1.49, 28.29)=0.31, p=.67).

Table 30. Mean standardised scores for reading rate from the YARC at each of the four time points in Study 3 (standard deviations are in parentheses)

	Time 1	Time 2	Time 3	Time 4
	M	M	M	M
N	22	23	23	22
YARC reading rate	108.59	108.13	107.91	108.73
	(11.13)	(11.52)	(10.16)	(12.66)

Note. Time 1 (pre-instruction), Time 2 (after 10 weeks RT), Time 3 (after 10 weeks RTV), and Time 4 (1 year follow-up).

Self-corrections

The number of self-corrections that were made during the YARC was recorded as in Study 2. Preliminary data analysis revealed the presence of an extreme outlier at Times 3 and 4, and as this had a large effect on the 5% trimmed mean that data point was excluded from the analysis. Two outliers remained at Time 3, but they had very little effect on the trimmed mean and were retained. The scores at the four time points were: Time 1 *M*=1.39, *SD*=1.34, Time 2 *M*=4.00, *SD*=1.78, Time 3 *M*=3.40, *SD*=3.40, Time 4 *M*=3.32, *SD*=2.87.

It was noted that there was a non-normal distribution of data at Time 1 (caused by the majority of children making none, or only one self-correction). For that reason a non-parametric test was used. The Wilcoxon signed-ranks test showed there was a significant increase in the number of self-corrections at Time 2 compared with Time 1 (*z=-3.97, p<.001, r=.59*) at Time 3 compared to Time 1 (*z*=-3.94, *p*<.001,r=.59), at Time 3 compared to Time 2 (*z*=-2.34, *p*=.02, *r*=.34) and at Time 4 compared to Time 1 (*z*=-3.19, *p*=.001, r=.48). There was no significant difference between the number of self-corrections made at Time 2 compared to Time 4 (z=-1.43, *p*=.15, *r*=.21) but there was a significant decline in the number of self-corrections made from Time 3 to Time 4 (*z*=-2.85, *p*=.01, *r*=.42).

5.4.4. Standardised Assessment Tasks (SATs)

The Standardised Assessment Tasks provide a check on the researcher's assessments as they were administered and marked made by the class teachers. There were significant one-tailed correlations between the scores obtained for reading comprehension using the YARC comprehension ability scores and the SATs standardised scores at all three time points (Time 1(N=23) *r*=.57, *p*=.002, Time2 (N=23) *r*=.75, *p*<.001, Time 4 (N=22) *r*=.57, *p*=.003). Preliminary data analysis showed normal distributions. There were two outliers at Time 1, but they had little effect on the 5% trimmed mean (103.5 and 103.55) so they were retained. A repeated measures ANOVA showed a significant difference between the three time points (Time 1, M=103.5, SD=7.84, Time 2, M=108.45, SD=10.24, (Time 4, M=118.09, SD=11.80, *F* (2, 42) =30.48, *p*<.001). Repeated paired sample t-tests showed these differences were significant at Time 1 to Time 2 (*t* (21) =3.98, *p*=.001, *r*=.66) from Time 1 to Time 4, (*t* (21) =7.20, *p*<.001, *r*=.84, and from Time 2 to Time 4 (*t* (21) =4.24, *p*<.001, *r*=.68).

5.4.5. The Qualitative Reading Inventory

The QRI was used to obtain a measure of comprehension using the questions provided at the end of each passage. It was also used to obtain think-aloud data. The think-aloud data

207

will be discussed first, followed by the answers to the comprehension questions. As in Study 2, two measures were taken for the think-aloud data, the number of comments per think-aloud opportunity and the number of strategies employed in the think-aloud as a whole. Analyses of the number of comments are presented first, followed by the number of strategies employed.

5.4.5.1. Number of comments

A summary of the number of comments per think aloud opportunity over the four time points is given in Table 31. Preliminary data analysis revealed the presence of an outlier at Time 3, which had an effect on the 5% trimmed mean, so that data point was excluded.

Table 31. Mean of the average number of comments per think-aloud opportunity, and the number of strategies reported at the four time points in Study 3 (standard deviations are in parentheses)

	Time 1	Time 2	Time 3	Time 4
	M	M	M	M
Average number of comments	2.73	5.05	5.18	5.47
	(1.25)	(2.80)	(2.19)	(2.23)
Number of strategies reported	2.91	4.65	6.26	6.09
	(1.50)	(2.04)	(1.54)	(1.51)

Note. Time 1 (pre-instruction), Time 2 (after 10 weeks RT), Time 3 (after 10 weeks RTV), and Time 4 (1 year follow-up).

A repeated measures ANOVA was conducted. Mauchly's test indicated that the assumption of sphericity had been violated ($\chi^2(5)=28.78$, $p=.002$) therefore the degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity ($\varepsilon=.60$). The results showed that there was a significant main effect of time ($F(1.81,36.10)=13.15$, $p<.001$). Post-hoc analyses using t-tests revealed that the significant differences were between Time 1 and Time 2 ($t(20)=5.85$, $p<.001$, $r=.79$) between Time 1 and Time 3

(t(20)=8.00, $p<$.001, $r=$.87) and between Time 1 and Time 4 (t(20)=5.74, $p<$.001, $r=$.79). The differences between the other time-points were non-significant.

5.4.5.2. Number of strategies

The number of reading strategies reported by the children in the think-alouds was analysed as the second measure of change. A summary of the results at the four time points is given in Table 31. The data were analysed using a repeated measures ANOVA. Preliminary data analysis excluded the presence of outliers, there was no violation of sphericity, and the distribution of scores was normal. There was a significant main effect of time (F (3,60)=42.76, $p<$.001). Post-hoc analyses using t-tests showed that there were significant differences in the number of strategies for all time points, with the exception of Time 3 to Time 4. There was a significant difference between Time 1 and Time 2 (t(22)=4.66, $p<$.001, $r=$.70) Time 1 and Time 3 (t(22)=11.20, $p<$.001, $r=$.92) Time 1 and Time 4 (t(21)=11.46, $p<$.001, $r=$.93) Time 2 and Time 3 (t(22)=4.22, $p<$.001, $r=$.82) and Time 2 and Time 4 (t(21)=3.82, $p<$ .001, $r=$.64).

5.4.5.3. The comprehension questions

The QRI provides six or eight questions at the end of each passage, according to the level. These questions are equally divided between what the authors call explicit and implicit questions. As the number of questions for each category is limited (three or four for each according to the length of the passage) the authors warn against drawing conclusions about the variation between the scores for each category. However, it does provide an additional measure of comprehension, which adds validity, so the results are considered here. If RT improves children's inferencing skills, then it would be useful to look at the sub-types of questions in the comprehension tests used. It is not possible to make any comparisons of question sub-types on the YARC, since they are not consistent across passages and forms.

Explicit questions
Explicit questions are literal, and focus on information explicitly stated in the text; they usually begin with such words as "who", "what", "where" and "when". Preliminary data analysis of the data showed the presence of outliers at Time 4 and a distribution that was not normal at Time 2 (D(23) =.30, $p<$.001) Time 3 (D(23)=.21, $p=$.013, and Time 4 (D(22)=.24, $p=$.002) so a non-parametric test was considered to be suitable. There were also possible ceiling effects at Times 3 and 4, since several children answered all the questions correctly. A

summary of the percentage of correct answers in both categories (explicit and implicit) is given in Table 32.

Table 32. Mean percentage of correct answers for the QRI at four time points in Study 3 (standard deviations in parentheses)

	Time 1	Time 2	Time 3	Time 4
	M	M	M	M
N	23	23	23	22
Explicit questions	46.68	59.04	71.74	80.68
	(30.26)	(26.36)	(26.42)	(24.31)
Implicit questions	38.75	56.13	54.34	71.59
	(29.28)	(25.01)	(25.73)	(23.52)

Note. Time 1 (pre-instruction), Time 2 (after 10 weeks RT), Time 3 (after 10 weeks RTV), and Time 4 (1 year follow-up).

Analysis with Wilcoxon signed-ranks test for paired samples showed that there was a significant difference between the scores for explicit questions at Time 1 to Time 3 (z=-1.89, p=.06, r=.44) at Time 1 to Time 4 (z=-3.33, p=.01, r=.50) and at Time 2 to Time 4 (z=-2.51, p=.01, r=.37). The differences at the remaining time points were not statistically significant, although there were small to medium effect sizes (Time 1 to Time 2 (z=-1.89, p=.06, r=.28) Time 2 to Time 3 (z=-1.37, p=.17, r=.20) and Time 3 to Time 4 (z=-1.49, p=.14, r=.22)).

Implicit questions
Implicit questions are inferential. The inferences involved can be relatively simple, such as when the answers are stated in the text but in different language, or they may require connecting text segments. Higher-order inferential questions ask children to move beyond the text to predict, hypothesise, reconstruct, form opinions or offer rationale (Leslie and Caldwell, 2011). An analysis of the QRI-4 (Applegate, Quinn, & Applegate, 2002) showed that 23.7% of the questions required low-level inferences, 17.5% asked for high level inferences and 18.4% for response-based inferences. As the QRI authors note, since questions that require higher-level reasoning tend to indicate a deeper level of comprehension than that

shown by literal questions, an improvement in the scores for implicit responses would tend to indicate an improvement in the situation model constructed.

A summary of the percentage of correct responses at each time point is given in Table 32. Preliminary data analysis of the answers to the implicit questions showed the presence of outliers at Times 2 and 3 and a distribution that was not normal at Time 2 (D(23)= .30, p<.001), Time 3 (D(23)=.26, p=<.001) and Time 4 (D(22)=.24, p=.01) so a non-parametric test was considered to be suitable. The Wilcoxon signed-ranks test showed there was a significant difference between the scores at some of the time points. The difference between Time 1 and Time 2 was significant (z=-2.66, p=.008, r=.39) as was the difference between Time 1 and Time 4 (z=3.36, p=.001, r=.50), the difference between Time 2 and Time 4 (z=-2.13, p=.034, r=.32) and between Time 3 and Time 4 (z=-2.63, p=.009, r=.39). The difference between Time 1 and Time 3 was not significant, but there was a small to medium effect size (z=-1.79, p=.07, r=.26). The difference between the scores at Time 2 and Time 3 was not significant (z=-.18, p=.86, r=.03).

5.4.6. The strategy interview

Strategy interviews (based on the MPIR, Keene & Goudvis, 1995) were conducted at the four time points. The data comprised scores for four of the strategies explicitly taught (predicting, questioning, summarising and visualising) plus the use of schema and comprehension monitoring, which were implicit in the programme. The interviews were scored according to the rubric provided (which can be found at Appendix H) with the maximum score possible being 30. A summary of the data is given in Table 33.

A repeated measures ANOVA was used to analyse the data. There was no violation of the assumption of sphericity, and a main effect of time was revealed (F(3,63)=58.02, p<.001). Post-hoc t-tests revealed that there were significant differences at all time points with the exception of Time 3 to Time 4. The difference between Time 1 and Time 2 was significant (t(22)=6.49, p<.001, r=.81) as was that between Time 1 and Time 3 (t(22)=11.71, p<.001, r=.88) and between Time 1 and Time 4 (t(21)=10.16, p<.001, r=.91). Between Time 2 and Time 3 the difference was significant (t(22)=5.30, p<.001, r=.75) as was the difference between Time 2 and Time 4 (t(21)=6,14, p<.001, r=90). The difference between Time 3 and Time 4 was not significant (t(21)=.17, p=.87, r=.03).

Table 33. Mean strategy interview scores at the four time points in Study 3 (standard deviations are in parentheses)

	Time 1	Time 2	Time 3	Time 4
	M	M	M	M
Strategy interview score (max. = 30)	11.61 (0.68)	15.59 (0.76)	18.96 (0.65)	15.93 (0.72)

Note. Time 1 (pre-instruction), Time 2 (after 10 weeks RT), Time 3 (after 10 weeks RTV), and Time 4 (1 year follow-up).

5.4.6.1. Evidence of reading more actively

When self-corrections on the YARC reading passages were examined it became clear that these were few and far between at Time 1. Standards for coherence appeared to be low, and children routinely read phrases which made no sense whatsoever. For example, one child (when reading about an aeroplane journey, where the plane has bumped in the sky and the child on the plane asks his mother what it was) said "it was just a *cold*" (instead of a *cloud*). Another child described robins' eggs as being "pale with *radish* spots" (instead of *reddish*) and another was happy to let a dog flee with "a *signal* sausage" instead of a *single* one. Reading about a burglar, one child said "he gazed at the *continents* of the room", rather than the contents; whilst another child said, "he felt the man's *steadily creep* on his shoulder", rather than his *steely grip*.

The mean number of self-corrections made on the YARC at the four time points was: Time 1 *M*=1.39, *SD*=1.34, Time 2 *M*=4.00, *SD*=1.78, Time 3 *M*=5.43, *SD*=3.39, Time 4 *M*=3.31, *SD*=2.87. Preliminary data analysis showed the presence of an extreme outlier (the same child) at Time 3 and Time 4. As these outliers had an effect on the 5% trimmed mean these data points were excluded. The remaining outliers were not extreme and had minimal effects on the trimmed mean, so they were retained. Analysis also showed that there were significant departures from normality (Time 1 *D*(23)=.19*, p*=.02, Time 2 *D*(23)=.20, *p*=.02, Time 3 *D*(23)=.22,*p*=.007, Time 4 *D*(22)=.22, *p*=.01) so a non-parametric test, Wilcoxon signed-rank, was used. This showed self-corrections were significantly higher at Time 2 than

at Time 1 (z=-3.86,p=<.001, r=.60) and significantly higher at Time 3 than at Time 1 (z=-3.88,p=<.001, r=.57). At Time 4 they were significantly higher than at Time 1 (z=-3.05, p=.002 r=.53) and at Time 3 they were significantly higher than at Time 2 (z=-2.16, p=.031, r=.32). However, at Time 4 there were significantly fewer self-corrections than at either Time 2 (z=-2.00, p=.045, r=.30) or at Time 3 (z=-2.71, p=.007, r=.40).

5.4.7. Spontaneous comments made during the YARC passage reading

As in Study 2, children's comments made during their reading of the YARC passages were recorded. Only 3 children out of the 23 made any comment when reading the two passages at Time 1, and they only made one short comment each. One child made a comment at the end of the fiction passage ("Ah, that was a bad house to break into") and two children made a brief comment about a word they could not decode ("what does that say?" and "I still can't do that word"). At Time 2 however, 10 children (the 3 from Time 1 plus an additional 7 children) made a total of 20 comments. Eleven comments were related to single words. Three of these comments were difficulties with decoding (e.g. "I don't know that word" when a word was refused) but the other eight related to words which were read aloud but not understood (e.g. "I'm puzzled" and "I don't know what threatened means"). Other comments were made about features of the text – the phrase *when Mum had had enough*, elicited two comments – "that's weird" and "there are two hads!", and the sentence *it was a glorious sunny day* drew the response that "they could of (*sic*) put gorgeous there". Two children laughed at the ending of the story about the burglar who broke into a policeman's house and one child commented "that's nice" when Mum gave Dad an ice-cream to apologise for blaming him for losing her handbag when she had left it in the toilets. One child reading that *honey bees never form an angry swarm because their nests are small* made the observation "If they're outside of their nests they could make a swarm" whilst the same passage on bees caused one child to admit that "they look scary". These comments all indicated children who were more actively engaged in the text, who were questioning their understanding and even the text itself.

At Time 3, there were 25 comments from 12 children, with 4 children making a comment for the first time. Two comments related to decoding difficulties (e.g. "I need help with this word"); whilst 11 comments related to children not understanding the meaning of a particular word (e.g. "What's a *drone*?"). One child went further and offered a substitution (a strategy learned in clarifying) after reading: *They use their sharp claws to excavate animals and eggs hidden in the ground*, he said "Does that mean kill them?" Six comments

related to the contents of the text and the children's background knowledge (e.g. "I've never heard of one of those lizards" and "Two metres long? That's big!"), whilst the same child asked a question ("can this animal fly?"). One child made a comment about the words in the text ("tracking and attacking, they rhyme") whilst three laughed and said that the burglar breaking into the policeman's house was funny.

At Time 4, 7 children made 9 comments. One child made a comment about decoding ("I don't know that word ") whilst 3 comments were made querying the meaning of a word (e.g. "What does sole mean?"). Two children lost their place and made an exclamation before going on to re-read, and one child commented on the repetition of the word *had*. One child said that she thought she had read the passage before and finally one child wondered about the main idea of the passage - "What's the difference between a honey bee and a bumble bee?" There is a possibility that the increase in comments may be attributable to the children becoming accustomed to making think-aloud comments during the QRI, but during the NARA they are not asked to comment. The standardised tests were always conducted before the qualitative measures at each time point in all three studies. As has been mentioned previously, the children were told they were being timed, which would be presumed to discourage them from making comments.

5.4.8. The children's evaluation questionnaires

As it was suggested in Study 1 and Study 2 that RT might improve children's reading through increasing their motivation, a questionnaire to measure this was used again. This time it was used at Time 2 (after 10 hours of RT) and again at Time 3 (after a further 10 hours of RT plus visualisation) to gather information about what the children thought they had learned and whether they had enjoyed the instruction. The group statistics for the responses to the first two Likert scale questions are given in Figures 25 and 26.

Figure 25. Questionnaire responses about whether anything was learned for Time 2 and Time 3 for Study 3.
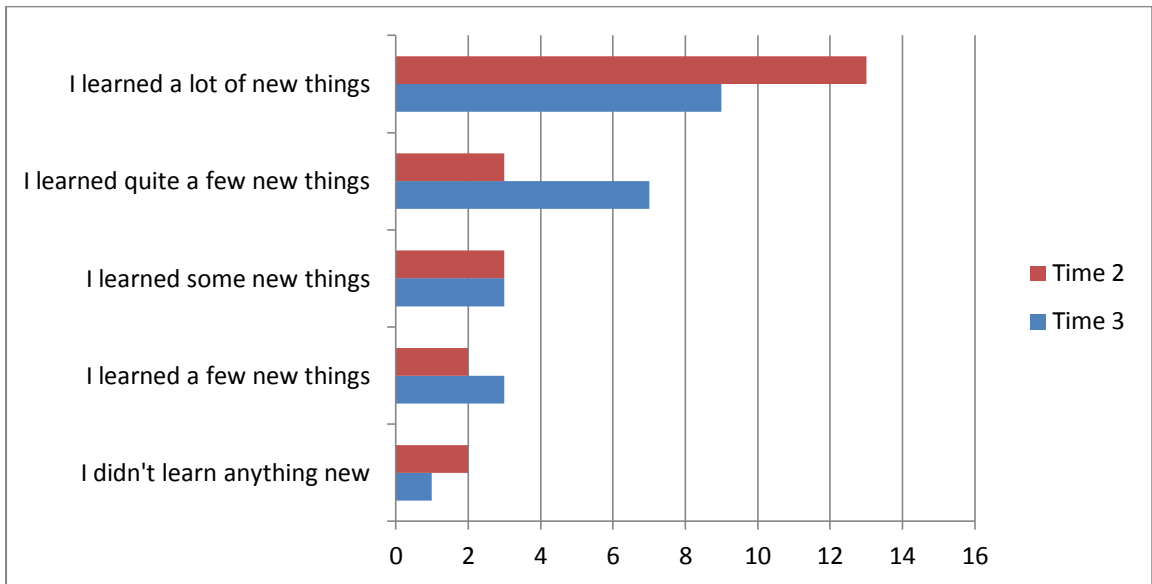
Figure 26. Questionnaire responses about whether children enjoyed the lessons at Time 2 and Time 3 for Study 3.
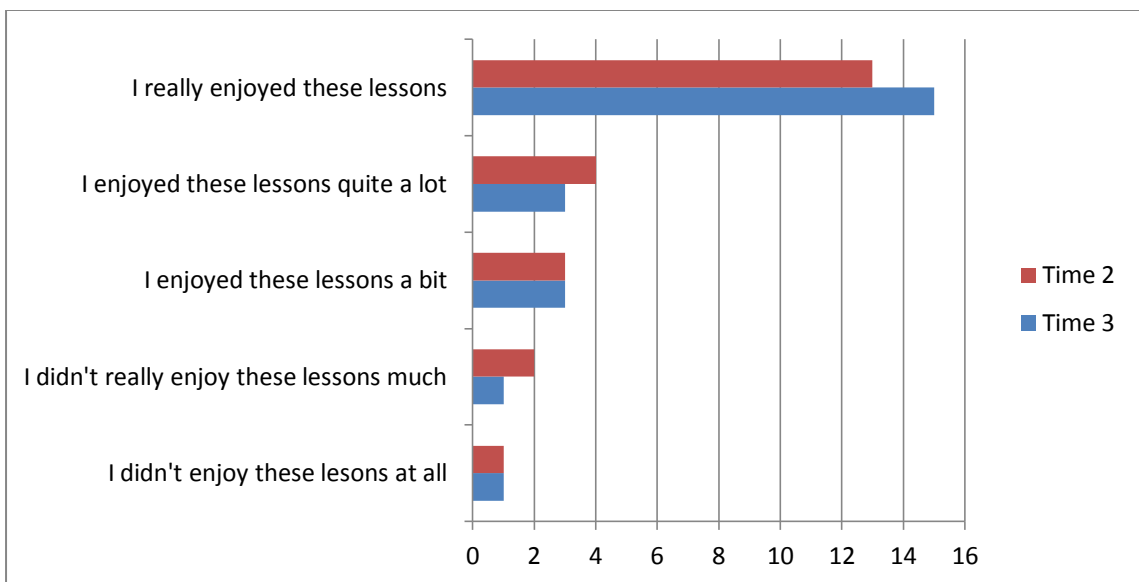
5.4.8.1. Time 2 responses for learning about reading

As Figure 25 shows, at Time 2, two (9%) children felt they had not learned anything new, but 91% felt they had learned at least a few new things, with 56% saying they had learned a lot of new things. The question about what the children had learned had answers which included: "I learned about using the four characters" (four responses), "You had to read between the lines" (three responses), "I have learned about Predicting Pete" (two responses), "I have learned new words, new stuff and how to make a summary", "I have learned how about Clarifying Clara", "I learned new things about thinking" and "I have learned what is the real thing about reading". Several children (11) responded with general comments about having learned "a lot" or "everything".

5.4.8.2. Time 2 responses for whether children enjoyed the lessons

When their enjoyment was questioned, as shown in Figure 26, 65% of the 23 children reported that they "really enjoyed" the lessons and only one child did not enjoy the lessons at all. The enjoyment came from "telling people what is in my brain" or using my brain (five responses), working as a group (five responses), learning new things or learning about the characters (six responses), summarising (one response), reading (two responses), high-lighting words (two responses), and drawing the characters (one response). As to what the children did not enjoy, 14 children replied "nothing", one said "I didn't enjoy reading when I got stuck", two said they did not like writing, one did not like the stories, and one felt that that Predicting Pete was too hard. Two commented on the more administrative aspect of the lessons, with one saying that they did not like the noise and one did not like using the talking stick.

5.4.8.3. Responses to 'Any further comments?' at Time 2

Space was provided at the end of the questionnaire for the children to make any further comments, but very few children wrote anything in this section. Of those that did, five gave positive affirmations of the instruction, saying "It was the best", "It was the best ever", "I really enjoyed learning new things" and "Good work". One child said "It was boring because I know how to read". Two referred to administrative issues: "I liked the folder to put our work in", and "I think you should give them stickers". Finally, there was one comment about the group work - "I think there should be one talking stick so there is not so much noise".

216

5.4.8.4. Time 3 responses for learning about reading

As shown in Figure 25, only one child felt they had not learned anything new, with the remaining 96% feeling that they had learned at least a few new things. There were 12 responses naming particular strategies or combinations of strategies, (four said predicting, four said clarifying, five said summarising, four said questioning, and eight said visualising or making pictures in their head), two mentioned the characters or strategies in general, two learned to read between the lines, one child had learned to think about books, and one learned that reading is enjoyable. Ways of monitoring comprehension were mentioned by five children (e.g., re-read, go back if you make a mistake or you do not understand) and nine responses concerned what had been learned from the texts (seven had learned about The Iron Man and two mentioned facts they had learned from the visualising texts).

5.4.8.5. Time 3 responses for enjoyment of the lessons

As shown in Figure 26, only one child said they did not enjoy the lessons at all (the same child who said they did not learn anything). Eighty-seven per cent of the children enjoyed them at least a bit, and 57% said they really enjoyed the lessons. When asked about what they had enjoyed about the lessons, 12 children said they enjoyed working in groups, four liked learning about the characters (in general), one enjoyed predicting, two enjoyed being Dennis the Detective or making up questions, nine enjoyed the stories they read, three enjoyed the activities, two enjoyed drawing picture, two enjoyed being the teacher, and one enjoyed everything.

5.4.8.6. Time 3 responses for how the lessons had improved their understanding of what they read

Three children said the lessons had improved their understanding of what they read by learning about the strategies in general, two children said they had learned how to figure out the meaning of new words or hard words, three said that making pictures helped them to understand, two were helped by asking questions, one by predicting and three by reading between the lines. Two said it had helped them by making them read more, three said it had made them read better, two said it made them look back and reread and one said it helped them with leadership.

217

5.4.8.7. Time 3 responses to the additional question

At Time 3 an additional question asked: "What do you do now when you're reading?" In response, two children mentioned using strategies in general, four children mentioned two or more strategies and twelve children mentioned one strategy (which was always visualising). Of the other strategies mentioned by name, there were two mentions for predicting, two for clarifying, two for questioning and two for summarising. There were 14 mentions in all for visualising or making pictures. Two children said they stop if they do not know what a sentence means and re-read. Two children said they think when they are reading.

5.4.8.8. Responses to 'Any further comments?' at Time 3

Space was provided at the end of the questionnaire for the children to make any further comments, but as at Time 2, very few children wrote anything in this section. Of those that did, one wrote a comment about how his group was constituted ("Next time I don't want to be the only boy in my group") whilst the rest wrote positive affirmations of the intervention ("Awesome", "Never give up", "Keep doing what you're doing" and "I don't want to stop doing these lessons"). One wrote "To read more instead of watching TV", although it was unclear whether they felt the instruction had encouraged them to do this or that they thought it was something they should be doing. Finally, one child suggested you could design your own character.

5.4.8.9. Time 4 – a one year follow-up

At Time 4 (a year after the second intervention phase had been completed) children were interviewed in preference to using a questionnaire. This was to try and elicit longer responses, as some children do not like writing and may contribute more when asked to respond verbally. This approach could not be used before owing to pressure on time, but on this occasion there was still some time left in the schedule. Twenty-two children were interviewed out of the original 23. One child had left the school. The children were asked: "Thinking back to last year, when Mrs N. and I did the Reading Project with you, what can you remember about what we did?" Then they were asked what they had learned from the Reading Project, and if there was anything they do now when they are reading because of what they had learned. And finally, they were asked if they thought they should still be doing the Reading Project in Year 4.

In response to the question about what they remembered, 12 children (55%) immediately responded that they had read The Iron Man, showing that the book used for the second half of the intervention year had been very memorable, leaving a lasting impression. Thirteen children mentioned at least one character by name, whilst 21 of the 22 children (95%) talked about at least one strategy. The strategies themselves (predicting, clarifying, questioning, summarising and visualising) where mentioned nearly twice as often as the characters (Predicting Pete, Clarifying Clara, Dennis the Detective, Summarising Susie and Visualising Vera/Vincent). Only one child who named characters did not also talk about strategies, showing that what the characters helped the children to do when they were reading had become more important than remembering their names. Of the characters that were mentioned by name, nine children named Predicting Pete, eight named Summarising Susie, seven named Clarifying Clara, six named Dennis the Detective and three named Visualising Vera/Vincent. However, in terms of strategies, the most frequently mentioned was clarifying (15 children) followed by predicting and visualising (8 children each) then summarising (5 children) and finally questioning (4 children). It was interesting that so few children mentioned visualising as a character, but far more mentioned the strategy. The character of Visualising Vera/Vincent was not introduced until the second half of the instruction period and it may have been that the children had assimilated the idea of using a strategy from their earlier instruction, and so this strategy was adopted more directly.

When the children were asked if there was anything that they do now when they are reading because of what they had learned in the Reading Project, thirteen children (59%) referred to understanding more ("It makes me understand", "How to understand now, which I didn't do before, but I do now") or using a comprehension monitoring strategy ("If we don't understand something we should read it again", "I would read it again if didn't understand it, and before I just left it"). Twelve children talked about clarifying words by highlighting words they did not know or using sticky notes. Eight children said they made pictures in their heads ("I make pictures so it's easier to understand stuff", "I learned you can make pictures in your head and make it better so you understand it"). Six children said they predicted what was going to happen ("I read the blurb and think about what might happen"). Four children said they summarised what had happened ("I learnt how to make stories shorter") and one child said they "asked questions". More general comments included two children who said it made them concentrate more, five that it improved their understanding, four that it had made them read more, and two that it helped them to learn more. One child said they had learned to read more slowly.

Finally, when the children were asked if they should still be doing the Reading Project in Year 4, one did not know, one said maybe, two said no (because they had already learned how to use the strategies) and 19 children (86%) said yes . Several of those who said yes referred to the fact that some children in the class had not had the opportunity as they were not at the school in Year 3, or that the current Year 3s were not doing the Reading Project. Reasons for wishing to continue included "because it was fun", "because it was helpful, before I didn't know how to understand properly", "it helps people with problems" and "it helps you understand more". Two children thought it would be useful to remind people who might have forgotten (although they did not put themselves in that category) and one child said they wished they were still doing it as they do not get much time for reading in Year 4. Two children also thought that what they had learned would help them in the Middle School when they had to do reading or comprehension exercises.

5.4.9. The teacher's comments about the intervention

The class teacher was given the same written questions about the intervention as in Study 2, but her responses were much less detailed, making it harder to evaluate. To compare with Study 2 her comments are given in the same order and in full. Her expectations at the beginning were far more target driven: "I wanted to ensure that all children make at least two sub levels progress in reading for the year." She did share the concerns about inferencing skills that were expressed by the teacher in Study 2 however: "I also wanted to ensure that children could understand how to get answers for inference questions."

When asked about what she thought went well she replied: "Once the children were used to the methods they responded well." As for what was more challenging, she said: "The most challenging aspect for the project for me was to ensure that all the tasks were completed as required and to keep all the children focused on the work required of them." This suggests the teacher was focused on achievement, measured as completion of tasks rather than improving understanding, at times. She did think she would use RT again however, as "it is a great way to share ideas", although it is unclear what she meant about this; was the sharing of ideas between the children or between the teacher and the researcher? When asked if she would do anything differently she said "I would reconsider if some SEN children should be added to the project". Again, her meaning is unclear. The child with ADHD was not included in the project, although the child with Down syndrome did participate in some lessons with his usual support systems. There was anecdotal evidence that this child did try and improve his strategy use, as his father reported that when he was

reading he would stop, and when questioned about why he had stopped he said, "I'm thinking about what that means".

When asked if she had learned anything about teaching reading comprehension the teacher replied: "No, but the use of new terms was a great boost." She also said: "The researcher gave of her time willingly and it was lovely to see the children respond so willingly as well."

The next section will look at the four case studies, selected as in Study 2, to look in more detail at any changes in strategy use during and after the intervention, for different types of comprehenders.

5.5. Case studies- the response of different types of comprehender

In Study 2, the responses of good and poor comprehenders were considered. In that study, using the criteria from Cain (2010) of children with age appropriate word reading skills but with comprehension skills of at least 12 months below their chronological age, two children were identified as poor comprehenders. However in this study, no child had a comprehension age on the YARC that was a year or more below their chronological age. In a recent book (Clarke, Truelove, Hulme, & Snowling, 2014), it was acknowledged that in practice the criteria used to define the poor comprender profile are not fully agreed, and the cut-off used to define an impairment are "to some degree arbitrary" (p. 13). Children with a discrepancy between the ability to read words aloud accurately (which is usually average) and the ability to answer questions which tap understanding (which is usually below average) can be considered to fit the profile (p.14).Given the different criteria, such children comprise between 3 to 10% of primary school aged children. In the class in Study 3, there was one child (3.2) with a comprehension age of eight months below that expected, and whose word reading ability was 1 year and 4 months above their chronological age, who although not fitting the criteria of Cain (2010) exactly, nevertheless showed a discrepancy in terms of reading comprehension when word reading ability was taken into account. A second case study (3.6) presents a good comprehender, whilst 3.7 is an example of a child who was a reluctant responder, as opposed to child 3.23, who was a poor responder.

5.5.1. A 'poor' comprehender: Pupil 3.2

Examination of the data for this child (Pupil 3.2) revealed an increase in the YARC standardised score for comprehension, from pre-instruction to Time 3 and then a small

decrease between Time 3 and Time 4. These data, accompanied by comments from the think-aloud and strategy interview, which showed an increasingly active response to the text, proceeding from the word level to the text level, are shown in Figure 27. Word reading ability remained constant (standardised scores for Time 1 = 103, Time 2 =104, Time 3 = 104, Time 4 = 102) so any increase in comprehension ability could not be attributed to an increase in word reading.

Figure 27. Standardised scores for YARC comprehension together with think-aloud and strategy interview comments at four time points for Pupil 3.2

5.5.2. A good reader: Pupil 3.6

As in Study 2, a good reader was selected to see if there was any benefit of the intervention for a child who already had good reading comprehension as well as good decoding skills. Pupil 3.6 had excellent word reading (standard score of 140 on the TOWRE word reading efficiency) and very good reading comprehension (standard score of 119 on the YARC) scores, making her the best reader overall in the class as measured by these two assessments. Her response to the intervention and think-aloud and strategy comments can be seen in Figure 28. Although she was already a strategic reader at Time 1, employing seven different strategies in the first think-aloud and scoring 18.5 in the strategy interview, she did change her primary focus from the word level to the text level, and she employed her background knowledge more effectively, using it to question new information and predict what she might learn. Finally, she showed an emotional response to the text and expressed disbelief at a new fact uncovered. Her ability to explain her strategy use, as measured by the MPIR, improved from 18.5 at Time 1 to 20 and Time 2, then remained fairly constant at Time 3 (MPIR score 19) before increasing to 23.5 at Time 4. This pattern was reflected in her reading comprehension standardised score as measured by the YARC of an initial increase from Time 1 to Time 2, a levelling out at Time 3, and then a further increase from Time 3 to Time 4. As with Pupil 2.10 in Study 2, this good comprehender had become an even better reader over the course of the intervention period and the follow-up period, despite no greater than expected improvement in word reading ability.

Figure 28. Standardised scores for YARC comprehension together with think-aloud and strategy interview comments at four time points for Pupil 3.6.

I never knew that whales didn't have gills. (*Using background knowledge to further what she has learnt from text*)

You should try and check the dictionary. *(Word level focus)*

Some snakes have so venom in their teeth, so maybe the octopus has venom in its beak. (*Combining background knowledge with a prediction*)

I had a question when I was reading. Can the octopus breath above water? *(Questioning, going beyond the information in the text)*

I've been to Africa. The Nile River isn't a very good place to swim in because the current is really fast*. (Using personal experience to comment on text)*

Understanding some of the ideas was a problem. I've never been to Africa when it's raining really hard, so how did the Nile start? *(Using personal experience to question the text)*

I felt sad because the farmers had to work for 8 to 10 hours a day, but I felt happy because they get paid with food and have a day of rest to see their families. *(Expressing an emotional connection to the text)*

I can't believe how long they had to spend on preserving the body…70 days*! (Expressing an opinion about new information learned from the text)*

5.5.3. A reluctant responder: Pupil 3.7

In Study 2 we saw that one very good reader was initially resistant to RT and persisted in using the same strategies she had always used. The researcher was reminded of that case study early in this study when she was helping the teacher by taking a reading group. This reading group consisted of four readers all reading at the same level on the reading scheme. As was the usual practice in group reading sessions, the children took it in turns to read a page each. After each child had read a page the researcher asked some questions about what they had been reading. They were fairly unresponsive and after the third child had read she was asked what she had been thinking about, to which she replied, "I wasn't thinking anything; this is a reading lesson, not a thinking lesson. We don't do thinking with Mrs N." At the end of the session the same pupil complained: "But we've only read four pages. We usually do lots more than that." These remarks indicated that this pupil was not reading actively, and that she measured success in terms of the amount read, rather than the amount understood. The YARC comprehension standardised scores for pupil 3:7 however, show that she did improve her reading comprehension over the course of the intervention (see Figure29) and this was accompanied by an increase in strategy use and total strategy interview score (see Figures 30 and 31). Extracts from the think-alouds and the strategy interview (in the same figure) confirm that she had become a more active reader and accepted that reading was more than just getting through the words, as she had indicated in the initial group reading session.

Figure 29. Standardised scores for YARC comprehension together with think-aloud and strategy interview comments at four time points for Pupil 3.7.

Figure 30. The number of strategies reported in the think-aloud at each of the four time points for Pupil 3:7
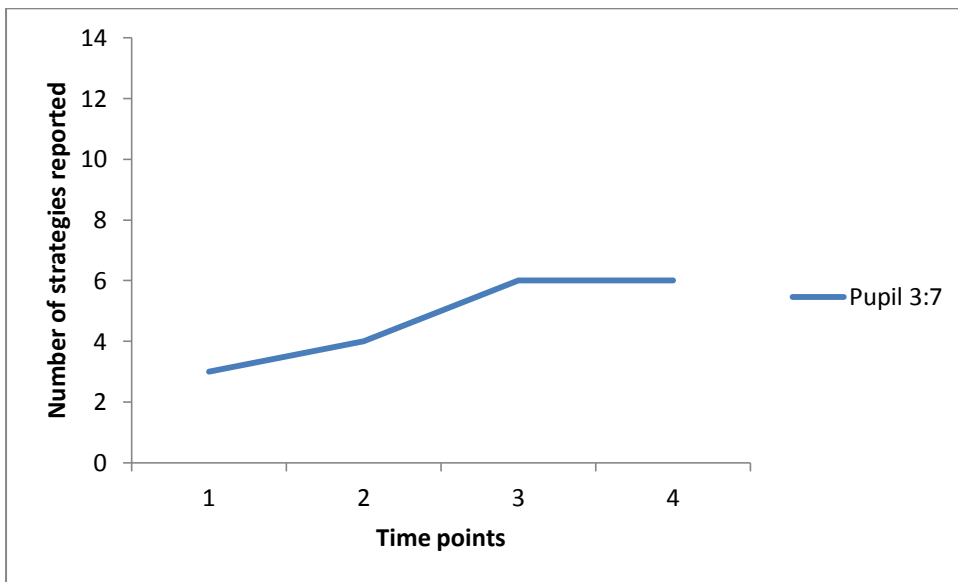
Figure 31. The total score for the strategy interview at each of the four time points for Pupil 3:7
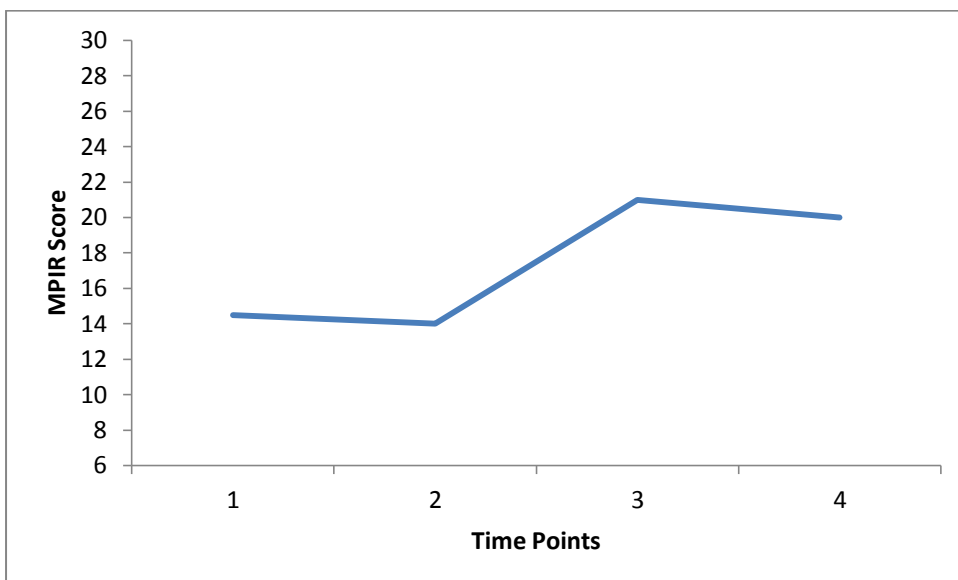
5.5.4. A poor responder: Pupil 3.23

There was one child in this study whose comprehension scores remained fairly constant (see Figure 32). Pupil 3.23, although not meeting the criteria for a poor comprehender as defined by Cain (2010) failed to show improvement in comprehension scores despite increased total

word reading efficiency scores (TOWRE standardised score Time 1=103, Time 2=102, Time 3=111, Time 4=117). In Year 4 (the time of the one year follow-up) his current class teacher expressed her concerns about his lack of comprehension and his reluctance to read outside school. The strategy interview and think-aloud data did show a child who had increased his declarative, procedural and conditional knowledge of reading strategies, albeit from a very low base, whilst remaining among the lowest scores for the class on all the qualitative measures (see Figure 33).

Given the increasing ability to decode single words efficiently, and the corresponding lack of improvement in reading comprehension and the teacher's observations, child 3.23 would seem to require further intervention to remedy the gap between word reading and passage comprehension. He may have appeared to be a poor responder on the basis of his comprehension scores, but as we have seen he did show an increase in strategy use (Figure 33b). RT may have prevented this child from slipping further behind as he appeared to have increased his use of strategies. However, there was little evidence that he was using these strategies to improve his comprehension in a situation where he was not reminded to use them, such as in the YARC reading test. Additionally, as poor comprehenders are known to be a heterogeneous group (Cain, 2010), alternative causes of poor comprehension need to be explored, for example, language or vocabulary difficulties. The YARC does give a breakdown of comprehension questions, and of the six comprehension questions relating to vocabulary asked across the eight passages read during the testing procedure, only two were answered correctly, which might suggest a difficulty which the class teacher was advised to pursue further.

Figure 32. Standardised scores for YARC comprehension together with think-aloud and strategy interview comments at four time points for Pupil 3.23.

Standardised Scores

120
110
100
90
80
70

1   2   3   4

Time Points

— YARC comprehension
— TOWRE word reading
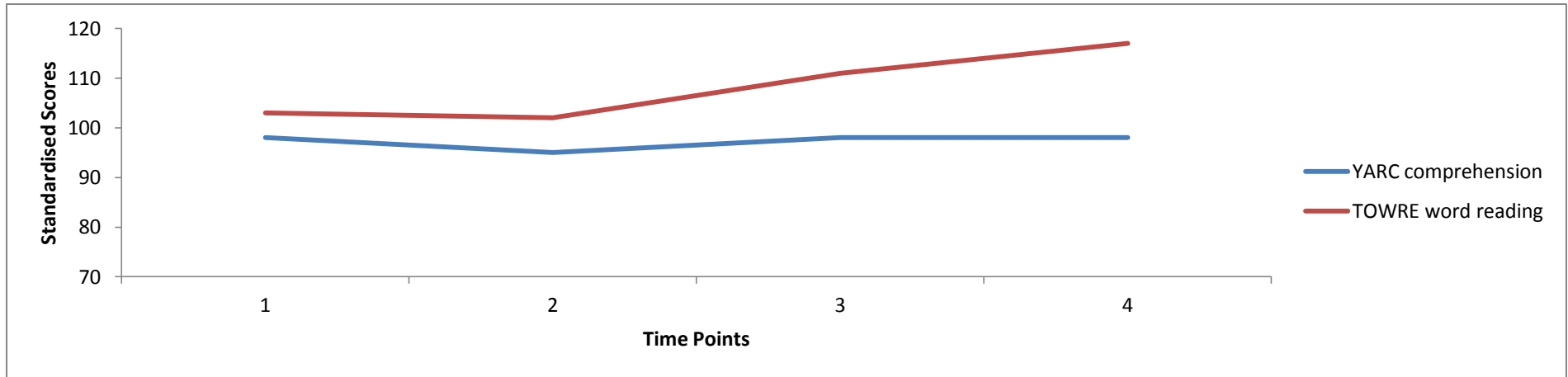
Air is there.

(Repeats words in text)

I had a problem with that word w-e-i-gh.

(Focus on decoding)

They have flippers and fins.

(Summarises in words of the text)

I don't have any problems. It was easy.

(Decoding was easy, therefore the passage is easy)

They crack the animals, the predator's neck. *(Summarises in own words)*

I had a problem understanding the ideas.

(Reading may be more then decoding)

They use the flat tail to pack the mud.

(Summarises in own words)

If it's a big word you can split in half.

(Reverts to decoding emphasis)

Figure 33. Pupil 3.23 compared with the mean score for the class (a) average number of comments per think-aloud opportunity (b) number of strategies reported (c) strategy interview score

(a)

(b)
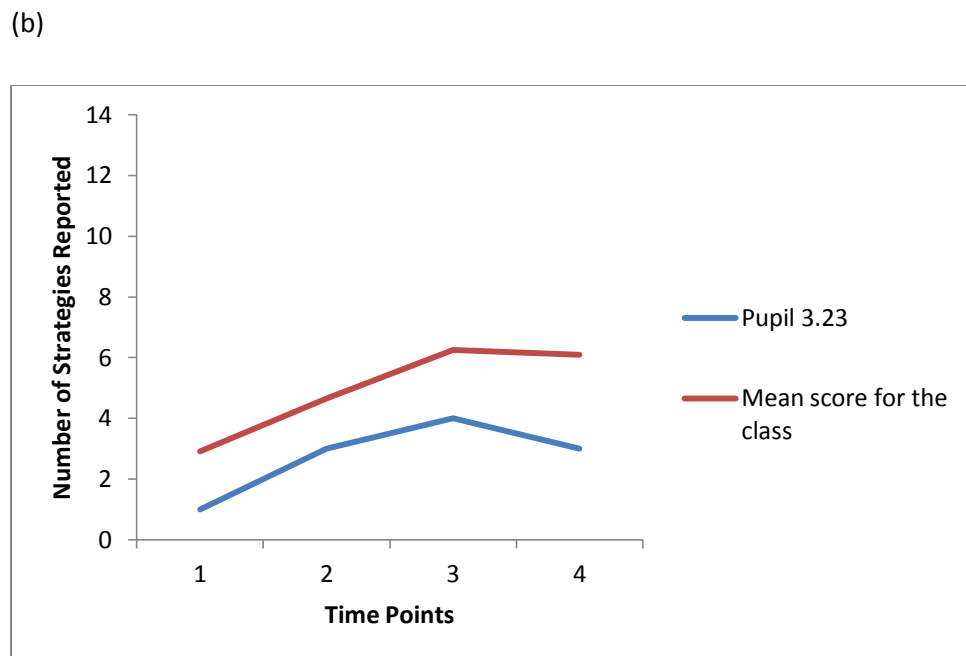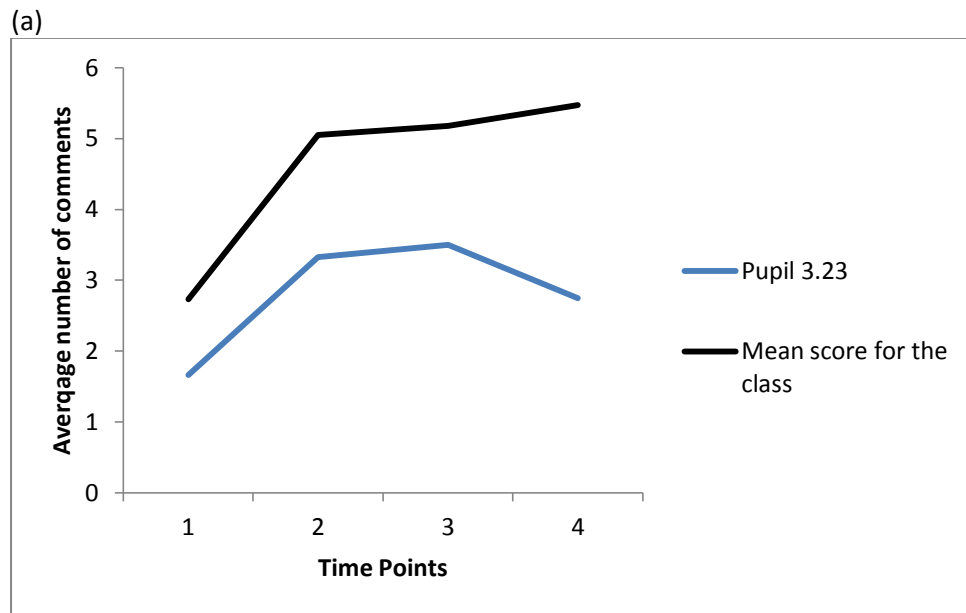
(c)
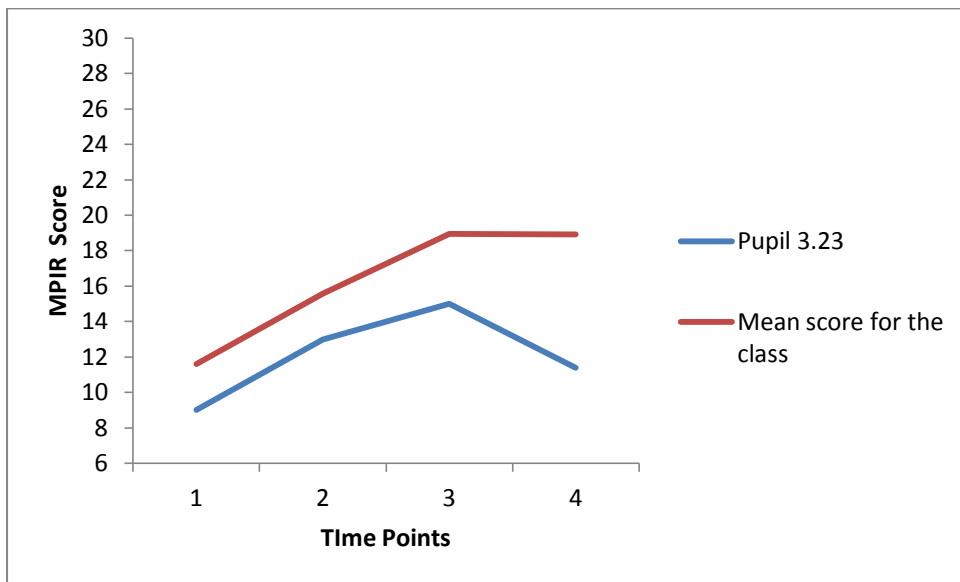
5.6. Discussion

The third study involved an examination of the changes in strategy use, as measured through think-alouds and reading interviews by children aged 7 to 8 who received two periods of RT instruction, from their class teacher. The first period of instruction (10 hours over 10 weeks) involved instruction and practice in the four strategies of predicting, clarifying, questioning and summarising, whilst the second period of the same length added another strategy – visualisation. The aims of Study 3 were to investigate whether RT could be an effective method of improving reading comprehension in a whole class situation, in the UK, for children in Year 3 (ages 7 to 8) and to investigate the changes in strategy use and motivation. A supplementary aim was to see whether the addition of visualisation might change strategy use. The study also investigated the delivery of RT by a different class teacher from Study 2, and investigated the intervention's use in a different school setting. Finally, the study aimed to investigate whether RT and/or visualisation had any lasting effects on reading comprehension, strategy use and motivation. These aims will be considered in turn.

5.6.1. Is Reciprocal Teaching effective in a whole class situation for children aged 7 to 8 in the UK?

The reading comprehension results from the YARC revealed large effect sizes in terms of improvement in scores from pre-test (Time 1) to immediately the instruction ended (Time 3). Although there was a subsequent decline in comprehension scores from Time 3 to Time 4

(the one year follow-up assessment) the difference in Time 1 and Time 4 scores still shows a large effect size (of .71). The effect size of .79, from Time 1 to the immediate post-test is in line with that reported by Galloway (2003) in her meta-analysis, of .74 for the 667 participants in the studies she included. The ratio gain for Time 1 to Time 4 in the present study is 1.63. Using Brooks' standard of 1.4 as educationally significant this indicates that the intervention was effective. In this study the most improvement in reading comprehension scores was observed between Time 2 and Time 3 (with a ratio gain of 4.07, "remarkable" according to Brooks), after the introduction of visualisation. This strategy seems to have been particularly successful with this group of children, as will be discussed in the section below about changes in strategy use.

As was discussed in the literature review, there have not been any previous studies with this age group which have examined RT, as defined by the guidelines of Rosenshine and Meister (1994) within a whole class setting. This is therefore a new finding concerning RT.

5.6.2. What were the changes in strategy use over the instruction period?

Group statistics and some of the case study material from Study 3 have shown that the children generally became more active in their reading after each instruction period. They made more comments about what they were reading in the think-alouds, and they increased the number of strategies they were using. They also scored more highly on the strategy interview. This increase in reading actively and therefore acquiring a higher standard for coherence was confirmed by the increase in the number of self-corrections as children read the YARC reading test passages, and the number of spontaneous comments they made as they were reading. The results of significance testing showed that the number of self-corrections increased as the instruction progressed, and although the numbers fell back between the end of the intervention and the one year follow up, they had still increased significantly since Time 1. Instead of being content to read a phrase which did not make sense children were increasingly inclined to re-read until it made sense. As the passages were graded and chosen to be within the child's word reading ability, the increase in self-corrections cannot be attributed to an increase in word difficulty, but rather an increase in the child's demand for coherence. And as the results of the word reading tests showed, there were no significant changes in the children's ability to read single words.

5.6.3. The impact of introducing visualisation

Although only one child mentioned making a picture in their head in the Time 1 (pre-instruction) think-alouds, from the reading interviews it was apparent that some children were already using visualisation successfully and were aware that they use it as a strategy to improve comprehension. For example, one child at Time 1 (pre-instruction) in response to the question "If someone in your class said 'I've read that but I don't understand it', what would you tell them to do?" replied: "Try and think what could be happening. If it's not a picture you could imagine it." And when asked about the pictures she made when reading the passage she said:

> I sort of made some pictures... like every bit I readed (*sic*) I thought of it………you could close your eyes and think of it…………………if you don't understand you can make the picture of it and you understand it a bit more.

At Time 2 (after 10 weeks of instruction in the original four strategies) the explanation is even clearer:

> I usually make pictures, I like to know the picture and I kind of wait a minute before I read more, so I can kind of make the picture……………if you like imagine the picture it helps you a bit more 'cos you know you can see what's happening………close your eyes and just think really hardly and then a picture just comes into your mind.

So what difference might instruction make to a child who is already visualising as they read? At Time 3 (after 20 weeks total instruction including 10 weeks with the addition of visualisation) the explanation seems very similar:

> I find pictures very easily, like if I look at the colour of the thing I'm trying to think of it like makes a picture in my head……..if you've got a picture in a book to look at it really helps 'cos you can see what's actually happening….pictures in my head really help much…. 'cos if there's no pictures in the book and you're kind of wondering what's the movement, you can kind of make it up. It's not the book's one, but you kind of take a bit from the book's one and then make it up in your mind.

However, there is now an explanation of what the child is trying to visualise which relates to some of the visualisation teaching, namely the use of the words 'colour' and 'movement' which are two of the things children are specifically taught to think about. Movement is particularly important as it is a step up from static images. A similar step up can be seen

from the way that at Time 3 the child is synthesising their images with others that may appear in the book.

In the reading interviews children were asked specifically about visualisation, but these questions came close to the end of the interview, so it was possible to look at the number of children who mentioned visualisation as a strategy previous to the word being used by the interviewer. At Time 1 three children mentioned making pictures in their heads. At Time 2 the number of children mentioning visualisation was again three (although a different three children). However, at Time 3, 13 children mentioned visualisation, with 10 of these not having mentioned that they used making pictures in their mind at either Time 1 or Time 2. At Time 4, 12 children mentioned using visualisation, one of these being for the first time ever. Visualisation therefore appears to have been adopted as a strategy and its use has been maintained over time, despite no further instruction having been given. Its effectiveness as an addition to RT seems to have been shown. However, an ideal design for establishing its effectiveness would involve parallel classes being taught different combinations of the five strategies. This would eliminate the problem in Study 1 of the extra strategy taking up more teaching time.

5.6.4. Improvements in reading accuracy

In addition to an increase in strategy use, there were significant improvements (both statistically and educationally) in accuracy over the instruction period. These improvements in accuracy were apparent from the YARC reading test, but they were not accompanied by similar improvements on the TOWRE word reading test. This suggests that the children became more accurate when reading connected text rather than when they were reading single words, implying that they were using context to read words that were beyond their decoding ability when presented individually. This in turn suggests that they were reading for meaning more effectively as the instruction proceeded. When the instruction ended it does appear that this improvement did not continue, suggesting that children need to be reminded of the importance of maintaining a high standard for coherence.

It was also evident that whilst children referred to the characters by name at Time 2, they referred to the strategies themselves at Time 3, so that "I used Clarifying Clara" became "I figure out what words mean" or "I used Dennis the Detective" became "I asked questions". It would seem that the children had assimilated the strategies into their comprehension monitoring and no longer required the support of a character.

5.6.5. What were the changes in motivation?

There is no doubt that the children enjoyed the RT intervention, as the questionnaire data from what the children thought about the intervention shows. This enjoyment was accompanied by the children's acknowledgement that they had learned new things from it. Reading the comments at the Time 4 interview, there is a definite sense that the children thought the project had been worthwhile and enjoyable. It had helped them improve their understanding of what they read and their acknowledgement of its benefit was also manifested in the concern expressed for other children who had not had the same opportunity.

5.6.6. What were the differences in implementation in Study 3 compared to Study 2?

Both teachers in Study 2 and Study 3 adapted RT to their own needs and the needs of the class, as has been seen in previous research (Marks, et al., 1993). In Study 2 the class teacher prepared her own lesson with minimal input from the researcher, as had been the intention. In Study 3, the researcher wrote the lesson plans from week to week, but they arose from discussion between the researcher and the teacher during the lunch hour on the same day as the previous lesson, and the teacher's ideas were adhered to. Both teachers were familiar with the Palincsar manual (1989) and the book *Reciprocal Teaching at Work* (Oczkus 2003). However these two books present the instruction in very different ways, so it is not surprising that the implementation by the two teachers differed. In the Palincsar manual the strategies are covered by the group as a whole, but Oczkus suggests assigning a character role to individual children within the group. The latter was the method chosen by the teacher in Study 2, whilst the teacher in Study 3 chose to have the whole group working on a particular strategy at any one time. The important thing is that, whichever method is chosen, instruction is provided on why, when and where these strategies might be applied - which is one of the criteria on the checklist for quality of RT instruction provided by Rosenshine and Meister (1994). As was discussed in chapter 2, strategy instruction has been criticised (Hirsch 2003, 2006) for being reductionist, and some strategy instruction has become all about the strategies and not about what we can learn about being an active reader from using them. Palincsar (2006) was scathing about using packs of cards or spinners to allocate the roles (one is provided in Oczkus 2003, p.172) and emphasised instead the need to connect the strategies to understanding the text. However, although the teacher in Study 2 did allocate roles she was careful to ensure that the texts used were suitable for

each lesson and adapted the allocation accordingly; for example, when using the short inferencing texts she did not allocate the role of Predicting Pete. What was common to both teachers was an understanding that it was important to impart the procedural and conditional knowledge as well as the declarative.

5.6.7. Did Reciprocal Teaching and/or visualisation have any lasting effects on reading comprehension, strategy use and motivation?

From the interviews at Time 4, when the children were asked what they had learned from the Reading Project, there were strong indications that these children were more active in their reading. They knew about a variety of strategies and they talked about how they used them in their reading a year after the instruction finished. Over half of them talked about how the project had helped them with understanding what they read. Their responses indicated that comprehension was being monitored and strategies were being used when comprehension broke down. Clarifying (15 children) and visualising (8 children) seemed to have been retained the most. Four children reported that they used sticky notes when reading at home, so that they could ask someone about an unfamiliar word, or look it up in a dictionary later. The sticky notes had been used in the instruction period and the children were given some to take home, but their use had not been reinforced for twelve months, so the children had themselves made a decision to carry on using them. Making pictures in your head was still being carried out spontaneously by 36% of the class, whilst it was only mentioned by one child (4% of the class) during the think-aloud at Time 1 (pre-instruction).

A teacher who was in the same room when the reading interviews were conducted commented that she was amazed that the children remembered so much from something they had not revisited for a year. And despite a decline from Time 3 to Time 4, the Time 1 to Time 4 comparisons for reading comprehension showed overall gains which were statistically and educationally significant. It would seem that whole class RT was again shown to be effective for children aged 7 to 8.

Chapter 6: Conclusions

6.1. An overview

The present research consists of three studies that investigated the effectiveness of RT for two age groups, employing whole class instruction in the UK. The literature review revealed that whole class teaching of this method is under-researched, and that there are no studies with the two target age groups (7-8 and 9-10 years) which provide quantitative data. The data collected were from 50 children aged 9 to 10 in Study 1, from 10 children aged 7 to 8 in Study 2 and from 23 children also aged 7 to 8 in Study 3. Qualitative data concerning changes in concurrent reading processes were also collected in Studies 2 and 3.

Study 1 compared three groups of 16/17 children. One received RT instruction as it was originally conceived; that is, they were taught four strategies predicting, clarifying, questioning and summarising, using explicit instruction by the teacher followed by modelling and a gradual release of responsibility to the children. In heterogeneous groups, the children took it in turns to fulfil the role of the teacher. A second group received the same content and method of instruction, but with the addition of visualisation as a strategy. Both these groups were taught by the researcher. A third group comprised a normal instruction control. Statistical analysis showed that both RT groups made a significant improvement in reading comprehension as measured by a standardised test, while the normal instruction group did not, and that there was no difference in scores between the group that received traditional RT and the group that received RT plus visualisation.

Studies 2 and 3 used larger class sizes (24 and 27 respectively). The children received traditional RT for 10 weeks and then RT plus visualisation for an additional 10 weeks. In both these studies the class teacher delivered the instruction, with the researcher acting as an observer. Reading comprehension was assessed in both studies pre and post instruction and at a one year follow-up assessment. The results in Study 2 showed a significant improvement in reading comprehension scores from pre-instruction to Time 2 (i.e., following ten weeks of RT) and Time 3 (i.e., following ten weeks of RT plus visualisation). Results also revealed a decrease in scores between Time 3 and Time 4 (the one-year follow-up). However, the improvement at Time 4 compared to the pre-instruction result approached significance and showed a ratio gain of 1.77. Results from Study 3 showed no significant improvement in reading comprehension scores at Time 2 compared to pre-instruction, but at Time 3 and Time 4 the improvement was significant, with the ratio gain between Time 1 and Time 4

being very similar to that seen in Study 2 at 1.63. Table 34 shows the ratio gains for comprehension at the different time points in the three studies.

Table 34. Ratio gain at all time points for all three studies.

	N	Time 1 to Time 2	N	Time 1 to Time 3	N	Time 1 to Time 4
Study 1 (RT)	16	7.04*				
Study 1(RTV)	17	5.60*				
Study 2	10	4.78	10	4.01	9	1.77
Study 3	22	2.55	22	3.31	22	1.63

* Children with ceiling scores at Time 1 omitted
Note. Time 1 (pre-instruction), Time 2 (after 10 weeks RT), Time 3 (after 10 weeks RTV), and Time 4 (1 year follow-up).

As can be seen from the differences in ratio gain, there are differences in the ways in which the children in the different studies appear to have responded to the instruction. There were improvements in reading comprehension after ten weeks of instruction, which Brooks (2007, 2013) would consider as showing 'remarkable' progress in Study 1, 'substantial' (and very close to 'remarkable') progress in Study 2, and 'useful' progress in Study 3. After 20 weeks instruction, the children in Studies 2 and 3 had made 'substantial' progress. And at a 1 year follow-up, with no intervening instruction, children in Studies 2 and 3 were still showing a 'useful' impact on their reading comprehension. Whilst it is difficult to say with any certainty why this might be, there are some possibilities which will be discussed next.

6.1.1. Possible reasons for differences in results across the three studies

The standardised test

In Study 1, the reading test used was the NARA (Neale, 1989); whilst in Studies 2 and 3 it was the YARC (Snowling, et al., 2011; Snowling, et al., 2009). The importance of the way

comprehension is tested was discussed in the literature review, and although both tests used the same format, it does not mean they are measuring the same thing. The NARA has been criticised in that Form B has an inappropriate gradation in the difficulty of comprehension questions (Stothard & Hulme, 1991). The NARA has also been shown to be over-reliant on questions which require literal comprehension (Hurry & Doctor, 2007). Differences between Study 1 and the other two studies might therefore be due to the method of measurement.

The difference in age groups

Study 1 involved children in Year 5 (aged 9 to 10) and Studies 2 and 3 involved children in Year 3 (aged 7 to 8). Both these groups are under-researched as the literature review outlined. However, the studies presented here do not enable us to say which age group benefits the most. Certainly the 9 to 10 year olds in Study 1 seemed to have improved more quickly than the 7 to 8 year olds in Study 2 and much more quickly than in Study 3, but without follow-up data from the first study, it is impossible to say whether that improvement would have been maintained. More research is needed for children aged 9 to 10 which includes follow-up data. However, despite misgivings about RT being too cognitively demanding for such a young age group (Cain 2010, p.185) Studies 2 and 3 have shown that these young children are capable of not only participating in such a programme but of benefitting from it in terms of improved comprehension; this has possibly come about through an improvement in strategy use, as shown by the correlation between the increase in strategy use (as measured by the MARSI questionnaire) and the improvement in reading comprehension scores in Study 1, and the correspondence of strategy use and comprehension scores in Study 3.

Differences in implementation

The instruction in the three studies was planned and implemented by different people. In Study 1 the instruction was planned and delivered by the researcher. In Study 2 the instruction was planned and delivered by the class teacher, whilst in Study 3 the instruction was planned by the researcher but delivered by a different class teacher. Reading research has often shown that interventions delivered by researchers are more effective than those delivered by teachers (for example Scammacca et al., 2007). Possible explanations given in that particular article were: (a) researchers implement interventions more consistently, (b) researchers implement interventions with greater fidelity, (c) the novelty of a different

teacher providing interventions positively influences students' response, and (d) studies implemented by researchers also use researcher-developed outcome measures, which are known to be associated with greater effects. In this thesis any differences cannot be attributed to (a) or (b) since the Rosenshine and Meister guidelines (1994, and given in full in chapter 2) were followed closely in all three studies. And whilst there were differences – for example, the children were given character roles in two of the studies, but not in the third study – the principles of both content and method were maintained. As for (c), the possibility of any Hawthorne effects were discussed in chapter 3, as the researcher taught the children in Study 1; the researcher was familiar to the children however, as she taught in the school. And as the studies did not involve researcher–developed outcome measures, then (d) can also be discounted. Additionally, the two class teachers involved in Study 2 and Study 3 were very experienced and so it cannot be said that any of the classes had the benefit of instruction from a more experienced teacher. In the case of these three studies, differences in implementation would therefore be an unlikely cause of variation in effectiveness. However, there were differences in the ways in which the children responded during the three interventions, which may help to explain why the improvements from Time 1 to Time 2 were much greater in Studies 1 and 2 compared to Study 3. The children in Study 3 found working in groups much more challenging and it was far more difficult to keep them on task. As was noted in the implementation section of the final study, from the outset the teacher felt that a more structured approach would be necessary, and more work was carried out in pairs and some written tasks were involved. The class was quite lively and the large desk size and more open-plan nature of the classroom made group discussions much noisier than those in Studies 1 and 2. It may have been that the intervention took longer to have an effect as the amount of time on task and working, and time spent working as group was shorter.

6.1.2. Why is it that strategy instruction does not always work?

As we have seen, strategy instruction can be delivered in a variety of ways and the importance of strategy use has become increasingly acknowledged. However, even when instruction has become more widespread, as in the United States, there is no guarantee that there will be a corresponding improvement in reading comprehension. In the Introduction we saw that some large scale studies have failed to find that strategy teaching does improve reading comprehension, indeed, with one intervention, reading comprehension scores became worse (James-Burdumy, et al., 2010). A doctoral thesis (Scott, 2009) concluded:

The findings in this study have repeatedly demonstrated that reading comprehension instruction in the United States classrooms and the resources available to teachers focus extensively on the use of comprehension strategies with little or no attention to constructing meaning of the content in the text (p. 207).

This conclusion echoes the criticisms of strategy instruction by Hirsch (2006) i.e. that such instruction is narrow and reductive. If the content of the text is ignored then the instruction *is* narrow and reductive. The focus must always be on the purpose of reading, which is also the purpose of the instruction. The purpose of RT is not to teach the readers about the four strategies, but to enable them to become more active readers by telling them about reading strategies, by letting them use them and by getting them to become actively engaged in what they read. As Palincsar and Schultz (2011) contend, strategy instruction in the USA needs to reconnect with its theoretical roots. RT "has been extracted from the theoretical context in which it was originated, rendering it anaemic" (p.88). In discussing the foundations of RT, Tennent (2011) feels this has not always been made clear – "what is absent from this theory base is an explicit focus on the central purpose of using these key strategies: to assist in the making of inferences" (p.127). He goes even further by stating "there is a rationale for re-focusing the theoretical base of RT so that the central importance of inference making is acknowledged and made explicit" (p.128). RT as it was originally conceived did include inference making as one of the six objectives, as was shown in the manual prepared for teachers (Palincsar, et al., 1989), but there may indeed be a case for re-emphasising its importance.

In a recent article about reading theory and interventions for children with reading disability, strategy instruction has again been criticised (Compton, et al., 2014). Reiterating the misgivings of Hirsch (2006) the authors feel that "quick fix" strategy interventions ignore "the fundamental role of reader knowledge in constructing the situation model" (p.57). Compton et al. contend that by failing to promote inference making children are led to a superficial reading of the text and fail to construct a deeper understanding (p.64). This may indeed be the case if strategies are taught in a reductive way, with the focus on strategy use and not on understanding, but the RT approach used in all three studies has involved inference making and constructing a situation model. For example, passages were used which required inferences to be made, and when the strategy of questioning was taught, a distinction was made between surface questions and the kinds of questions a teacher would

ask, which required reading between the lines. Additionally, Studies 2 and 3 showed that children improved their responses to implicit questions on the QRI assessment.

Compton et al. propose that a better approach than strategy instruction would be to teach children to "read between the lines" but this is what RT does, and in post-instruction interviews in the present research this is one of the skills children said they had learned. The authors also put forward the idea that what is important for enabling deeper understanding amongst children with a comprehension difficulty is knowledge building. However, after controlling for background knowledge, poor comprehenders may still have difficulties. Cain Oakhill, Barnes and Bryant (2001) using material originally designed by Barnes and colleagues (Barnes, Dennis, & Haefele-Kalvaitis, 1996) explored differences in inference making between good and poor comprehenders which tested the importance of prior knowledge; they taught the same knowledge base to groups of 7 to 8 year old children. The multi-episode stories were about an imaginary planet, and the knowledge base was taught until perfect recall had been achieved. The poor comprehenders found it more difficult to acquire the knowledge base, but even when they had learned it, differences in the ability to make inferences remained, with poor comprehenders being significantly worse at making inferences than their better comprehending peers. The less skilled comprehenders experienced most failures at the level of retrieving the relevant textual premise. Thus, for poor comprehenders, there may be more to poor inferencing skills than a lack of background knowledge. It is not enough to have the background knowledge available; we need to know how to access that knowledge and how to integrate it into a text. Compton et al. go on to propose that for poor comprehenders it may be necessary to provide an intervention that combines knowledge building with inference making, which indeed may be a way forward for some of that very heterogeneous group, but if we are not to risk throwing out the baby with the bath water, then strategy instruction, properly constituted and conceived still has a place in comprehension instruction.

In the introduction to this thesis it was shown that in the UK strategy instruction is lagging behind, so this may be an ideal time to learn from the way that strategy instruction has been implemented in the USA. Teachers should be told that the goal of strategy instruction is "to equip learners with the means to undertake complex problem solving in more efficient ways" (Palincsar and Schutz, p. 85) and that it is far more than just giving them knowledge about formal skills. As Palincsar contended in her response to criticisms of

RT (Palincsar, 2006) children should be taught that reading is for knowledge building (Bereiter & Scardamalia, 2003), and that the strategies are not an end in themselves.

Having considered that although there may have been differences in how effective RT intervention was in the different studies, there were indications that it may have benefitted both age groups, in bringing about an increase in reading comprehension scores greater than that expected by normal progress. The following sections will look at how the studies might have been improved, and what are the implications for reading comprehension instruction arising from the findings. Finally, there is a section reflecting on what I have learnt as a teacher of reading.

6.2. Improvements in the methodology for these studies

Problems with the methodology in Study 1 were discussed in chapter 3. It was felt that the marking of the reading comprehension tests by the researcher could result in bias in scoring. This was corrected in Studies 2 and 3 where the tests were recorded and double-marked by an experienced teacher who was blind to the identity of children and the timing of testing. Secondly, there were no online measures of reading taken in Study 1. The comprehension test and reading strategy questionnaire used were offline measures. In Studies 2 and 3, a think-aloud was used as an online measure, and the questionnaire was replaced by a strategy interview which was directly related to a text which had just been read. The questionnaire was also replaced because there were some doubts as to the validity of the children's responses, as shown by the answers to some of the foils.

As we have seen, Studies 2 and 3 used online measures of reading to assess the processes taking place whilst reading; this had been missing in Study 1. However, there may be additional online measures to the think-aloud and strategy interview which were used in the current research and which could be informative. One such measure could be eye-tracking. A recent study (Ponce & Mayer, 2014) has shown how college students using different strategies in reading the same passage searched the text differently. Thus students who used a compare and contrast graphic organiser made more eye movements between the top and bottom of the passage than students who used a read only and a note-taking strategy. They also scored more highly on comprehension. Tracking the eye movements of children before and after RT instruction may provide information about whether they are using a more generative approach to reading (for example, by re-reading difficult phrases or

looking back at important sections) rather than reading in a purely linear way, from the first word to the last.

6.3. Improvements in Reciprocal Teaching implementation

6.3.1. The use of gesture

The use of gesture could reinforce the strategy learning (Shanahan & Roof, 2013). The focus of the presented studies has been very much on what the teacher says when the strategies are introduced, and although supports were offered in the form of bookmarks, with young children gesture may have an important part to play. In their research, Shanahan and Roof noticed how a class of elementary school children responded to a question about which strategy they had been using that day with an incorrect strategy name, as the teacher was using a gesture she had previously used for their now incorrect suggestion. They decided to include multiple modes of communication in their next analysis and showed how pairing images with metaphoric gestures helped to reinforce the spoken messages. To reinforce the use of gesture, puppets are available in the USA which depict each of the four characters, in much the same way the children dressed up in Study 3. These puppets could be used with the younger age group in Studies 2 and 3 (age 7 to 8) to provide visible reminders of the strategies being used.

6.3.2. The inclusion of decoding strategies

As has been discussed earlier in the introduction, with the focus on phonics teaching, it is important that we do not forget about the importance of teaching comprehension. This has been highlighted by Rose, in a review of a book about developing reading comprehension (2013):

> In recent years the debate about teaching young children to read has tended to focus upon equipping them with the crucially important knowledge and skills they need to read words accurately in and out of context, that is to say, teaching them how the alphabet works for reading and spelling. While such knowledge and skills are essential, more is required for children to become literate, fluent readers who understand what they read. In short, the goal of reading is comprehension.

Children need to be able to decode in order to read, but they need to be able to comprehend in order to learn from reading and to be able to enjoy it. If RT is used with

young children who have not yet mastered decoding then it may be advantageous to combine phonics teaching and reading comprehension instruction. This has been suggested by Miller (2002) who combines phonics and instruction by teaching strategies that are common to both and emphasising the connections (Figure 34). RT instruction encourages decoding strategies, specifically through clarifying, where children often highlight words they cannot decode. In the studies in the present research, clarifying was largely confined to explaining the meaning of words, but it could be extended to include decoding strategies when words cannot be read at all.

Figure 34. An in-progress classroom chart that shows the side-by-side teaching of decoding and comprehension (Miller 2002, p. 51)
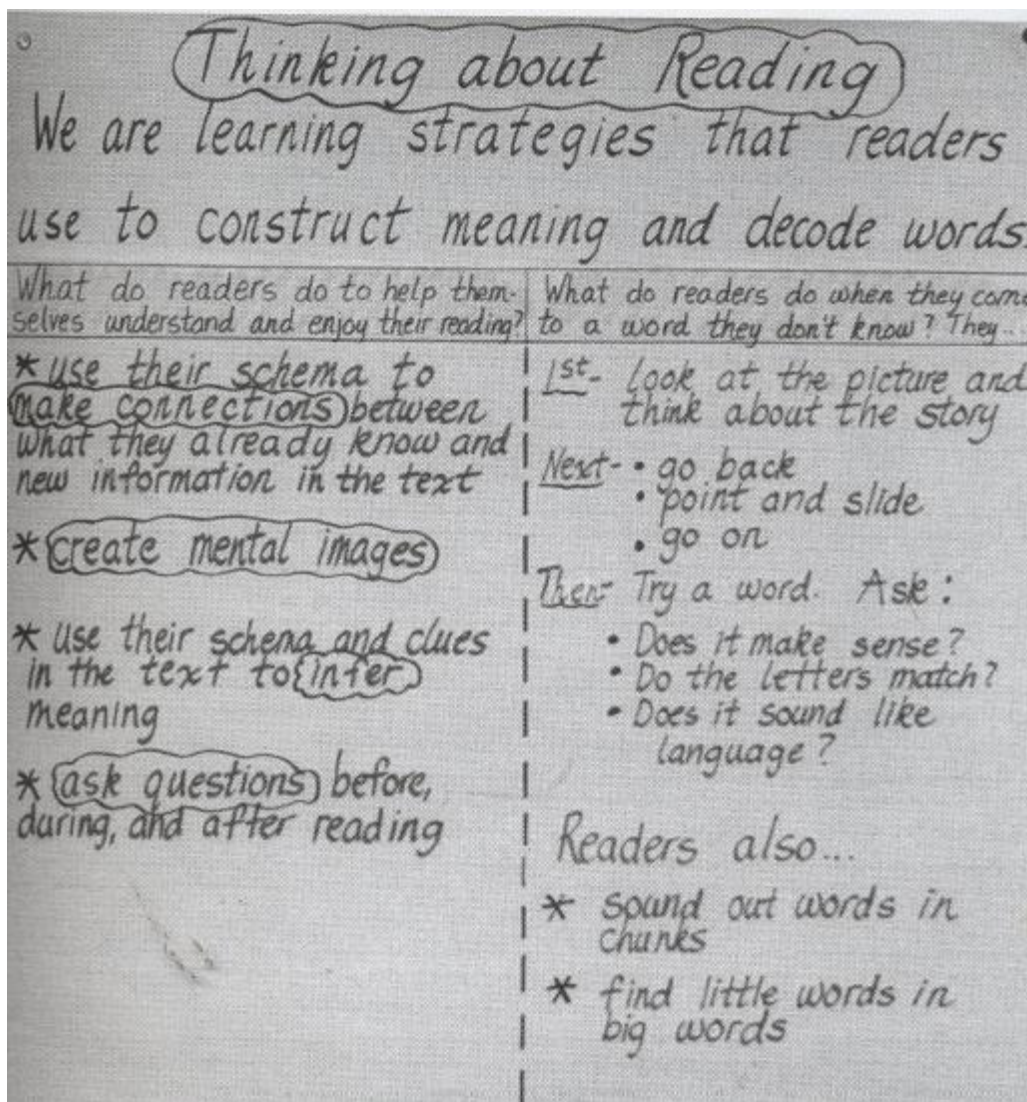
6.4. The problem of assessment

As was discussed at the beginning of this thesis, the way in which comprehension is assessed will have a bearing on the outcome of research into the effectiveness of reading comprehension instruction. Cain (2010) emphasises the difficulty:

> Because reading comprehension is a complex process, and because so many factors affect students' acquisition of comprehension skills, researchers and practitioners may approach the topic of the assessment of reading comprehension with trepidation (Carlisle & Rice, 2004, p. 521).

The researcher needs to think carefully about how comprehension is measured and whether the test used will assess the aspect of comprehension in which they are interested. Thus, if the interest is in a child's response to connected text then this will require the selection of a test which uses passages of an adequate length rather than just sentences. Similarly, as Cain also points out, some tests are over-reliant on background knowledge (the 'what colour was the banana?' problem). Therefore, in the second and third studies in the present thesis the NARA was replaced by the YARC, since the authors aimed to reduce over-reliance on background knowledge. The NARA may also assess literal comprehension, rather than inferential (Hurry & Doctor, 2007). The authors of the YARC aimed to remedy this by providing questions which measure different types of comprehension skill. Some questions tap information that is explicit in the text, whilst others require a variety of inferencing skills. This can provide valuable information on an individual basis, but as the different types of questions are not equally distributed across passages or forms (as the authors acknowledge) it is not possible to use an analysis of correct responses by question type.

If RT does improve inferencing skills then a test which measures these skills more specifically is needed. Tennent (2010) devised such a test for his study, and although the test suffered from ceiling effects, such a test would have been of value in the present research. The QRI responses did distinguish between explicit and implicit question types, but there is a much greater variety among inferences than is incorporated there. Indeed, a standardised test which specifically identifies problems with inferencing in a manner which is amenable to group effects (i.e., consistent across passages and forms) would be of great benefit to researchers. There are inherent difficulties; as the authors of the YARC say, their passages generate different kinds of inference as the difficulty increases. Given that there is now more understanding of how inferencing skills develop however, this should be feasible.

247

6.5. Contribution

As the literature review highlighted, despite a vast amount of research, there are no studies which provide quantitative data concerning RT's effectiveness with children of ages 7 to 8 and 9 to 10, in a whole class situation. Brooks (2013) talks about "a general dearth of research on improving comprehension" (p.109) and Cain (2010) suggested that more research is needed to determine how effective strategy instruction is for different age groups (p.185). Tennent (2010) demonstrated the effectiveness of RT for average comprehenders in his study of small group instruction with the same age group as sampled Studies 2 and 3 in the present research (i.e., 7 to 8 years). This has now been extended to whole class teaching through the use of heterogeneous small groups. The present research has helped to show that for these younger children RT may indeed be effective. In her book about reading development Cain also expressed the hope for more research determining which types of reader benefit from such instruction, and whilst this research has not shown the effectiveness of RT for all poor comprehenders, it has shown that it can form a valuable part of normal instruction, which may benefit average and good comprehenders too.

6.5.1. Reciprocal Teaching as part of normal classroom practice

In the past, RT had been promoted as an intervention for children with comprehension difficulties, and as the literature review shows, it has largely been used for small groups of children whose comprehension is not as good as their word reading skills would indicate. As the case study material in Study 2 showed, one poor comprehender appeared to improve as a result of the intervention, but in Study 3 we also saw a non-responder who had a deficit in comprehension compared to word reading ability; it was suggested that this child might have additional difficulties with language and vocabulary. Following recent research on the heterogeneous nature of comprehension deficits (Cain 2010), it may be that that the place for RT is as a part of normal comprehension instruction within the classroom, and that other interventions are more appropriate for poor comprehenders whose difficulties arise from more specific language difficulties. The results of the three studies reported in the thesis indicate that it is practically possible to deliver RT to the whole class. The second two studies have shown that classroom teachers can deliver the instruction with the kind of input that might be expected from in-service training. As a classroom teacher I would use RT sessions in the time allocated for Guided Reading activities. I would use heterogeneous groups, to expose poorer readers (both poor comprehenders and poor decoders) to a higher grade of reading material than they are exposed to in homogeneous groups. However, by careful

monitoring I would be alert to the possibility that among the expected 10% of poor comprehenders in the class there may be children who need additional support above and beyond that offered by RT, in the same way that there may be 10% of children who need additional support with phonics above and beyond that offered by good practice in the classroom as a whole. The present studies have also shown the value of different methods of assessment, and the role of think-alouds and strategy interviews will be discussed next.

6.5.2. How think-alouds could help in the classroom

Listening to children think aloud could give insight into their thought processes in a way which is not possible with the usual 'listening to reading' which still takes place daily in classrooms. For example, pupil 3:6 was an active reader from the outset, using a variety of strategies, but particularly strong at comparing her background knowledge to what she was reading and commenting on what she had read that was contradictory to her perceptions. However, at Times 2 and 3 there was a tendency to talk about more associative recollections, which did not contribute to increasing her understanding. For example, the octopus changing colour reminded her of a necklace she had been given which also changed colour (Time 2) and the passage about the Nile (Time 3) reminded her of a visit to Egypt. Rich as these associations are she did not use them to further her understanding by relating those experiences to what she was reading. Her reading comprehension could be aided by encouraging her to think more deeply about how experiences she has had can help her understanding in a way that is more than just associative. Without a think-aloud we would never know what she was thinking about as she was reading, nor how to help her to use these associations to enrich her understanding.

6.5.3. Strategy Interviews

Studies 2 and 3 have shown how strategy interviews give rich information about what children are doing and thinking as they read, and that they can provide valuable information about how to help a child cultivate and improve an active involvement with text. I would now use a strategy interview at the beginning of each year to determine which strategies are being used by the readers in my class. This would give an individual and a global picture of the children and the class as readers and enable me to plan the necessary support for each child and the class as a whole.

6.6. Implications for the class teacher

When I began working on this thesis I was concerned about how best to teach comprehension and how to help those children I had seen who could decode well, but who did not understand what they were reading. I think I have learned a lot about teaching comprehension, but that the problem of teaching poor comprehenders still requires more research. As we have seen they are a heterogeneous group and there may be no simple answer to this complex question and there is certainly no quick fix, as Compton et al. (2014) have intimated. One group of poor comprehenders may struggle with inhibiting alternative meanings of words (Borella, Carretti, & Pelegrina, 2010), whilst for another group, combining knowledge building and inference training as Compton et al. suggest may be effective, just as oral language training (Clarke, et al., 2014) may be effective for others. In the same way, RT has been shown to be effective for some poor comprehenders but not all. Trying to identify the source of poor comprehension in an individual may be the key for a classroom teacher.

6.7. Implications for instruction in the UK

If RT has been shown to be effective, and we are now clearer about the mechanisms behind its success, then how is RT to be promoted in the classroom? There is a problem in that teachers are very much concerned with teaching what they have been told to teach. For example, in the USA, Miller (2013) has found that the introduction of the Common Core means that teachers are reluctant to teach things which are not made explicit in the standards – for example, she cites teachers saying "I'd like to try this lesson, but we're doing Common Core, and well, it doesn't really include teaching children about mental images". One way to use RT within the existing structure in the UK is to use the programme during Guided Reading. Guided Reading was intended to replace the traditional 'listening to children read' practices which have been prevalent for decades. When used as it was intended, Guided Reading provides an ideal opportunity for teaching reading comprehension strategies (Whitehead, 2002), but in my experience in three different schools, Guided Reading has become listening to children read in a group. The expectations of children within these groups have remained that they are there for them to read aloud, just as they have always done, except that now they take in turns to read to the teacher rather than individually. Evidence of this was provided by a pupil in Study 3, as previously quoted, who when asked what she was thinking about when reading her paragraph, indignantly replied to the researcher "I wasn't thinking anything, this is a reading lesson, not

a thinking lesson. We don't do thinking with Mrs N". At the end of the lesson she commented "But we've only read four pages. We usually do lots more than that." The expectation was that reading would take place, not talking about reading, and that what mattered was the amount of reading done in a certain amount of time, not how much had been learned or understood.

6.8. Dissemination

Despite 30 years of studies on comprehension instruction and the use of strategies, the findings do not seem to have found their way into the classroom. This gap between what we know and what is being done in classrooms is acknowledged in the USA (Klingner, Urbach, Golos, Brownell, & Menon, 2010) and the UK (Parker & Hurry, 2007). Where strategy instruction programmes exist, as in some core reading programmes in the USA, they recommend a long list of strategies, do not utilise a gradual release-of-responsibility model, and do not provide the necessary amount of practice (Dewitz, et al., 2009). In the UK, even with the advent of Guided Reading (which came with the introduction of the National Literacy Strategy in 1998) it would appear that a large amount of the time allocated is spent in listening to children read aloud, the very practice it sought to replace with a focus on interpretative and critical comprehension (Fisher, 2008). The only difference is that now children read aloud in groups rather than individually. As we have seen, the opportunity exists within the time provided for Guided Reading to implement RT, but more would need to be done to promote its use amongst class teachers, for example through in-service training.

Finally, besides the initial desire to show whether or not RT could be an effective way of improving reading comprehension in whole class situations, these studies have helped me develop as a teacher. The next section will outline some of the ways in which these studies will influence my practice in the future.

6.9. The role of the teacher; what I have learned that will inform my practice in the future

In addition to discovering how effective RT can be and increasing my understanding of the best ways in which to implement it in my own classroom, I have also learned about myself as a teacher. One of the most important things I have learned from these studies is the importance of realising that I do not have all the answers, and neither should I have all the

questions. This was underlined by Miller (2002), when she discussed how to involve children in deciding for themselves about classroom rules:

> Why not just post a list of rules on the first day and be done? I remember those days, but that was when the room was mine, not ours; that was when I was the only teacher, and they were the only learners; that was when I asked all the questions, and had all the answers , too. (p. 33)

RT is about recognising that children can be teachers, and that they can teach us. They can teach us about ways to respond to a book or a way of understanding a book which is completely different to our own. They can pose questions we had not thought of, and they can surprise us with their insight. I remember how surprised the teacher was in Study 2 at the things of which the children were capable and her comments about how she would approach her teaching of reading in the future in the light of that experience. A recent article by Miller shows that this is where her thinking has gone too, more than ten years after the book cited above:

> My expectations for children and for me as a teacher and learner are forever and significantly raised. I still remember when I wasn't sure how to go about determining importance or synthesizing information, let alone showing children as young as 6, 7 and 8 how to go about doing it. However, with time and patience and practice, learn we did. (Miller, 2013, p. 360)

As a class teacher I would use RT for heterogeneous groups within the time allocated for Guided Reading. I would try using puppets and gesture in addition to the methods used in these studies, and I would make RT, including visualisation, a part of the pattern of teaching that continues throughout the week. Connections need to be made to enable the transfer of strategies to be achieved, and modelling should be something that happens at every available opportunity. Strategy teaching should not be confined to the RT lessons, but should permeate reading aloud, listening to children read, discussing a maths problem – in short, whenever reading takes place.

Miller also clarified a feeling I had about readers: "I've learned that the best decoders aren't necessarily the most thoughtful readers, nor are the most thoughtful readers necessarily the best decoders" (2002, p. 80). In a rush to teach decoding we must be careful to bear this in mind. Decoding is not an end in itself. Reading is more than decoding.

As the Simple View of Reading shows, reading is about comprehension too, and RT may be able to help us redress the balance.

References

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print.* Cambridge, MA: MIT Press.

Afflerbach, P., & Johnston, P. (1984). Research methodology: On the use of verbal reports in reading research. *Journal of Reading Behavior, 16*, 307 - 322. doi: 10.1080/10862968409547524

Alexander, R. (2004). Towards dialogic teaching: Rethinking classroom talk. York: Dialogos.

Alfassi, M. (1998). Reading for meaning: The efficacy of reciprocal teaching in fostering reading comprehension in high school students in remedial reading classes. *American Educational Research Journal, 35*(2), 309-332.

Alison, H. P., & Paris, S. G. (2003). Assessing narrative comprehension in young children. *Reading Research Quarterly, 38*(1), 36-76. doi: 10.1598/RRQ.38.1.3

Allan, V. (1976). *Children as teachers: Theory and research on tutoring*. New York: Academic.

Allington, R. (2009). *What really matters in fluency: Research-based practices across the curriculum*. Boston: Allyn and Bacon.

Anderson, R., Wilson, P., & Fielding, L. (1988). Growth in reading and how children spend their time outside school. *Reading Research Quarterly, 13*, 285 - 303. doi: 10.1598/RRQ.23.3.2

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

Andrews, J. (1985). Deaf children's acquisition of pre-reading skills using the reciprocal teaching procedure. Technical Report No. 350. Urbana, University of Illinois: Center for the Study of Reading.

Antoniou, F., & Souvignier, E. (2007). Strategy instruction reading comprehension; an intervention study for students with learning disabilities. *Learning Disabilities: A Contemporary Journal, 5*(1), 41 -57.

Applegate, M. D., Quinn, K. B., & Applegate, A. (2002). Levels of thinking required by comprehension questions in informal reading inventories. *The Reading Teacher, 56*, 174 - 180.

Baker, H. J., & Leland, B. (1959). *Detroit Test of Learning Aptitude*. Indianapolis: Bobbs-Merrill.

Baker, L. (1984). Children's effective use of multiple standards for evaluating their comprehension. *Journal of Educational Psychology, 76*(4), 588. doi: 10.1037/0022-0663.76.4.588

Baker, L. (2002). Metacognition in comprehension instruction. In C. Block & M. Pressley (Eds.), *Comprehension instruction: Research-based best practices* (pp. 77 -95). New York: Guildford.

Baker, L., & Wigfield, A. (1999). Dimensions of children's motivation for reading and their relations to reading activity and reading achievement. *Reading Research Quarterly, 34*(4), 452 - 477. doi: 10.1598/RRQ.34.4.4

Baker, L., & Zimlin, L. (1989). Instructional effects on children's use of two standards for evaluating their comprehension. *Journal of Educational Psychology, 81*(3), 340 - 346. doi: 10.1037/0022-0663.81.3.340

Barnes, M. A., Dennis, M., & Haefele-Kalvaitis. (1996). The effects of knowledge availability and knowledge accessibility on coherence and elaborative inferencing in children from six to fifteen years of age. *Journal of Experimental Child Psychology, 61*, 216 - 241. doi: 10.1006/jecp.1996.0015

Barrs, M., Pradl, G., Hall, K., & Dombey, H. (2008). Literacy: Whose complex activity? *Literacy, 42*(2), 55 - 58. doi: 10.1111/j.1741-4369.2008.00498.x

Baumann, J. F., Edwards, E. C., Boland, E. M., Olejnik, S., & Kame'enui, E. J. (2003). Vocabulary tricks: Effects of instruction in morphology and context on fifth-grade students' ability to derive and infer word meaning. . *American Educational Research Journal, 40*(2), 447 - 494. doi: 10.3102/00028312040002447

Bell, N. (1991). *Visualising and verbalising for language comprehension and thinking* (Revised ed.). San Luis Obispo, CA: Gander Educational Publishing.

Bereiter, C., & Scardamalia, M. (2003). Learning to work creatively with knowledge. In E. De Corte, L. Verschaffel, N. Entwistle & K. van Merrienboer (Eds.), *Powerful learning environments*. Amsterdam: Elsevier.

Biancarosa, C., & Snow, C. E. (2006). *Reading next: A vision for action and research in middle and high school literacy: A report to Carnegie Corporation of New York (2nd ed.)*. Washington, DC: Alliance for Excellent Education.

Biemiller, A., & Meichenbaum, D. (1992). The nature and nurture of the self-directed learner. *Educational Leadership, 50*(October), 75 - 80.

Block, C., & Paris, S. G. (Eds.). (2008). *Comprehension instruction: Research-based best practices* (2nd ed.). New York, London: Guildford Press.

Block, C., & Pressley, M. (2007). Best practices in teaching comprehension. In L. Gambrell, L. Morrow & M. Pressley (Eds.), *Best practices in literacy instruction* (3rd ed., pp. 220 - 242). New York: Guildford Press.

Borella, E., Carretti, B., & Pelegrina, S. (2010). The specific role of inhibition in reading comprehension in good and poor comprehenders. *Journal of Learning Disabilities, 43*(6), 541 - 552. doi: 10.1177/0022219410371676

Brady, P. L. (1990). *Improving the reading comprehension of middle-school students through reciprocal teaching and semantic mapping strategies.* Unpublished doctoral thesis. University of Alaska.

Braun, C., Rennie, B., & Labercane, G. (1986). A conference approach to the development of metacognitive strategies. *National Reading Conference Yearbook, 35*, 204 - 209.

Brooks, G. (2002). What works for children with literacy difficulties: The effectiveness of intervention schemes. London: DfES Research Report No. RR380.

Brooks, G. (2013). What works for children and young people with literacy difficulties: The effectiveness of intervention schemes: Dyslexia-SpLD Trust.

Broussard, S. C., & Garrison, M. E. B. (2004). The relationship between classroom motivation and academic achievement in elementary-school-aged children. *Family and Consumer Sciences Research Journal, 33*(2), 106 - 120. doi: doi:10.1177/1077727X04269573

Brown, A. L., & Campione, J. C. (1996). Psychological theory and the design of innovative learning environments: On procedures, principles, and systems. In L. Schauble & R. Glaser (Eds.), *Innovations in learning: New environments for education* (pp. 289 - 325). Mahwah, NJ: Erlbaum.

Brown, A. L., & Palincsar, A., S. (1985). Reciprocal teaching of comprehension strategies: A natural history of one program for enhancing learning. Cambridge, Mass: Illinois University, Urbana, Center for the Study of Reading.

Brown, A. L., Palincsar, A., S, & Armbruster, B. B. (1984). Instructing comprehension-fostering activities in interactive learning situations. In H. Mandl, N. Stein & T. Trabasso (Eds.), *Learning and comprehension of text* (pp. 255 - 286). Hillsdale, NJ: Lawrence Erlbaum Associates.

Brown, A. L., & Palincsar, A. S. (1989). Guided cooperative learning and individual knowledge acquisition. In L. B. Resnick (Ed.), *Knowing and learning: Issues for a cognitive psychology of learning. Essays in honor of Robert Glaser* (pp. 393 - 451). Hillsdale, NJ: Erlbaum.

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

Bruner, J. (1978). The role of dialogue in language acquisition. In A. Sinclair, R. Jarvella & W. Levelt (Eds.), *The child's conception of langauge* (pp. 241 - 256). New York: Springer-Verlag.

Bryant, D. P., Sharon, V., Linan-Thompson, S., Ugel, N., Allison, H., & Hougen, M. (2000). Reading outcomes for students with and without reading disabilities in general education middle-school content area classes. *Learning Disability Quarterly, 23*(4), 238 - 252. doi: 10.2307/1511347

Cain, K. (1999). Ways of reading: How knowledge and use of strategies are related to reading comprehension. *British Journal of Developmental Psychology, 17*(2), 293 - 309. doi: 10.1348/026151099165285

Cain, K. (2010). *Reading development and difficulties*. Oxford: Wiley Blackwell.

Cain, K., & Oakhill, J. (2006). Profiles of children with specific reading comprehnsion difficulties. *British Journal Of Educational Psychology, 76*, 683 - 696. doi: DOI:10.1348/000709905SX67610

Cain, K., & Oakhill, J. (2011). Matthew effects in young readers. *Journal of Learning Disabilities, 44*(5), 431-443. doi: 10.1177/0022219411410042

Cain, K., & Oakhill, J. (Eds.). (2007). *Children's comprehension problems in oral and written language: A cognitive perspective.* New York: Guildford Press.

Cain, K., Oakhill, J., Barnes, M., & Bryant, P. (2001). Comprehension skill, inference making ability, and the relation to knowledge. *Memory and Cognition, 29*(6), 850 - 859. doi: 10.3758/BF03196414

Caldwell, J., & Leslie, L. (2010). Thinking aloud in expository text: processes and outcomes. *Journal of Literacy Research, 42*(3), 308 - 340. doi: 10.1080/1086296X.2010.504419

Canney, G., & Winograd, P. (1979). *Schemata for reading and reading comprehension performance (Technical Report No. 120)*. Urbana, University of Illinois: Center for the Study of Reading.

Carlisle, J. F., & Rice, M. S. (2004). Assessment of reading comprehension. In C. A. Stone, E. R. Silliman, B. J. Ehren & K. Apel (Eds.), *Handbook of language and literacy: Development and disorders* (pp. 521 -525). New York: The Guildford Press.

Carter, C. J. (1997). Why reciprocal teaching? *Educational Leadership, 54*(6), 64 - 68.

Carter, C. J., & Fekete, D. (2001). *Reciprocal teaching: The application of a reading improvement strategy on urban students in Highland Park, Michigan*. Geneva, Switzerland: International Bureau of Education.

Catts, H., & Adolf, S. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language and Hearing Research, 49*(April), 278 - 293. doi: 10.1044/1092-4388(2006/023)

Center, Y., Freeman, L., Robertson, G., & Outhred, L. (1999). The effect of visual imagery training on the reading and listening comprehension of low listening comprehenders in Year 2. *Journal of Research in Reading, 22*(3), 241 - 256. doi: 10.1111/1467-9817.00088

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132*, 354 - 380. doi: 10.1037/0033-2909.132.3.354

Cervetti, G., Pearson, P., Bravo, M., & Barber, J. (2006). Reading and writing in the service of inquiry-based science. In R. Douglas, M. Klentschy & K. Worth (Eds.), *Linking science and literacy in the K-8 classroom* (pp. 221 - 244). Arlington, VA: NSTA Press.

Chan, L., & Cole, G. (1986). The effects of comprehension monitoring training on the reading competence of learning disabled and regular class students. *Remedial and Special Education, 7*(4), 33 - 40. doi: 10.1177/074193258600700407

Clarke, P. J., Henderson, L. M., & Truelove, E. (2010). The poor comprehender profile: understanding and supporting individuals who have difficulties extracting meaning from text. *Advances in Child Development and Behavior: Developmental Disorders and Interventions, 39*(39), 79-129. doi: 10.1016/B978-0-12-374748-8.00003-2

Clarke, P. J., Snowling, M. J., Truelove, E., & Hulme, C. (2010). Ameliorating children's reading comprehension difficulties. *Psychological Science, 21*(8), 1106-1116. doi: 10.1177/0956797610375449

Clarke, P. J., Truelove, E., Hulme, C., & Snowling, M. J. (2014). *Developing reading comprehension*. Chichester, West Sussex: Wiley-Blackwell.

Clarke, P. J., Truelove, E., Snowling, M. J., & Hulme, C. (2011). Using speaking and listening to support the development of reading comprehension skills. PowerPoint presentation. Retrieved April 4th, 2014, from http://readingformeaning.co.uk/overview1/

Clay, M. (1973). *Reading: The patterning of complex behaviour*. Auckland, NZ: Heinemann Educational Books.

Cohen, E. G. (1994). Restructuring the classroom: conditions for productive small groups. *Review of Educational Research, 61*(1), 1 - 35.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.

Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education* (6th ed.). Abingdon: Routledge.

Cohen, P. A., Kulik, J. A., & Kulik, C. (1982). Educational outcomes of tutoring: A meta-analysis of findings. *American Educational Research Journal, 19*, 237 - 248. doi: 10.3102/00028312019002237

Coley, J., DePinto, T., Craig, S., & Gardner, R. (1993). From college to classroom: Three teachers' accounts of their adaptations of reciprocal teaching. *The Elementary School Journal, 94*(2), 255 - 266. doi: 10.1086/461765

Collins, A., & Smith, E. (1982). Teaching the process of reading comprehension. In D. K. Detterman & R. J. Sternberg (Eds.), *How and how much can intelligence be increased* (pp. 173 - 185). Norwood, NJ: Ablex Publishing Company.

Collins, K. M. T., Onwuegbuzie, A. J., & Sutton, I. L. (2006). A model incorporating the rationale and purpose for conducting mixed methods research in special education and beyond. *Learning Disabilities: A Contemporary Journal, 4*, 67-100.

Compton, D., Miller, C., Elleman, A., & Steacy, L. (2014). Have we forsaken reading theory in the name of 'quick fix' interventions for children with reading disability. *Scientific Studies of Reading, 18*(1), 55 - 73. doi: 10.180/10888438.2013.836200

Conderman, G., Bresnahan, V., & Hedin, L. (2011). Promoting active involvement in today's classrooms. *Kappa Delta Pi Record, 47*(4), 174-180. doi: 10.1080/00228958.2011.10516587

Condron, D. J. (2008). An early start: Skill grouping and unequal reading gains in the elementary years. *Sociological Quarterly, 49*(2), 363-394. doi: 10.1111/j.1533-8525.2008.00119.x

Cooper, H., & Rosenthal, R. (1980). Statistical versus traditional procedures for summarising research findings. *Psychological Bulletin, 87*, 422 - 449. doi: 10.1037/0033-2909.87.3.442

Cummings, C., Stewart, M., & Block, C. (2005). Teaching several metacognitive strategies together to increase students' independent metacognition. In S. E. Israel, C. Block, K. L. Bauserman & K. Kinnucan-Welsch (Eds.), *Metacognition in literacy learning: Theory, assessment, instruction and professional development* (pp. 227 - 296). Mahwah, NJ: Erlbaum.

Cutting, L. E., & Scarborough, H. S. (2006). Prediction of reading comprehension: relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. *Scientific Studies of Reading, 10*(3), 277 - 300. doi: 10.1207/s1532799xssr1003_5

Davey, B. (1983). Think-aloud: Modelling the cognitive processes of reading comprehension. *Journal of Reading, 27*, 44 - 47.

Davis, D. S. (2011). Internalization and participation as metaphors of strategic reading development. *Theory into Practice, 50*(2), 100 - 106. doi: 10.1080/00405841.2011.558434

DCSF. (2006). *Primary National Strategy*. Department of Children, Schools and Families.

De Corte, E., Verschaffel, L., & Van De Ven, A. (2001). Improving text comprehension strategies in upper primary school children: A design experiment *British Journal of Educational Psychology, 71*, 531 -559. doi: 10.1348/000709901158668

Demmrich, A. (2005). *Improving reading comprehension by enhancing metacognitive competences: an evaluation of the reciprocal teaching method.* PhD thesis, University of Potsdam. Retrieved June 28th, 2014 from http://opus.kobv.de/ubp/volltexte/2005/524/pdf/demmrich_neu.pdf

Dempster, F. N. (1988). The spacing effect: A case study in the failure to apply the results of psychological research. *American Psychologist, 43*, 627 - 634. doi: 10.1037/0003-066X.43.8.627

Dempster, F. N. (1989). Spacing effect and their implications for theory and practice. *Educational Psychology Review 1*, 309 - 330. doi: 10.1007/BF01320097

Denny, T., & Weintraub, S. (1963). Exploring first-graders' concepts of reading. *Reading Teacher, 16*, 363 - 365.

Denny, T., & Weintraub, S. (1966). First-graders' responses to three questions about reading. *Elementary School Journal, 66*, 441 - 448. doi: 10.1086/460326

Department for Children Schools and Families. (2006). Developing reading comprehension: Comprehension strategies. Primary framework for literacy and mathematics Retrieved April 4th, 2014, from http://webarchive.nationalarchives.gov.uk/20110902085210/http://www.nsonline.org.uk/node/19579?uc=force_uj

Department for Children Schools and Families. (2010). Key leaflet: Guided reading Ref: 00464-2010PDF-EN-02 Retrieved April 4th, 2014, from http://webarchive.nationalarchives.gov.uk/search/?lang=en&noneW=&format=all&site=&search_type=category&x=40&y=21&exactW=&where=text&query=Department+for+Children+School+and+Families.+%282010%29.+Key+leaflet%3A+Guided+reading+&x=43&y=12

Department for Education. (2010). The Importance of Teaching: Schools White Paper 2010 Retrieved April 4th, 2014, from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/175429/CM-7980.pdf

Department for Education and Skills. (2005a). Understanding reading comprehension:1- What is reading comprehension? Retrieved April 4th, 2014, from http://www.teachfind.com/national-strategies/understanding-reading-comprehension-1-%E2%80%93-what-reading-comprehension

Department for Education and Skills. (2005b). Understanding reading comprehension:2- Strategies to develop reading comprehension Retrieved April 4th, 2014, from http://www.teachfind.com/national-strategies/understanding-reading-comprehension-2-%E2%80%93-strategies-develop-reading-comprehension

Department for Education and Skills. (2005c). Understanding reading comprehension: 1 Retrieved April 4th, 2014, from http://nationalstrategies.standards.dcsf.gov.uk/node/64507?uc=force_uj

Dermody, M. (1988). *Metacognitive strategies for development of reading comprehension for younger children.* Paper presented at the Annual meeting of the American Association of Colleges for Teacher Education, New Orleans, LA.

Dewitz, P., Jones, J., & Leahy, S. (2009). Comprehension strategy instruction in core reading programs. *Reading Research Quarterly, 44*(2), 102 -126. doi: 10.1598/RRQ.44.2.1

Dole, J. A., Duffy, G. G., Roehler, L. R., & Pearson, P. D. (1991). Moving from the old to the new: Research on reading comprehension instruction. *Review of Educational Research, 61*((2)), 239 - 264. doi: 10.3102/00346543061002239

Donovan, M. S., Wigdor, A. K., & Snow, C. E. (2003). *Strategic education research partnership*. Washington, DC: The National Academies Press.

Duff, F. J., & Clarke, P. J. (2011). Practitioner Review: Reading disorders: what are the effective interventions and how should they be implemented and evaluated? *Journal of Child Psychology and Psychiatry, 52*(1), 1 -108. doi: 10.1111/j.1469-7610.2010.02310.x

Duffy, G., Miller, S., Howerton, S., & Williams, J. B. (2010). Comprehension instruction - merging two historically antithetical perspectives. In D. Wyse, R. Andrews & J. Hoffman (Eds.), *The Routledge international handbook of English, language and literacy teaching* (pp. 58 - 73). Abingdon, Oxon: Routledge.

Duke, N. K., & Carlisle, J. F. (2011). The development of comprehension. In M. Kamil, P. Pearson, E. Moje & P. Afflerbach (Eds.), *Handbook of Reading Research* (Vol. IV, pp. 199 - 288). New York: Routledge.

Durkin, D. (1978). What classroom observations reveal about reading comprehension instruction. *Reading Research Quarterly, 14*((4)), 481 - 533. doi: 10.1598/RRQ.14.4.2

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values and goals. *Annual Review of Psychology, 53*(1), 109-132. doi: doi:10.1146/annurev.psych.53.100901.135153

El-Dinary, P. B., Pressley, M., Coy-Ogan, L., & Schuder, T. (1994). The teaching practices of transactional strategies instruction teachers as revealed through collaborative interviewing. *Technical Report, 23*.

El-Dinary, P. B., & Schuder, T. (1993). Seven teachers' acceptance of Transactional Strategies Instruction during their first year using it. *The Elementary School Journal, 94*(2), 207 - 219. doi: 10.1086/461761

Fisher, A. (2008). Teaching comprehension and critical literacy: Investigating guided reading in three primary classrooms. *Literacy, 42*(1), 19 - 28. doi: 10.1111/j.1467-9345.2008.00477.x

Fisher, D., & Frey, N. (2008). What does it take to create skilled readers? Facilitating the transfer and application of strategies. *Voices from the Middle, 15*(4).

Fuchs, D., & Fuchs, L. S. (2005). Peer-assisted learning strategies: Promoting word recognition, fluency, and reading comprehension in young children. *Journal Of Special Education, 39*(1), 34-44. doi: 10.1177/00224669050390010401

Fuchs, D., Fuchs, L. S., Mathes, P. G., & Simmons, D. C. (1997). Peer-assisted learning: making classrooms more responsive to diversity. *American Educational Research Journal, 34*(1), 174 - 206. doi: 10.3102/00028312034001174

Fuchs, D., Fuchs, L. S., Thompson, A., Svenson, E., Yen, L. L., Al Otaiba, S., et al. (2001). Peer-assisted learning strategies in reading - Extensions for kindergarten, first grade, and high school. *Remedial And Special Education, 22*(1), 15-21. doi: 10.1177/074193250102200103

Fuchs, L. S., Fuchs, D., Karns, K., Hamlett, C.L., Dutka, S., Katzaroff, M. (1996). The relation between student ability and the quality and effectiveness of explanations. *American Educational Research Journal, 33*(3), 631 - 664. doi: 10.3102/00028312033003631

Funder, D. C., Levine, J. M., Mackie, D. M., Morf, C. C., Vazire, S., & West, S. G. (2013). Improving the dependability of research in personality and social psychology: Recommendations for research and educational practice. *Personality and Social Psychology Review*. doi: 10.1177/1088868313507536

Galloway, A. (2003). *Improving reading comprehension through metacognitive strategy instruction. Evaluating the evidence for the effectiveness of the reciprocal teaching procedure.* Doctoral thesis. University of Nebraska.

Gambrell, L., & Bales, R. (1986). Mental imagery and the comprehension monitoring performance of fourth- and fifth-grade poor readers. *Reading Research Quarterly, 21*(4), 454 - 464. doi: 10.2307/747616

Gambrell, L., & Bales, R. (1987). Visual imagery: A strategy for enhancing listening, reading, and writing. *Australian Journal of Reading, 10*(3), 147 - 153

Gambrell, L., & Koskinen, P. (2002). Imagery: A strategy for enhancing comprehension In C. Block & M. Pressley (Eds.), *Comprehension instruction: Research-based best practices* (pp. 305 - 319). New York: Guildford.

Garner, R. (1981). Monitoring of passage inconsistency among poor comprehenders. A preliminary test of the 'Piecemeal Processing' explanation. *Educational Research Quarterly, 75*, 159 - 162.

Garner, R. (1992). Self-regulated learning, strategy shifts, and shared expertise: reactions to Palincsar and Klenk. *Journal of Learning Disabilities, 25*(4), 226 - 229. doi: 10.1177/002221949202500403

Gersten, R., Keating, T., Yovanoff, P., & Harniss, M. (2001). Working in special education: Factors that enhance special educators' intent to stay. *Exceptional Children, 64*(7), 549 - 567. doi: 10.1177/001440290106700408

Gillies, R. M. (2006). Teachers' and students' verbal behaviours during cooperative and small-group learning. *British Journal Of Educational Psychology, 76*, 271-287. doi: 10.1348/000709905X52337

Gilroy, A., & Moore, D. W. (1988). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities with ten primary school girls. *Educational Psychology, 8*(1), 41 - 49. doi: 10.1348/000709905X52337

Glaser, C., & Brunstein, J. C. (2007). Improving fourth-grade students' composition skills: Effects of strategy instruction and self-regulation procedures. *Journal of Educational Psychology, 99*(2), 297 - 310. doi: 10.1037/0022-0663.99.2.297

Gough, P., & Tunmer, W. (1986). Decoding, reading and reading disability. *Remedial and Special Education, 7*, 6 - 10. doi: 10.1177/074193258600700104

Graesser, A. C. (2007). An introduction to strategic reading comprehension. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions and technologies*. New York: Lawrence Erlbaum Associates.

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

Greenway, C. (2002). The process, pitfalls and benefits of implementing a reciprocal teaching intervention to improve the reading comprehension of a group of Year 6 pupils. *Educational Psychology in Practice, 18*(2), 113-137. doi: 10.1080/02667360220144557

Guthrie, J. T., McGough, K., Bennett, L., & Rice, M. E. (1996). Concept-oriented reading instruction: An integrated curriculum to develop motivations and strategies for reading In L. Baker, P. Afflerbach & D. Reinking (Eds.), *Developing engaged readers in school and home communities* (pp. 165 - 190). Hillsdale, NJ: Erlbaum.

Guthrie, J. T., Taboada, A., & Coddington, C. (2007). Engagement practices for strategy learning in concept-oriented learning instruction. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions and technologies* (pp. 241 - 266). New York: Lawrence Erlbaum Associates.

Guthrie, J. T., Van Meter, P., Hancock, G. R., Alao, S., Anderson, E., & McCann, A. (1998). Does concept-oriented reading instruction increase strategy use and conceptual learning from text? *Journal of Educational Psychology, 90*(2), 261 - 278. doi: 10.1037/0022-0663.90.2.261

Guthrie, J. T., Van Meter, P., McCann, A. D., Wigfield, A., Bennett, L., Poundstone, C., et al. (1996). Growth of literacy engagement: Changes in motivations and strategies during concept-oriented reading instruction. *Reading Research Quarterly, 31*(3), 306-332. doi: 10.1598/RRQ.31.3.5

Hacker, D. J. (2004). Self-regulated comprehension during normal reading. In R. Ruddel & N. Unrau (Eds.), *Theoretical models and processes of reading* (5th ed., pp. 755 - 779). Newark, DE: International Reading Association.

Hacker, D. J., & Tenent, A. (2002). Implementing reciprocal teaching in the classroom: Overcoming obstacles and making modifications. *Journal of Educational Psychology, 94*(4), 697-718. doi: 10.1037/0022-0663.94.4.699

Haertle, E. (1986). The valid use of student performance measures for teacher evaluation. *Educational evaluation and policy analysis, 8*(1), 45-61. doi: 10.3102/01623737008001045

Hager, W., & Hasselhorn, M. (2000). Psychologische Interventionsmassnahmen: Was sollen sie bewirken konnen. In W. Hager, J.-L. Patry & H. Brezing (Eds.), *Evaluation psychologischer Interventionsmassnahmen. Standards und Kriterien* (pp. 41 - 85). Bern: Huber.

Halberstam, M. (2008). *Reciprocal teaching: The effects on reading comprehension of third grade students.* Unpublished Doctor of Education thesis, Northcentral University, Arizona.

Hart, E. R., & Speece, D. L. (1998). Reciprocal teaching goes to College: Effects for post-secondary students at risk for academic failure. *Journal of Educational Psychology, 90*(4), 670 - 681. doi: 10.1037/0022-0663.90.4.670

Harvey, S., & Goudvis, A. (2007). *Strategies that work: teaching comprehension for understanding and engagement* (2nd ed.). York, ME: Stenhouse.

Hashey, J. M., & Connors, D. J. (2003). Learn from our journey: Reciprocal teaching action research. *Reading Teacher, 57*(3), 224 - 232.

Hirsch , E., D. (2006). The case for bringing content into the language arts block and for a knowledge-rich curriculum core for all children. *American Educator, 30*(1). Retrieved June 28th, 2014, from http://www.aft.org/pubs-reports/american_educator/issues/spring06/h

Hirsch , E. D. (2003). Reading comprehension requires knowledge of words and the world: Scientific insights into the fourth-grade slump and stagnant reading comprehension. *American Educator, 27*(10), 12 - 13, 16 - 22, 28 - 29.

Hobsbaum, A., Gamble, N., & Reedy, D. (2002). *Guiding reading: A handbook for teaching guided reading at Key Stage 2*. London: Institute of Education.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage.

Hurry, J., & Doctor, E. (2007). Assessing literacy in children and adolescents. *Child and Adolescent Mental Health, 12*(1), 38 - 45. doi: 10.111/j.1475-3588.2006.00420.x

Jackson, J. W. (1967). *Life in classrooms*. New York: Holt, Rinehart and Winston.

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

James-Burdumy, S., Deke, J., Lugo-Gil, J., Carey, N., Hershey, A., Gersten, R., et al. (2010). Effectiveness of selected supplemental reading comprehension interventions: Findings from two student cohorts-Executive summary (NCEE 2010-4016). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. .

James-Burdumy, S., Mansfield, W., Deke, J., Carey, N., Lugo-Gil, J., Hershey, A., et al. (2009). Effectiveness of selected supplemental reading comprehension interventions: Impacts on a first cohort of fifth-grade students-Executive summary (NCEE 2009-4058). Wasington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. .

Janssen, T., Braaksma, M., & Rijlaarsdam, G. (2006). Literary reading activities of good and weak students: A think-aloud study. *European Journal of Psychology of Education, 21*(1), 35 - 52. doi: 10.1007/BF03173568

Janssen, T., Braaksma, M., & Rijlaarsdam, G. (2010). Reading and teaching short stories, based on process studies and experimental research. In D. Wyse, R. Andrews & J. Hoffman (Eds.), *The Routledge International Handbook of English, Language and Literacy Teaching* (pp. 45 - 57). Abingdon, Oxon: Routledge.

Jenkins, J. R., Johnson, E., & Hileman, J. (2004). When is reading also writing: Sources of individual differences on the new reading performance assessments. *Scientific Studies of Reading, 8*(2), 125 - 151. doi: 10.1207/s1532799xssr0802_2

Joffe, V. L., Cain, K., & Maric, N. (2007). Comprehension problems in children with specific language impairment: Does mental imagery training help? *International Journal of Language and Communication Disorders, 42*(6), 648 - 664. doi: 10.1080/13682820601084402

Johns, J., & Ellis, D. (1976). Reading: Children tell it like it is. *Reading World, 16*(2), 115 - 128. doi: 10.1080/19388077609557350

Johnson-Glenberg, M. (2000). Training reading comprehension in adequate decoders/poor comprehenders:Verbal versus visual strategies. *Journal of Educational Psychology, 92*(4), 772 - 782. doi: 10.1037/0022-0663.92.4.772

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

Johnson, D., & Johnson, R. (1990). Cooperative learning and achievement. In S. Sharan (Ed.), *Cooperative learning: Theory and research* (pp. 173 - 202). New York: Praeger.

Keene, E., & Goudvis, A. (1995). Major point interview for readers. In E. Keene & S. Zimmerman (Eds.), *(1997) Mosaic of thought: Teaching comprehension in a reader's workshop* (pp. 227 - 235). Portsmouth, NH: Heinemann.

Keene, E., & Zimmermann, S. (1997). *Mosaic of thought; teaching comprehension in a reader's workshop*. Portsmouth, NH: Heinemann.

Kelly, M., Moore, D. W., & Tuck, B. (1994). Reciprocal teaching in a regular primary school classroom. *Journal Of Educational Research, 88*, 53 - 61. doi: 10.1080/00220671.1994.9944834

King, C. M., & Johnson, L. M. P. (1999). Constructing meaning via reciprocal teaching. *Reading Research And Instruction, 38*(3), 169-186. doi: 10.1080/19388079909558287

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*, 163 - 182. doi: 10.1037/0033-295X.95.2.163

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

Kirby, J. R., & Savage, R. S. (2008). Can the simple view of reading deal with the complexities of reading? *Literacy 42*(2), 75 - 82. doi: 10.1111/j.1741-4369.2008.00487.x

Klingner, J. K., Urbach, J., Golos, D., Brownell, M., & Menon, S. (2010). Teaching reading comprehension in the 21st century: A Glimpse at how Special Education teachers promote reading comprehension. *Learning Disability Quarterly, 33*(2), 59 - 74.

Klingner, J. K., & Vaughn, S. (1996). Reciprocal teaching of reading comprehension strategies for students with learning disabilities who use English as a second language. *The Elementary School Journal, 96*(3), 275. doi: 10.1086/461828

Klingner, J. K., Vaughn, S., Arguelles, M. E., Hughes, M. T., & Leftwich, S. E. (2004). Collaborative Strategic Reading: "Real-world" lessons from classroom teachers. *Remedial & Special Education, 25*(5), 291 - 302. doi: 10.1177/07419325040250050301

Klingner, J. K., Vaughn, S., Hughes, M. T., & Arguelles, M. E. (1999). Sustaining research-based practices in reading: A 3-year follow-up. *Remedial & Special Education, 20*, 263 - 274. doi: 10.1177/074193259902000502

Klingner, J. K., Vaughn, S., & Schumm, J. S. (1998). Collaborative Strategic Reading during social studies in heterogeneous fourth-grade classrooms. *The Elementary School Journal, 99*(1), 3 - 22. doi: 10.1086/461914

Kratochwill, T., & Stoiber, K. (2002). Evidence-based interventions within school psychology: Conceptual foundations of the procedural and coding manual of Division 16 and the Society for the Study of School Psychology Task Force. *School Psychology Quarterly, 17*(4), 341 - 389. doi: 10.1521/scpq.17.4.341.20872

Kucan, L., & Beck, I. L. (1996). Four fourth graders thinking aloud: An investigation of genre effects. *Journal of Literacy Research, 28*, 259 - 287. doi: 10.1080/10862969609547921

Kulhavy, R. W., & Swenson, I. (1975). Imagery instructions and the comprehension of text.

Kvale, S. (1996). *Interviews: An introduction to qualitative research interviewing*. Thousand Oaks, CA: Sage.

Labercane, G., & Battle, J. (1987). Cognitive processing strategies, self-esteem, and reading comprehension of learning disabled students. *B.C. Journal Of Special Education, 11*(2), 167 - 185.

Landsberger, H. (1958). *Hawthorne revisited*. Ithaca, NY: Cornell University.

Lave, J., & Wenger, E. (1991). *Situated learning: legitimate peripheral participation*. New York: Cambridge University Press.

Lederer, J. M. (2000). Reciprocal teaching of social studies in inclusive elementary classrooms. *Journal of Learning Disabilities, 33*(1), 91-106. doi: 10.1177/002221940003300112

Leslie, L., & Caldwell, J. (1995). *Qualitative Reading Inventory-II*. Glenview, IL: Scott, Foresman.

Leslie, L., & Caldwell, J. (2006). *Qualitative Reading Inventory-4*. Boston: Pearson.

Leslie, L., & Caldwell, J. (2011). *Qualitative Reading Inventory-5*. Boston: Pearson Education Inc.

Linden, M., & Wittrock, M. (1981). The teaching of reading comprehension according to the model of generative learning. *Reading Research Quarterly, 17*, 44 - 57. doi: 10.2307/747248

Lysynchuk, L., Pressley, M., & Vye, N. (1990). Reciprocal teaching improves standardised reading comprehension performance in poor comprehenders. *The Elementary School Journal, 1990*(5), 469 - 484. doi: 10.1086/461627

Mandel, E. (2007). *Vocabulary acquisition techniques for Grade One: An experimental investigation of shared reading vs. reciprocal teaching*. MA thesis. The Department of Education. Concordia University. Montreal. Retrieved June 28th, 2014, from http://spectrum.library.concordia.ca/975741/1/MR34708.pdf

Marcell, B., DeCleene, J., & Juettner, M. R. (2010). Caution! Hard hat area! Comprehension under construction: Cementing a foundation of comprehension strategy usage that carries over to independent practice. *Reading Teacher, 63*(8), 687-691. doi: 10.1598/RT.63.8.8

Marks, M., Pressley, M., Coley, J., Craig, S., Gardner, R., DePinto, T., et al. (1993). Three teachers' adaptations of reciprocal teaching in comparison to traditional reciprocal teaching. *The Elementary School Journal, 94*(2), 267 - 283. doi: 10.1086/461766

Marston, D., & Deno, S. L. (1995). Comparison of reading intervention approaches for students with mild disabilities. *Exceptional Children, 62*(1), 20.

Marton, F. (1976). On non-verbatim learning II. The erosion of a task-induced learning algorithm. *Scandinavian Journal of Psychology, 17*, 41 - 48. doi: 10.1111/j.1467-9450.1976.tb00209.x

Mayo, E. (1933). *The human problems of an industrial civilisation*. New York: Macmillan.

Mayo, E. (1949). *Hawthorne and the Western Electric Company: The social problems of an industrial civilisation*. London: Routledge and Kegan Paul.

McKenna, M., Kear, D., & Ellsworth, R. (1995). Children's attitudes towards reading: A national survey. *Reading Research Quarterly, 30*, 934 - 955. doi: 10.2307/748205

McLellan, E., MacQueen, K. M., & Neidig, J. L. (2003). Beyond the qualitative interview: Data preparation and transcription. *Field Methods, 15*(1), 63-84. doi: 10.1177/1525822X02239573

McNamara, D. S. (2007). *Reading comprehension strategies; Theory, interventions and technologies*. Mahwah, NJ: Erlbaum.

McNamara, D. S., & Magliano, J. (2009). Toward a comprehensive model of comprehension. In B. H. Ross (Ed.), *Psychology of Learning and Motivation* (Vol. 51, pp. 297-384): Academic Press.

Mercer, N., & Littleton, K. (2007). *Dialogue and the development of children's thinking*. London: Routledge.

Mergenthaler, E., & Stinson, H. E. (1992). Psychotherapy transcription standards. *Psychotherapy Research, 2*(2), 125 - 142. doi: 10.1080/10503309212331332904

Miller, D. (2002). *Reading with meaning: Teaching comprehension in the primary grades*. Portsmouth, ME: Stenhouse Publishers.

Miller, D. (2013). I can create mental images to retell and infer big ideas. *The Reading Teacher, 66*(5), 360-364. doi: 10.1002/trtr.01135

Mokhtari, K., & Reichard, C. (2002). Assessing students' metacognitive awareness of reading strategies. *Journal of Educational Psychology, 94*(2), 249 - 259. doi: 10.1037/0022-0663.94.2.249

Moore, P. J. (1988). Reciprocal teaching and reading comprehension: A review. *Journal of Research in Reading, 11*(1), 3-14. doi: 10.1111/j.1467-9817.1988.tb00144.x

Morgan, M. (Ed.). (2004). *Buffalo Bert*. London: A and C Publishers Ltd.

Myers, M., & Paris, S. (1978). Children's metacognitive knowledge about reading. *Journal of Educational Psychology, 70*, 680 - 690. doi: 10.1037/0022-0663.70.5.680

Myers, P. A. (2005). The Princess Storyteller, Clara Clarifier, Quincy Questioner, and the Wizard: Reciprocal teaching adapted for kindergarten students. *The Reading Teacher, 59*(4). doi: 10.1598/RT.59.4.2

Myhill, D. (2006). Talk, talk, talk; teaching and learning in whole class discourse. *Research Papers in Education, 21*(1), 19 - 41. doi: 10.1080/02671520500445425

Nation, K., Adams, J., Bowyer-Crane, C., & Snowling, M. (2004). Hidden language impairments in children; parallels between poor reading comprehension and specific language impairment. *Journal of Experimental Child Psychology, 73*, 139 - 158. doi: 10.1006/jecp.1999.2498

National Governors Association Center for Best Practices: Council of Chief State School Officers. (2010). *Common Core State Standards*. Washington, D.C.: National Governors Association Center for Best Practices, Council of Chief State School Officers.

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.

National Research Council. (2002). Scientific research in education. In R. J. Shavelson & L. Towne (Eds.). Washington, DC: The National Academies Press.

Neale, M. D. (1989). *The Neale analysis of reading ability: Revised British edition*. Windsor, United Kingdom: NFER-Nelson.

Nilsson, N. (2008). A critical analysis of eight Informal Reading Inventories *The Reading Teacher, 61*(7), 526 - 536. doi: 10.1598/rt.61.7.2

Oakhill, J., & Cain, K. (2007). Issues of causality in children's reading comprehension. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions and technologies* (pp. 47 - 72). New York: Lawrence Erlbaum Associates.

Oakhill, J., & Patel, S. (1991). Can imagery training help children who have comprehension problems? *Journal of Research in Reading*(2), 106 - 115. doi: 10.1111/j.1467-9817.1991.tb00012.x

Oczkus, L. (2003). *Reciprocal teaching at work: Strategies for improving reading comprehension* Newark, NJ: International Reading Association.

Oldfather, P., & Wigfield, A. (1996). Children's motivation to read. In L. Baker, P. Afflerbach & D. Reinking (Eds.), *Developing engaged readers in school and home communities* (pp. 89 - 113). Mahwah, NJ: Erlbaum.

Oliver, M. (1982). *Improving reading comprehension with mental imagery.* Paper presented at the Annual Meeting of the Washington Organization for Reading Development of the International Reading Association, Seattle, WA.

Oster, L. (2001). Using the think-aloud for reading instruction. *The Reading Teacher, 55*(1), 64 - 70.

Padron, Y. N. (1985). *Utilizing cognitive reading strategies to improve English reading comprehension of Spanish-speaking bilingual Students.* Unpublished doctoral dissertation. University of Houston.

Palincsar, A. S. (1982). *Improving the reading comprehension of junior high students through the reciprocal teaching of comprehension-monitoring strategies*. Unpublished doctoral dissertation. University of Illinois at Urbana-Champaign.

Palincsar, A. S. (1985). *Unpacking 'multicomponent, metacognitive training packages'.* Paper presented at the Annual meeting of the American Research Association, Chicago.

Palincsar, A. S. (1986). *Interactive cognition to promote listening comprehension.* Paper presented at the Annual meeting of the American Educational Research Association, San Francisco.

Palincsar, A. S. (1987). *Collaborating for collaborative learning of text comprehension*. Paper presented at the Annual Meeting of the American Educational Research Association, Washington, DC.

Palincsar, A. S. (2006). *Reciprocal teaching 1982 to 2006: The role of research, theory and representation in the transformation of instructional research. Oscar Causey Address.* Paper presented at the National Reading Conference, 2006, Los Angeles.

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

Palincsar, A. S., & Brown, A. L. (1986). Interactive teaching to promote independent learning from text. *The Reading Teacher, 39*, 771 - 777.

Palincsar, A. S., & Brown, A. L. (1989). Classroom dialogues to promote self-regulated comprehension. In J. Brophy (Ed.), *Advances in research on teaching* (Vol. 1, pp. 35 - 72). Greenwich, CT: JAI Press.

Palincsar, A. S., Brown, A. L., & Martin, S. (1987). Peer interaction in reading comprehension instruction. *Educational Psychologist, 22*(3), 231 - 253. doi: 10.1207/s15326985ep2203&4_3

Palincsar, A. S., David, Y. M., & Brown, A. L. (1989). *Using reciprocal teaching in the classroom: A guide for teachers*. Unpublished manual.

Palincsar, A. S., David, Y. M., Winn, J., Stevens, D., & Brown, A. L. (1990). *Examining the differential effects of teacher- versus student-controlled activity in comprehension instruction*. Paper presented at the Annual meeting of the American Research Association, Boston.

Palincsar, A. S., & Klenk, L. (1992). Fostering literacy learning in supportive contexts. *Journal of Learning Disabilities, 25*(4), 211 - 229. doi: 10.1177/002221949202500402

Palincsar, A. S., & Schutz, K. (2011). Reconnecting strategy instruction with its theoretical roots. *Theory Into Practice, 50*, 85 - 92. doi: 10.1080/00405841.2011.558432

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. doi: doi: 10.1037/0022-0663.76.6.1239

Paris, S. G., Lipson, M. Y., & Wixson, K. (1983). Becoming a strategic reader. *Contemporary Educational Psychology, 8*, 293 - 316. doi: 10.1016/0361-476X(83)90018-8

Parker, M., & Hurry, J. (2007). Teachers' use of questioning and modelling comprehension skills in primary classrooms. *Educational Review, 59*(3), 299 - 314. doi: 10.1080/00131910701427298

Pashler, H., Rohrer, D., Cepeda, N. J., & Carpenter, S. K. (2007). Enhancing learning and retarding forgetting: Choices and consequences. *Psychonomic Bulletin and Review, 14*, 187 - 193. doi: 10.3758/BF03194050

Paivio, A. (1986). *Mental representations: A dual coding approach*. New York: Oxford University Press.

Pearson, P. (2011). Toward the next generation of comprehension instruction: A coda. In H. Daniels (Ed.), *Comprehension going forward* (pp. 243 - 253). Portsmouth, NH: Heinemann.

Pearson, P. D., & Fielding, L. (1991). Handbook of reading research. In R. Barr, M. L. Kamil, P. B. Mosenthal & P. D. Pearson (Eds.), *Comprehension instruction.* (Vol. 2, pp. 815 - 860). White Plains, NY: Longman.

Perry, N., & Winne, P. H. (2006). Learning from learning kits: Study traces of students' self-regulated engagements with computerised content. *Educational Psychology Review, 18*(3), 233 - 237. doi: 10.1007/s10648-006-9014-3

Pilonieta, P., & Medina, A. L. (2009). Reciprocal teaching for the primary grades: "We can do it, too!". *Reading Teacher, 63*(2), 120-129. doi: 10.1598/RT.63.2.3

Pintrich, P. R. (2000). The role of goal orientation in learning and achievement. *Journal of Educational Psychology, 92*, 544 - 555. doi: 10.1037/0022-0663.92.3.544

Ponce, H. R., & Mayer, R. E. (2014). Qualitatively different cognitive processing during online reading primed by different study activities. *Computers in Human Behavior, 30*(1), 121 - 130. doi: 10.1016/j.chb.2013.07.054

Pressley, M. (1976). Mental imagery helps eight-year-olds remember what they read. *Journal of Educational Psychology, 68*, 355 -359. doi: 10.1037/0022-0663.68.3.355

Pressley, M. (2000). What should comprehension instruction be the instruction of? In M. L. Kamil, P. B. Mosenthal, P. D. Pearson & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 545 - 561). Mahwah, NJ: Earlbaum.

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

Pressley, M., Graham, S., & Harris, K. (2006). The state of educational intervention research as viewed through the lens of literacy instruction. *British Journal of Educational Psychology, 76*(1), 1 - 19. doi: 10.1348/000709905X66035

Pressley, M., & Hilden, K. (2004). Verbal protocols of reading. In N. K. Duke & M. H. Mallette (Eds.), *Literacy research methodologies* (pp. 308 - 321). New York: Guildford.

Pressley, M., Wharton-McDonald, R., Mistretta-Hampton, J., & Echevarria, M. (1998). The nature of literacy instruction in ten fourth-grade classrooms in upstate New York. *Scientific Studies of Reading, 2*(2), 159 - 194. doi: 10.1207/s1532799xssr0202_4

Protopapas, A., Sideridis, G. D., Mouzaki, A., & Simos, P. G. (2011). Matthew effects in reading comprehension. *Journal of Learning Disabilities, 44*(5), 402-420. doi: 10.1177/0022219411417568

Puntambekar, S., & Hubscher, R. (2005). Tools for scaffolding students in a complex learning environment: What have we gained and what have we missed? [Article]. *Educational Psychologist, 40*(1), 1-12. doi: 10.1207/s15326985ep4001_1

Purcell-Gates, V., Duke, N. K., & Martineau, J. A. (2007). Learning to read and write genre-specific text: Roles of authentic experience and explicit teaching. *Reading Research Quarterly, 42*, 8 - 45. doi: 10.1086/461599

QSR International Pty Ltd. (2010). NVivo qualitative data analysis software (Version 9).

Rapp, D. N., Van den Broek, P., McMaster, K., Kendeou, P., & Espin, C. A. (2007). Higher-order comprehension processes in struggling readers: A perspective for research and intervention. *Scientific Studies of Reading, 11*(4), 289 - 312. doi: 10.1080/10888430701530417

Ricketts, J. (n.d.). Reading: Assessment and research. Retrieved September 26, 2013 http://www.yarcsupport.co.uk/documents/reading-research.pdf

Rimrodt, S., Lightman, A., Roberts, L., Denckla, M. B., & Cutting, L. E. (2005). *Are all tests of reading comprehension the same?* Paper presented at the Annual meeting of the International Neuropsychological Society, St. Louis, MO.

Roberts, K., & Duke, N. K. (2010). Comprehension in the elementary grades: The research base. In K. Ganske & D. Fisher (Eds.), *Comprehension across the curriculum: Perspectives and practices K-12* (pp. 23 - 45). New York: The Guildford Press.

Rogoff, B., Gauvain, M., & Gardner, W. (1987). The development of children's skills in adjusting plans to circumstances. In S. Friedman, E. Skolnick & R. Cocking (Eds.), *Blueprints for thinking: The role of planning in cognitive development* (pp. 303 - 320). Cambridge, England: Cambridge University Press.

Rojas-Drummond, R., & Mercer, N. (2003). Scaffolding the development of effective collaboration and learning. *International Journal of Educational Research, 39*, 99 - 111. doi: 10.1016/S0883-0355(03)00075-2

Rose, J. (2006). *Independent review of the teaching of early reading*. London: DfES.

Rose, J. (2013). A review of 'Developing reading comprehension'. Retrieved October 25, 2013 http://eu.wiley.com/WileyCDA/WileyTitle/productCd-1118606760.html

Rosenshine, B., & Meister, C. (1994). Reciprocal teaching: a review of the research. *Review of Educational Research, 64*(4), 479-530. doi: 10.2307/1170585

Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions; a review of the intervention studies. *Review of Educational Research, 66*(2), 181 - 221. doi: 10.3102/00346543066002181

Rush, R. T., & Milburn, J. L. (1988). *The effects of reciprocal teaching on self-regulation of reading comprehension in a postsecondary Technical School program.* Paper presented at the 38th Annual Meeting of the National Reading Conference, Tucson, AZ.

Ryan, M. (2006). *Scratch and Sniff*. London: A and C Publishers Ltd.

Sadoski, M. (1983). An explanatory study of the relationship between reported imagery and the comprehension and recall of a story. *Reading Research Quarterly, 19*, 110 - 123. doi: 10.2307/747341

Scammacca, N., Roberts, G., Vaughn. S., Edmonds, M., Wexler, J., R., C. K.,, et al. (2007). *Interventions for adolescent struggling readers: A meta-analysis with implications for practice.* Portsmouth, NH: RMC Research Corporation, Center on Instruction.

Schellings, G., Aarnoutse, C., & van Leeuwe, J. (2006). Third-grader's think-aloud protocols: Types of reading activities in reading an expository text. *Learning And Instruction, 16*(6), 549-568. doi: DOI: 10.1016/j.learninstruc.2006.10.004

Schneider, W., & Schlagmüller, M. (2002). *Metakognitives wissen über textverarbeitung*: Universität Würzburg: Lehrstuhl für Psychologie IV.

Schünemann, N., Spörer, N., & Brunstein, J. C. (2013). Integrating self-regulation in whole class reciprocal teaching: A moderator-mediator analysis of incremental effects on fifth graders' reading comprehension. *Contemporary Educational Psychology, 38*(4), 228 - 305. doi: 10.1016/j.cedpsych.2013.06.002

Schunk, D. (2001). Social cognitive theory and self-regulated learning. In B. Zimmerman & D. Schunk (Eds.), *Self-regulated learning and academic achievement. Theoretical perspectives* (2nd ed., pp. 125 -151). Mahwah, NJ: Lawrence Erlbaum Associates.

Schunk, D., & Zimmerman, B. (1997). Social origins of self-regulatory competence. *Educational Psychologist, 32*, 195 - 208. doi: 10.1207/s15326985ep3204_1

Scott, S. (2009). *Knowledge for teaching reading comprehension: Mapping the terrain.* Doctoral dissertation, University of Michigan, Ann Arbor. Retrieved June 28th, 2014, from http://deepblue.lib.umich.edu/bitstream/2027.42/62201/1/sarascot_1.pdf

Seymour, J. R., & Osana, H. P. (2003). Reciprocal teaching procedures and principles: Two teachers' developing understanding. *Teaching and Teacher Education, 19*(3), 325. doi: 10.1016/S0742-051X(03)00018-0

Shanahan, T., Callison, K., Carriere, C., Duke, N. K., Pearson, P., Schatschneider, C., et al. (2010). Improving reading comprehension in kindergarten through 3rd grade: A practice guide (NCEE 2010-4038) Retrieved June 28th, 2014, from http://ies.ed.gov/ncee/wwc/PracticeGuide.aspx?sid=14

Shortland-Jones, B. (1986). *The development and testing of an instructional strategy for improving reading comprehension based on schema and metacognitive theories.* Unpublished doctoral dissertation, University of Oregon.

Shriffin, 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. doi: 10.1037/0033-295X.84.2.127

Siegler, R. S. (2006). Microgenetic analyses of learning. In D. Kuhn & R. S. Siegler (Eds.), *Handbook of child psychology: Vol. 2. Cognition, perception , and language* (6th ed., pp. 464 - 510). Hoboken, NJ: John Wiley and Sons.

Slavin, R. E. (1996). Research for the future. Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology, 21*(1), 43 - 69.

Slavin, R. E. (2008). What works? Issues in synthesizing education program evaluations. *Educational Researcher, 37*(1), 5 - 11. doi: 10.3102/0013189X08314117

Slavin, R. E., Lake, C., Chambers, B., Cheung, A., & Davis, S. (2009). Effective reading programs for the elementary grades: A best-evidence synthesis. *Review Of Educational Research, 79*(4), 1391-1466. doi: 10.3102/0034654309341374

Smith, F., Hardman, F., Wall, K., & Mroz, M. (2004). Interactive whole class thinking in the National Literacy and Numeracy Strategies. *British Education Research Journal, 30*(3), 395 - 411. doi: 10.1080/01411920410001689706

Smith, N. B. (1965). *American reading instruction*. Newark, DE: International Reading Association.

Snow, C. E. (Ed.). (2002). *Reading for understanding: Toward an R and D program in reading comprehension*. Santa Monica, CA: RAND.

Snowling, M., Stothard, S., Clarke, P. J., Bowyer-Crane, C., Harrington, A., Truelove, E., et al. (2011). *York Assessment of Reading for Comprehension: Second edition*. London: GL Assessment.

Snowling, M., Stothard, S., Clarke, P. J., Nation, K., Truelove, E., & Hulme, C. (2009). *York Assessment of Reading for Comprehension*. London: GL Assessment

Souvignier, E., & Trenk-Hinterberger, I. (2010). Implementation of a program to foster reading competence: Improving long-term effects using booster-sessions. *Zeitschrift fur Padagogische Psychologie, 24*, 207 - 220. doi: 10.1024/1010-0652/a000017

Spache, G. (1972). *Diagnostic Reading Scales*. San Francisco: McGraw-Hill.

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

Stahl, S., & Fairbanks, M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research, 56*, 72 - 100. doi: 10.3102/00346543056001072

Stanovich, K. E. (1986). Matthew effects in reading; some consequences of individual differences in the acquisition of literacy *Reading Research Quarterly, 21*, 360 - 407. doi: 10.1598/RRQ.21.4.1

Stanovich, K. E. (1991). Discrepancy definitions of reading disability: Has intelligence lead us astray? *Reading Research Quarterly, 26*, 7 - 29. doi: 10.2307/747729

Stone, C. A. (1998a). The metaphor of scaffolding: Its utility for the field of learning disabilities. *Journal of Learning Disabilities, 31*, 344 - 364. doi: 10.1177/002221949803100404

Stone, C. A. (1998b). Should We Salvage the Scaffolding Metaphor? *Journal of Learning Disabilities, 31*(4), 409-413. doi: 10.1177/002221949803100411

Stothard, S., & Hulme, C., G. (1991). A note of caution concerning the Neale Analysis of Reading Ability (revised). *British Journal Of Educational Psychology, 61*(2), 226 - 229.

Stuart, J. (1979). A ribbon for Baldy *Adventures for readers. Book One* (pp. 60). New York: Harcourt Brace Jovanovich.

Stuart, M., Stainthorp, R., & Snowling, M. (2008). Literacy as a complex activity: Deconstructing the simple view of reading. *Literacy, 42*(2), 59. doi: 10.1111/j.1741-4369.2008.00490.x

Swanson, E. (2008). Observing reading instruction for students with learning disabilities: A synthesis. *Learning Disability Quarterly, 31*(115 - 131).

Taboada, A., & Guthrie, J. T. (2006). Contributions of student questioning and prior knowledge to construction of knowledge from reading information text. *Journal of Literacy Research, 38*, 1 - 35. doi: 10.1207/s15548430jlr3801_1

Takala, M. (2006). The effects of reciprocal teaching on reading comprehension in mainstream and special education. *Scandinavian Journal of Educational Research, 50*(5), 559 - 576. doi: 10.1080/00313830600953824

Tanner, E., Brown, A. L., Day, N., Kotecha, M., Low, N., Morrell, G., et al. (2011). Evaluation of Every Child a Reader (ECaR) Research Report DFE-RR114: Department for Education.

Taylor, B. M., & Frye, B. J. (1992). Comprehension strategy instruction in the intermediate grades. *Reading Research and Instruction, 32*(1), 39 - 48. doi: 10.1080/19388079209558104

Taylor, B. M., Peterson, D. S., & Pearson, P. (2002). Looking inside classrooms: Reflecting on the "how" as well as the "what" in effective reading instruction. *The Reading Teacher, 56*(3), 270 -279. doi: 10.1598/RT.56.3.5

Tennent, W. (2011). *Reading comprehension and its acquisition at Key Stage 2; Intervention and assessment*. Unpublished doctoral thesis. Institute of Education, London.

Tomeson, M., & Aarnoutse, C. (1998). Effects of an instructional programme for deriving word meaning. . *Educational Studies, 24*(1), 107 - 122. doi: 10.1080/0305569980240108

Tomlinson, J. (2002). *The owl who was afraid of the dark*. London: Egmont Books.

Torgesen, J., Wagner, R. K., & Rashotte, C. A. (1999). *Test of word reading efficiency (TOWRE)*. Austin, Texas: Pro-Ed.

Trabasso, T., & Bouchard, E. (2002). Teaching readers how to comprehend text strategically. In C. Block & M. Pressley (Eds.), *Comprehension instruction: Research-based best practices* (pp. 176 - 200). New York: Guildford.

Trabasso, T., & Magliano, J. P. (1996). Conscious understanding during comprehension. *Discourse Processes, 21*, 255 - 287. doi: 10.1080/01638539609544959

Trease, G. (2002). *Cue for treason*. London: Penguin Books Ltd.

Van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher-student instruction: A decade of research. *Educational Psychology Review, 22*, 271 - 296. doi: 10.1007/s10648-010-9127-6

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic.

Van Hout-Wolters, B. (2009). Leerstrategieën meten: soorten meetmethoden en hun bruikbaarheid in onderwijs en onderzoek. (Measuring learning strategies. Different kinds of assessment methods and their usefulness in education and research). *Pedagogische Studiën, 86*(2), 103 -110.

Van Keer, H. (2004). Fostering reading comprehension in fifth grade by explicit instruction in reading strategies and peer tutoring. *British Journal of Educational Psychology, 74*, 37 - 70. doi: 10.1348/000709904322848815

Vaughn, S., Chard, D. J., Bryant, D. P., Coleman, M., Tyler, B., & Thompson, S. (2000). Fluency and comprehension interventions for third-grade students: Two paths to improved fluency. *Remedial & Special Education, 21*, 325 - 335. doi: 10.1177/074193250002100602

Vaughn, S., Hughes, M. T., Schumm, J., & Klingner, J. K. (1998). A collaborative effort to enhance reading and writing instruction in inclusion classrooms. *Learning Disability Quarterly, 21*, 55 - 74.

Vaughn, S., Klinger, 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 and Special Education, 22*(2).

Veenman, M. V. J., Prins, F., & Verheij, J. (2003). Learning styles; Self-reports versus thinking-aloud measures. *British Journal of Educational Psychology, 73*, 357 - 372. doi: 10.1348/000709903322275885

Veenman, M. V. J., & Spaans, M. A. (2005). Relation between intellectual and metacognitive skills: Age and task differences. *Learning and Individual Differences, 15*(2), 159 -176. doi: 10.1016/j.lindif.2004.12.001

Vygotsky, L. (1978). *Mind in society*. Cambridge: Harvard University Press.

Webb, N., & Palincsar, A. S. (1996). Small group processes in the classroom. In D. Berliner & R. Calfee (Eds.), *Handbook of educational psychology* (pp. 841 - 873). New York: Simon and Schuster.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: The Psychological Corp.

Westera, J., & Moore, D. W. (1995). Reciprocal teaching of reading comprehension in a New Zealand high school. *Psychology in the Schools, 32*(3), 225 - 232. doi: 10.1002/1520-6807(199507)32:3<225::AID-PITS2310320310>3.0.CO;2-F

What Works Clearinghouse. (2008). Procedures and standards handbook (version 2.0) Retrieved July 20th 2010, from http://ies.ed.gov/ncee/wwc/references/idocviewer/doc.aspx?docid=19&tocid=1

What Works Clearinghouse. (2010). Intervention: Reciprocal Teaching Retrieved July 20th 2010, from http://ies.ed.gov/ncee/wwc/pdf/intervention_reports/wwc_rec_teach_091410.pdf

Whatmuff, A. (n.d.). Inference training Retrieved June 5th, 2014, from http://removingbarriers.lls.leicester.gov.uk/Media/Removing%20Barriers/LLS%20web%2003-14%20V2.pdf

Wheldall, K., Colmar, S., Wenban-Smith, J., Morgan, A., & Quance, B. (1992). Teacher-child oral reading interactions: how do teachers typically tutor? *Educational Psychology, 12*, 177 - 194.

Whitehead, D. (2002). 'The story means more to me now': Teaching thinking through Guided Reading. *Reading, Literacy and Language, 36*(1), 33 - 37.

Wigfield, A., Eccles, J. S., Schiefele, U., Roeser, R., & Davis-Kean, P. (2006). Development of achievement motivation. In N. Eisenberg, W. Damon & R. M. Lerner (Eds.), *Handbook of Child Psychology. Vol 3. Social, emotional and personality development* (pp. 933 - 1002). Hoboken, NJ: Wiley.

Wiig, E., & Secord, W. (1992). *Test of Word Knowledge (TOWK)*. USA: The Psychological Corporation.

Williams, J. P., Hall, K. M., & Lauer, K. D. (2005). Expository text comprehension in the primary grade classroom. *Journal of Educational Psychology, 97*, 538 - 550.

Williamson, R. A. (1989). *The effect of reciprocal teaching on student performance gains in third grade basal reading instruction*. Unpublished doctoral dissertation. Texas A and M University.

Winograd, P., & Johnston, P. (1982). Comprehension monitoring and the error detection paradigm. *Journal of Literacy Research, 14*(1), 61-76. doi: 10.1080/10862968209547435

Wood, D., & Middleton, D. (1975). A study of assisted problem-solving. *British Journal of Psychology, 66*, 181 - 191. doi: 10.1111/j.2044-8295.1975.tb01454.x

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

Yuill, N., & Jocelyne, T. (1988). Effect of organizational cues and strategies on good and poor comprehenders' story understanding. *Journal of Educational Psychology, 80*(2), 152 - 158. doi: 10.1037/0022-0663.80.2.152

Yuill, N., & Oakhill, J. (1988). Effects of inference awareness training on poor reading comprehension. *Applied Cognitive Psychology, 2*(1), 33 - 45.

Zimmerman, B. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich & M. Zeider (Eds.), *Handbook of self-regulation* (pp. 13 - 39). San Diego, CA: Academic Press.

Zimmerman, B. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice, 41*, 64 - 70. doi: 10.1207/s15430421tip4102_2

Zimmerman, B., & Kitsantas, A. (1997). Developmental phases in self-regulation. Shifting from process goals to outcome goals. *Journal of Educational Psychology, 89*, 29 - 36. doi: 10.1037/0022-0663.89.1.29

Zimmerman, B., & Schunk, D. (Eds.). (2011). *Handbook of self-regulation of learning and performance*. New York, NY and Abingdon, Oxon: Routledge

Appendices

Appendix A: Collaborative Strategic Reading compared to Reciprocal Teaching.

CSR uses four strategies – 'Preview', 'Click and clunk', 'Get the gist' and 'Wrap up'. The way that these relate to the four strategies of predicting, clarifying, summarising and questioning, can be seen below (Table from J. K Klingner, Vaughn, Dimino, Schumm, & Bryant, 2001).

Reciprocal teaching	Collaborative Strategic Reading
Designed for use with narrative as well as expository text.	Designed primarily for use with expository text.
No brainstorming before reading.	Students **brainstorm** to activate prior knowledge as part of **preview** (before reading).
Students **predict** what they think will happen next before reading each paragraph or segment of text.	Students only **predict** as part of the Preview strategy (before reading), making informed hunches about what they think they will learn.
Students **clarify** words or chunks of text they don't understand by rereading the sentences before and after the sentence they don't understand, and/or asking a peer for assistance.	Students use 'fix-up strategies' to clarify '**clunks**' (words they don't understand): • Reread the sentence… • Reread the sentences before & after… • Break apart the work and look for smaller words you know • Look for a prefix or suffix you know • (Look at the picture . . .) • (Ask for help)
Students **summarize** the paragraph or segment of text they have just read.	Students **get the gist** of the paragraph or segment of text they have just read, identifying 'the most important who or

273

	what' and the most important thing about the who or what. They say the gist in 10 words or less.
Students **generate questions** after each paragraph or segment of text they have just read.	Students only **generate questions** as part of a **wrap up** after they have read the entire day's selection. Students answer each other's questions.
No **review** after reading.	Students **review** what they have learned after reading the day's selection.
8-12 students in the group, with the teacher in the group.	An entire class is divided into **cooperative groups** of 2-5; the teacher circulates rather than staying with a group.
No learning logs.	Students each record their previews, clunks, questions, and what they've learned in individual **CSR Learning Logs**.
The 'leader' (a student) facilitates the discussion about a paragraph or section of text; this role rotates after each paragraph.	Every student in the group has a meaningful role; one of these roles is to be the 'leader.' Roles are assigned for an entire lesson (only rotating biweekly in some classes).
No cue cards.	Students use **Cue Cards** to help them implement their roles and the comprehension strategies.

The major difference in strategy instruction is the more prominent role given to background knowledge in the predicting phase. Rather than just using prior knowledge to think about what the author might say next, the emphasis is on making associations with what children already know, and learning from their peers already know too. The clarifying phase encourages not only to look for what they do not understand (the clunks), but to think about what enhances their understanding (the clicks). This encourages the self-monitoring process. The 'Get the gist' phase equates to summarising. Children are taught to identify the most

important points in the text and to paraphrase them. The only difference is that a limit of 10 words is set, to ensure that extraneous details are omitted.

Finally, 'Wrap up' is the questioning phase, where children identify the questions a good teacher would ask about what they have just read. Children are taught that there are different levels of questions –

1. 'right there' questions, which have answers contained in one sentence
2. 'think and search' questions which require the integration of several events or facts from different sections of the passage
3. 'author and you' questions, requiring inference.

The major difference with the questioning procedure in CSR is that it takes place at the end of the text rather than after very paragraph in Reciprocal Teaching. That positioning and the use of question levels does encourage integration of idea across the text as a whole, and may deter the lists of questions about easily accessible facts which children find easier to produce that the types of questions which require more integration and inference.

CSR does utilise heterogeneous collaborative groups as in Reciprocal Teaching, but children are given specific roles within each group, and they retain that role for the entire lesson. As has already been discussed in the Introduction, this taking of roles has been criticised, and it is necessary to ensure that each role has a meaningful part to play in each discussion by choosing text carefully.

In an article discussing several years of research (Vaughn, Klinger, & Bryant, 2001)we learn that CSR began with two researchers implementing Reciprocal Teaching with Spanish speaking middle school students with learning difficulties. Klingner and Vaughan (1996) achieved statistically significant gains on standardised reading tests with small groups of 7th and 8th Grade students (aged 12-14). When these student said they should have been taught the strategies earlier, the researchers moved on to fourth-Grade heterogeneous classrooms (children aged 9 to 10) with the intention of showing that younger children could benefit, and that it was possible to implement the intervention in student-led groups within an intact class, rather than the teacher-led small groups of adolescents which had characterised much of the Reciprocal Teaching research (Klingner, et al., 1998).This second study showed CSR students out performing a control group on a standardised reading test and a test of content knowledge. However, an analysis of the discourse showed too many of the traditional initiation-response-evaluation patterns and little engagement in higher level

discussions, which lead to the introduction of the different levels of questions which were discussed above. Klingner and Vaughan also discovered that 25% of the student's utterances were concerned with procedural discussion about implementation and role-taking; this lead then to adapt Reciprocal Teaching further so that students were pre-assigned a fixed role for the day.

After the researchers themselves had implemented and refined their approach, they implemented CSR on a wider scale via a yearlong collaborative professional development programme (Vaughn, Hughes, Schumm, & Klingner, 1998) and the followed the teachers for 3 years to study their implementation (Klingner, Vaughn, Hughes, & Arguelles, 1999). Modifications which came about as a result of this research included the introduction of learning logs and the refinement of cue cards.

A second professional development study took place in 2000 (Bryant et al., 2000) when CSR was one of three interventions carried in 10 6th Grade (aged 11 to 12) classes. Again significant gains were made in comprehension by the CSR group, and lessons were learned about the difficulties involved in implementation. In particular, poor decoders found the content text very challenging. The authors acknowledge that readers with severe difficulties in decoding at the age of 11 to 12 will need more targeted support in terms of special reading education to improve their fluency before CSR can be effective. CSR, like Reciprocal Teaching, is not aimed at the basic skills reading requires, but at the higher level skills.

After a decade of designing, implementing, refining and evaluating CSR the researchers conducted a randomized controlled trial to determine its efficacy. This large scale project involved 782 7th and 8th Grade students (aged 12 to 14) in six schools. Seventeen teachers taught CSR and 'business as normal' to their randomly assigned English/language arts classes. Fidelity measures were used to ensure adherence to CSR and that no crossover took place, and the students were given standardised tests of word reading and reading comprehension. Multilevel modelling suggested that the CSR outperformed the control students on the comprehension test (Gates-MacGinitie), when the effects of clustering and pre-treatment differences were explicitly modelled. Thus, CSR has been shown to be effective in whole class situations. The modifications which have taken place have made Reciprocal Teaching applicable to whole class situations, although the age groups have largely been older than the target group of this thesis. The question remains as to why CSR has not been recommended in this country? It may be that the labels used for

the strategies are seen as American English. However, the refinement of the strategies, the use of cue cards and the taking of roles, may be adaptations needed in whole class teaching, and the CSR research has shown that such an approach is effective in that situation, although the target age group has been consistently older than the age group of interest in the present research.

Institute of Education

University of London
20 Bedford Way
London
WC1H OAL

Date

Dear Parent/s,

This letter is to introduce myself and to explain about some research which I will be carrying out in Year 5.

My name is Frances Hampson-Jones. I have been connected with ……………. School as a parent and a teacher, from1998 to the present. I am currently working part-time in the Special Needs Department, whilst studying for a University of London PhD at the Institute of Education. I am undertaking a research project concerning interventions to improve reading comprehension.

I am writing to ask permission for your son/daughter to take part in this research. This will only involve the kinds of activities which children already do in school as part of the English curriculum.

All three Year 5 classes will be observed in a number of lessons, and all children will be assessed using standardised tests and researcher designed tests. After the assessments have been completed, I will be teaching two classes for an hour a week in place of their normal literacy lesson. This teaching phase will take place between February and May of 2008. After the teaching phase, all children will be reassessed. These assessments will be very useful to the Year 5 teachers in their future planning, and in informing the Year 6 teachers about progress.

 Research undertaken by students at the Institute is subject to regulation by an Ethics Committee and is conducted using guidelines from the British Psychological Society (see www.bps.org.uk for further details).

Please be assured that all research is strictly confidential. No names will be kept on any computer system and names will not be used when reporting findings.
The results of the research will be made available to the school, and to anyone who is interested, after I have completed the pilot study.

Mrs ………….and the Year 5 teachers have given me their full support. I did carry out the research for my MSc at ……………….., in 2006, and the staff and children were enormously helpful.

I am very much looking forward to renewing my research connection with the……….. School.

If you do not wish your son/daughter to take part in this research please let Mrs ………..know by (date of letter + one week)

Yours faithfully,

Frances Hampson-Jones

Appendix C. The adapted form of the MARSI (Mokhtari & Reichard, 2002) used in Study 1.

Below is a list of some of the things people do when reading. I am interested in what you do when you read a book at school, so please be honest and tell me if you do these things.

There are no right or wrong answers!

Please circle a number next to each question.

1 means **'I never or almost never do this'**

2 means **'I do this only occasionally'**

3 means '**I sometimes do this' (about half of the time)**

4 means '**I usually do this'**

5 means '**I always or almost always do this'**

Example (not to do with reading!)

I brush my teeth before I go to bed ……………………………..	1	2	3	4	5

1. When text becomes difficult I read it aloud to help me understand.	1	2	3	4	5
2. I make notes to help me understand what I have read.	1	2	3	4	5
3. I count the number of verbs in a passage before I start reading.	1	2	3	4	5
4. I think about what I already know to help me understand what I read.	1	2	3	4	5
5. I look over the text quickly to see what it is about before I start reading.	1	2	3	4	5

6. When text becomes difficult I reread to help me understand.	1	2	3	4	5
7. Before I begin reading I think about the meanings of words with more than one meaning.	1	2	3	4	5
8. I summarise what I have read to help me make sense of what I am reading.	1	2	3	4	5
9. I read slowly but carefully to make sure I understand what I'm reading	1	2	3	4	5
10. I discuss what I have read with other people to check my understanding	1	2	3	4	5
11. I try to get back on track when I lose concentration	1	2	3	4	5
12. I underline or circle information to help me remember important points.	1	2	3	4	5
13. While I'm reading I count how many words I know already.	1	2	3	4	5
14. I change how fast I read according to what I am reading.	1	2	3	4	5
15. I decide what to read closely and what to ignore.	1	2	3	4	5
16. I use a dictionary to help me understand what I read.	1	2	3	4	5
17. When I've finished reading I count how many words I could read without a mistake.	1	2	3	4	5
18. I look at the pictures to help me understand what is going on.	1	2	3	4	5
19. I stop from time to time to think about what I am reading.	1	2	3	4	5
20. I think about the ideas in my own words to better understand what I have read.	1	2	3	4	5
21. I try to make pictures in my head of what is happening.	1	2	3	4	5
22. I have a purpose in mind when I read a book at school.	1	2	3	4	5

23. I look at headings to see what is important.	1	2	3	4	5
24. If I don't understand something I read the sentence again.	1	2	3	4	5
25. I ask myself questions about what I would like to find out in the text.	1	2	3	4	5
26. I try to guess the meanings of unknown words or phrases.	1	2	3	4	5
27. I change the ending so that it makes sense.	1	2	3	4	5
28. I check to see if the pictures have anything missing.	1	2	3	4	5
29. I try to predict what is going to happen from the title.	1	2	3	4	5
30. I skim through the text first to look at how long it is and how it is organised.	1	2	3	4	5

Appendix D. Post-instruction questionnaire for Study 1 and Study 2.

Thinking about the group reading lessons, please circle the number which sums up your opinion the best:

Scales for how much you have learned

1. I learned a lot of new things.
2. I learned quite a few new things.
3. I learned some new things.
4. I learned a few new things.
5. I didn't learn anything new.

Scales for how much you enjoyed the lessons

1. I really enjoyed these lessons
2. I enjoyed these lessons quite a lot.
3. I enjoyed these lessons a bit.
4. I didn't really enjoy these lessons much.
5. I didn't enjoy these lessons at all.

Please tell me what you have enjoyed about these lessons:

Please tell me if there is anything you haven't enjoyed about these lessons? Be honest. I won't mind!

Do you think these lessons have helped you to improve your understanding of what you read?

Please tell me any other comments you would like to make:

Appendix E. The script for the presentation made by the Reciprocal Teaching groups in Study 1.

Four children, C., CH, T. and L. read a page from 'Cue for Treason' and then discuss the text:

C. (the 'teacher') We need to make sure we understand what's happening, so can we clarify first? Who's Clara?

L. I found two words to clarify: 'halberdiers' and 'sovereign' (point out the words in the text first and then clarify).

C. I think you worked those words out really well. Is there anything else anyone needs to clarify?

CH. I'm not sure about these lines:

'I understand that you and this girl have been at some pains to prolong my wearisome existence'.

C. Can anyone help?

T. I think 'you and this girl' must mean Peter and Kit, and the queen must be talking to them a bout how they saved her and that must be why she wanted to see Peter. The queen wouldn't be talking to just a boy otherwise.

L. And Peter and Kit have saved her life. 'Prolong' has the word 'long' in it, so I think it means to make something longer. And 'wearisome' is like the word 'weary' which means tired. We know the queen is old as it says she has brown knobbly fingers. Maybe she is tired of life.

C. Now that's clear, we need to see if we can make up some questions about the text. Who's Dennis?

CH. That's me. My question about the main point is, 'Who is the girl in the flame coloured dress, and how does Peter recognise her?'

L. The girl's Kit.

C. How do you know that?

L. (read where it says in the text). I think I know how he recognised her (explain your point about knowing your Mum from warm feelings. Peter and Kit have been through a lot together).

C. That was a good question as it made us go beneath the surface and work things out. It wasn't an easy question, like 'what colour was her dress'. Do you have any more tricky questions CH?

C. We could ask what Peter is feeling?

T. He must be excited to meet the queen and maybe a bit nervous too.

C. Can you show me any evidence for that in the text?

T. He says his knees are weak, and he combed his hair before he went in. And I can imagine it would be very scary to meet the queen. I would be nervous doing that.

L. He must be excited too, as he knows he has saved the queen's life, even though he is only a boy.

C. Could we have a summary of what's happened on this page? Who's Susie?

T. I'm Susie and this is my summary (read it please).

C. I think that's a very good summary. You used your own words and told us what happened. But do you think there are too many details? Can anyone think of a detail that could be left out, without losing the main points?

L. Do we need to know that Peter neatened his hair, as you said he was nervous?

T. No, that's a good point. Maybe I could leave that bit out. And I didn't need to say the girl was wearing a flame-coloured dress. That's a detail too.

C. I'm Mystic Mike, and this is my prediction (read it). What do you think?

L. I think that's a good prediction as in other books I've read something similar (explain about only wanting to go home after a great adventure).

C. Shall we read on to see if I'm right? Whose turn is it to be the teacher now?

Institute of Education

University of London

20 Bedford Way

London

WC1H OAL

19/06/09

Dear Parent/s,

This letter is to introduce myself and to explain about some research which I will be carrying out in Year 3.

My name is Frances Hampson-Jones. I have been connected with …………..School as a parent and a teacher, from1998 to 2008. I am currently studying for a University of London PhD at the Institute of Education. I am undertaking a research project concerning interventions to improve reading comprehension.

I am writing to ask permission for your son/daughter to take part in this research. This will only involve the kinds of activities which children already do in school as part of the English curriculum.

Following a successful pilot study in Year 5 last year, I am interested in observing how a class teacher can improve reading comprehension, and how the children respond to that teaching. Your child will continue to be taught by their normal class teacher, but I will be involved in observing the class, and talking to the children about what they have learned. I would like to begin with some interviews and assessments of reading ability. The teaching phase will continue for two terms, after which, all children will be reassessed. These assessments will be very useful to the class teacher and to your child's teacher/s in the following academic year.

The interviews and some observations will include the use of a tape recorder. Your child will not be identified on any tape recording. Please be assured that all research is strictly confidential. No names will be kept on any computer system and names will not be used when reporting findings.

Research undertaken by students at the Institute is subject to regulation by an Ethics Committee and is conducted using guidelines from the British Psychological Society (see *www.bps.org.uk* for further details), of which I am a graduate member.

The findings of the research will be made available to the school, and to anyone who is interested, after I have completed the study. The findings (using anonymised data) will also be written up for submission to a peer reviewed journal.

You have the right to withdraw your child from this research project at any time.

Mrs…………….. and Mrs ………………… have given me their full support.
When I carried out the research for my MSc at …………. in 2006, and the pilot study for my PhD in 2008, I found the staff and children at the school to be enormously helpful.

I am very much looking forward to renewing my research connection with the …………..School.

If you do not wish your son/daughter to take part in this research please let Mrs ………… know by September 6th 2009.

Yours faithfully,

Frances Hampson-Jones

Appendix G. Strategy interview adapted from Major Point Interview for Readers, from Keene and Zimmerman (1997), for Study 2 and Study 3.

1. USES SCHEMA

When you read that passage did it remind you of anything you know about?

What did it remind you of?

(If response is no, ask ………..Did it remind you of any things that have happened to you.)

We have just talked about what this passage reminds you of (restate child's response).

How does it help a reader to understand a passage if they think about what they already know as they read?

2. INFERS, PREDICTS

In addition to what you have read so far, if we continued to read, what else do you think the author would like to tell you?

3. ASKS QUESTIONS

What questions did you have while you were reading?

What do you wonder about now that you've read this passage?

We have just discussed the questions you asked when you were reading. Do questions help you to understand what you have read? Do questions come into your head when you're reading?

4. DETERMINES WHAT IS IMPORTANT IN THE TEXT

Are there some parts of this passage that are more important than the others? Which ones? Why do you think they were the most important?

We have just talked about what you thought was important. How does thinking about the more important parts help you understand better?

Do you think about, or do anything, while you're reading to help you remember the important parts?

Do you ever have trouble remembering what is important after you have read?

How do you solve that problem?

5. MONITORS COMPREHENSION

What problems did you have when reading this passage?

Were your problems mainly in saying the words, or in understanding ideas?

How do you usually solve problems?

When you understand something really well, how do you know?

What would you tell another child about what a reader should try to understand each time he/she reads?

We have talked about problems you had when reading and how you solved them. What would you tell another child who didn't realise when what he/she was reading did not make sense to them?

6. CREATES MENTAL IMAGES

When you were reading this passage, did you make pictures in your mind?

Tell me everything you can about the pictures in your mind while you were reading just now.

What is in your picture that is not in the words in the passage?

Can you think of another book where you made pictures to help you understand the ideas? Tell me everything you can about that picture.

How do these pictures help you to understand more about what you read?

What would you tell another reader if he/she were trying to learn how to make pictures help them understand what they read?

7. SYNTHESISES

If you were to tell another person about the passage you have just read and you could only use a few sentences, what would you tell them?

When you are reading, do you ever think of other books/movies/people you know?

Does that help you understand the book you're reading?

Appendix H. Scoring rubric for MPIR (Keene and Zimmermann, 1997).

Uses schema	
1	No response/schematic connection
2	Can talk about what text reminds him or her of, but cannot explain; reference to schema may not be clearly connected to text
3	Relates background knowledge/experience to text
4	Expands interpretation of text using schema; may discuss schema related to author. Text structure; may pose questions based on apparent discrepancies between text and background knowledge
5	Explains how schema enriches interpretation of text; talks about use of schema to enhance interpretation and comprehension of other texts; connections extend beyond life experience and immediate text.
Infers	
1	No response/inference
2	Attempts a prediction or conclusion, inaccurate or unsubstantiated with text information
3	Draws conclusions or makes predictions that are consistent with text or schema
4	Draws conclusions and/or makes predictions and can explain the source of the conclusion or prediction
5	Develops predictions, interpretations, and/or conclusions that include connections between the text and the reader's background knowledge or ideas and beliefs
Questions	
1	No questions/irrelevant questions
2	Poses literal question/s
3	Poses questions to clarify meaning
4	Poses questions to enhance meaning of text (critical response/big idea), may explain how posing questions deepens comprehension
5	Uses questions to challenge the validity of print/author's stance, motive or point of view
Determines what is important in text	
1	No response, random guessing, inaccurate attempts to identify important elements
2	Identifies some elements (primarily pictures) as more important to text meaning

3	Identifies words/characters, and/or events as more important to overall meaning – makes some attempt to explain reasoning
4	Identifies at least one key concept, idea or theme as important in overall text meaning, and clearly explains why
5	Identifies multiple ideas or themes , may attribute them to different points of view, discuss author's stance or purpose and its relation to key themes and ideas in the text

Monitors comprehension

1	Little or no conscious awareness of reading process
2	Identifies difficulties- problems are often at word level; little or no sense of the need to solve the problem; does not articulate strengths; identifies need to concentrate; says sound it out
3	Identifies problems at word, sentence or schema level; can articulate and use a strategy to solve problems, usually at the word or sentence level
4	Articulates and uses more than one strategy for solving problems; focuses on problems at the schema (more global) level
5	Identifies problems at all levels; uses a variety of strategies flexibly and appropriately given the context and the problem

Visualises

1	No response
2	Describes some visual or other sensory images: may be tied directly to text or a description of the picture in the text
3	Describes own mental images, usually visual; images are somewhat elaborated from the literal text or existing picture
4	Creates and describes multisensory images that extend and enrich the text
5	Elaborates multisensory images to enhance comprehension; can articulate how the process enhances comprehension

Synthesises

1	Random or no response; may give title
2	Identifies some text events; random or nonsensical order
3	Synthesises with some awareness of event sequence
4	Enhances meaning in text with synthesis; may incorporate own schema; uses story

	elements to enhance their synthesis, may identify key themes
5	Succinct synthesis using internalised story/genre structure, identifies key themes; may articulate how synthesising promotes deeper comprehension

Appendix I. The administration of the word reading test for the QRI in Study 2 and Study 3.

Each child was introduced to these word lists, in accordance with the authors' instructions, in the following way:

> I have some lists of words for you to read. Some of the words will be easy for you and I expect some to be hard. Don't worry, you are not expected to know all of them. If you don't know a word right away, try your best to sound it out[17]. I cannot help you in any way and I cannot tell you whether you have got each word right or wrong. Just do your best. Are you ready?'

The word lists were designed by the QRI authors to provide a quick estimate of word identification ability. They suggest a beginning point two or more years below the child's grade placement. In using the lists in the UK, grades were replaced by years, so that a child in Year 3 – which equates to Grade 2 - was given the Primer (Year 1) level first. Correct responses in the word list were considered to be automatic if they were read within one second of exposure. If the response took longer it was marked as 'identified'. The lists were read one at a time and scored immediately. All correct answers (automatic and identified) were counted and a level calculated. Lists where 18 out of 20 words were read correctly were scored as at an independent level. Fourteen out of 20 marks gave an instructional level and below 14 was considered to be frustration level. Again, according to the authors' instructions, if the child scored at an instructional or frustration level, the researcher moved down a level until the child attained an independent level. Then the researcher moved up the lists until a frustration level was reached. The list which the child could read at an instructional level was taken as the level for selection of the reading passage. If the child read more than one passage at instructional level then the lower level was chosen, since the aim was for the child to read a passage that would ensure success.

---

[17] The words 'sound it out' were substituted for the original 'figure it out' which was a phrase which reflected the test's origin in the United States.

Appendix J. The administration of the QRI passages and think-aloud in Study 2 and Study 3.

Each child was introduced to the task in the following way:

> Thank you for helping me. I am going to be taping this, as I won't be able to remember everything you say, but just pretend it isn't there!

> There are no right or wrong answers today. Only you know what is going on inside your head. It will be very helpful to me if you can try and let me inside your head as much as possible.

> I am going to ask you to do some reading in a minute. But first, I just need to check what you already know about the subject.

The researcher then asked the concept questions from the QRI. These questions were designed to determine how familiar the child was with the subject matter in the text. After the concept questions, the researcher gave the following instructions for the think-aloud procedure:

> I want you to read this passage aloud. I am going to stop you as you read so you can tell me what you're thinking about as you read. Then, I want you to tell me exactly what you were thinking about as you read the passage. The important thing is that you pay attention and remember, so that you can tell me what you were thinking about while you were reading.

> You can tell me anything the passage makes you think about, any problems you had when reading it and what you think it is about.

Appendix K. Miscue analysis procedure and worksheet, adapted from QRI (Leslie and Caldwell, 1995, 2011) for Study 2 and Study 3.

Procedure

Miscues were transferred onto a sheet, adapted slightly from that provided in the QRI, in that the column for similar letter-sound patterns was not included (Appendix L) as the analysis was not intended to provide information about decoding skills. From the form it was possible to calculate Total Accuracy, which involved counting all miscues and also Total Acceptability, which counts only those miscues that distort or change the meaning of the passage. By counting the total miscues and subtracting them from the number of words in the passage it is possible to find the number of words read correctly. If this is then divided by the number of words in the passage, a percentage score for Total Accuracy can be calculated. This score can help to determine whether the passage was read at independent level (98% accuracy) instructional level (90 to 97% accuracy) or frustration level (less than 90% accuracy). The aim of the QRI is to administer a passage at the pupil's instructional level.

For Total Acceptability, only the meaning-change miscues are counted, and then the percentage is calculated in the same way. The authors give the percentages for levels of word identification in context as follows – independent level (98% Total Acceptability) instructional level (95-97%) and frustration level (less than 94%).

An analysis of the miscues, looking at the difference between Total Accuracy and Total Acceptability, will show if a pupil is self-correcting and/or reading for meaning, and therefore paying attention to the passage content rather than just attending to individual words. If a large number of the non-meaning-change miscues are corrected it would imply the opposite – that the reader is more focused on reading words accurately than in deriving meaning from the text.

Worksheet

Name……………………………………………………………………Text……………………………………………

Date…………………………………………………………………….

Level of Miscues: Accuracy   Independent (98%)   Instructional (95-97%)   Frustrational (<94%)

Miscue	Text	Ref.	Sub	Mispronunciation	Omission	Addition	Semantically acceptable	Self-corrected
					Meaning-change miscues	Corrected meaning-change miscues	Non-meaning change miscues	Corrected non-meaning change miscues
Total Miscue								
	Total accuracy Number of words in the passage minus total miscues/number of words in the passage							

296

Appendix L. Think-alouds: Coding schedule adapted from Schellings et al (2006) for Study 2 and Study 3, and examples of the categories.

Category	Text	Comment
B1 Comment on own reading behaviour		"I don't know that word."
B5 Summarising primarily in words of the text	Cats: lions and tigers in your house.	"It's about house cats and lions and tigers."
B6 Paraphrasing summary in own words	The octopus has a special sac.... that holds a dark, ink-like fluid. When an enemy comes close, the octopus squirts some of this fluid. It then swims away.	"The octopus swims away from the dying ink."
B6a Paraphrasing including info. not explicit	A cat drinks milk by lapping it.	"When the cat swallows the milk goes away into their stomachs."
B7 Paraphrasing incorrectly*	They don't worry that it will rain.	"They didn't know that it was going to rain."
B8 Predicting		No predictions made
B9 Visualising	You can find a river to fish in.	"I'm seeing pictures in my head of all these people fishing."
B9a Visualising that goes beyond text	Scientists have found a new type of light.	"I see a big building and on the top the door it says 'Science lab' and in there there's loads of computers and chemicals and lots of scientists."
B10a Questioning: a question that helps figure out what is happening or what the main topic is	People live in between the country and the city. They live in suburbs.	"Is it so the country's like here and the city's here and do they live like here?"- gestures to two distant point s and then the middle.
B10b Questioning: questions posed to help a deeper understanding of the text	People live in the city to be near their jobs. Cities have lots of factories, schools and offices. People work in these buildings.	"Why do people sometimes work and some people don't work?"
B13 Adding information connecting previous knowledge/experience in an associative way	One bug that likes to come out in the summer likes to bite.	"I was thinking about when I was in Turkey and these bugs were biting me"

B13a Adding information connecting previous knowledge/experience in a more active way , combining text information with previous knowledge	They think the octopus is a mean creature who attacks people and other animals.	It only eats fish like pretty much all sea animals eat fish.
B14 Evaluating text style	A cat drinks milk by lapping it. Because of the bumps, the milk stays on the tongue.	"You shouldn't really use 'because' to start a sentence."
B14a Evaluating text content	Suppose an octopus sees a crab. Patches of pink, purple or blue will appear on the octopus's skin	Laughs and says "That's weird".
B15 expressing nothing **		"I can't think of anything."
B16 other		Child pointed to a picture on the wall and asked "What's that wriggly thing?"
B17 clarifying word or idea	Farmers raise cows, pigs and chickens.	"Farmers something cows, pigs and chickens. ….Farmers …feed? Cows, pigs and chickens. It kind of connects to the next sentence – 'The main thing that these animals eat is grain' - so it probably does mean that."
B18 inferring incorrectly ***	We can weigh air. We can weigh two balloons. The one with a lot of air weighs more.	"If you put more air into a balloon it weighs more and it can go up."

* Not included in count when another summarising strategy is used, as it is using the same strategy, only comment is incorrect
** Not included in count of strategies
*** Not included in count if B17 is counted, as it is using the same strategy, only comment is incorrect.

Appendix M. MPIR Examples of levels for Study 2 and Study 3.

Schema

Level One

No response

Level Two

"It reminded me that whales can breathe air." Q. How did you know that? R." Whales don't lay eggs. I don't remember how I know." (Can say what text reminds him of, but cannot explain).

Level Three

"It reminded me of things we did in class about Egypt. I knew words that helped me understand the passage." (Relates background knowledge to text).

Level Four

"It reminds me cats I've seen all over the place and my auntie's cat and it reminds me of seeing lions and tigers at a safari park. 'Cos like if you've seen something or you know a bit about something already it makes it easier to understand." (Expands interpretation of text using background knowledge).

" 'Cos like the um.. well. You know about octopuses, and you how they like work and everything and you know stuff and you have to think of stuff that like reminded you of it and the story and the passage…. after you've read the story you know all what the octopus is doing 'cos you've read about it already."

Infers/Predicts

Level One

No response

Level Two

(Passage on air). Predicts passage will go on "to tell us what eggs turn into or some thing like that." (No apparent connection with text information).

Level Three

(Passage about cats, lions and tigers). Predicts "Maybe about how they sleep, how long they sleep. Stuff like that." (Prediction consistent with the text).

Level Four

"It said it can move a house and so it can move a car 'cos it weighs less." (Makes an inference and explains how it was made.)

Questions

Level One

(On Air). "Why do I have to brush my teeth?" (No connection with passage)

Level Two

"How do they move?" (Literal question).

Level Three

"How do the baby whales drink milk when it's water and milk?" (Clarifying concept).

Level Four

"Thinking of a question makes me think about what I'm reading." (Attempts to explain how the strategy helps).

Determines what is important

Level One

No response

Level Two

"The house what blew away." (One important element).

Level Three

"The top part is the most important because it's telling us all about the octopus. It's like the kind of setting where it tells you all about it." (Attempts to explain reasoning).

Level Four

"The middle paragraph about lights 'cos when I read it, it had all the main parts I didn't know about, all these other bits I knew already." (Key idea and why it was important explained).

Monitors comprehension

Level One

Q. How do you know you've understood something really well? R. "If there were books on the wall I could climb up them. I can read and climb at the same time." (Irrelevant response)

Level Two

"There was one word that was tricky. I tried to sound it out." (Problems at word level).

Level Three

"I'd use Clarifying Clara….if I had a word like 'predator', I'd read the sentence without the word, like…'it is hard for the something to find the octopus' and you see what would fit." (Uses a strategy to fix a word level problem).

Level Four

"I try and think about what could be happening. If it's not a picture book you could imagine it. For every little bit I would think of a picture. I would slow down; think of a picture in my head and then speed up again." (Uses a strategy to fix a global problem).

Visualising

Level One

Q. Do you ever make any pictures in your head when you're reading? R. "No. Just when I'm drawing pictures." (Does not make pictures when reading).

Level Two

"I can see cats, lions and tigers licking themselves." (Directly linked to the words of the text).

Level Three

"Saw a cat strangling a bunny and a lion hissing and spitting at a human." (Extended from text as text does not mention a rabbit or a human).

Level Four

"I can see me like sailing. It's windy, it's sunny, it's hot." (Extended to sensory experience).

Level Five

"It helps me to understand because what I can do is, I can focus on the passage and to see what it actually means. Does that actually make sense, or shall I do that again because it didn't make sense. I couldn't see it." (Explains how visualising is used to monitor comprehension)

Synthesising (Question 1)

Level One

"Cats, lions and tigers." (Just gives the title)

Level Two

"A busy beaver and some birds and squirrels and what they can do." (Some details from the text)

Level Three

"It's about the river Nile and how it floods and how they get their food." (Main idea)

Level Four

"Whales and fish and they're similar and not similar because the similar way is they've got the same fins and flippers and the non-similar way is the whales give birth alive whales and the fish give birth to eggs and the fishes have to find their own food straight away and the baby whales have to drink milk for a year." (Main idea with explanation).

Appendix N. Examples of the lesson plans used for the implementation of Reciprocal Teaching in Study 2.

Year 3 Literacy. Week 3				Term: Autumn 2009
	Learning Objectives	Activities	Speaking and Listening/ICT	Resources and texts
Monday	Strand 9. Creating and shaping texts Write short description of a known place	Shared Writing – write a basic description of the playground. Ask children to add golden adjectives and exciting verbs. Task: Begin to write description of the playground according to differentiated targets. XX:XX* to support. XX: At least 6 sentences XX: At least 10 sentences. Plenary: Choose a child to read out their best bit of work.		Differentiated targets Text for LA Stimulus photos if required (Printed from Easiteach)
Tuesday	Introduce the four strategies needed to read successfully – Predicting, questioning, clarifying, summarising.	Ask children what strategies they need to read well and what good readers do to help them to understand what they have read. Use PowerPoint presentation to introduce the four main strategies – teacher to dress accordingly. Task: XX: Choose one strategy and illustrate and say what it does. XX: Choose two strategies. XX: Draw all four strategies.	PowerPoint presentation.	PowerPoint presentation, Easiteach text.
Wednesday	Strand 2 - to introduce reciprocal strategies of predicting and questioning	Look at pg 27 – Dark is Fun. Think, Pair Share – predict content of chapter from title and illustration. Teacher to scribe children's suggestions. Read pg 27, 28, 29. Discuss if predictions were correct Read pg 30, 31. Ch imagine they were teacher, what would the teacher ask – teacher to model first Task – In pairs, write down 3 questions Plenary – read rest of chapter. Were your questions answered, chn reviewing learning .	Strand 1 – Speaking Strand 3 – Group interaction. Think, Pair, Share	Class reader – The Owl who was Afraid of the Dark P/C Pg 30 / 31

| Thursday | Strand 9 – select and use range of technical and descriptive vocabulary | Drama
Re-read pg 27-28. Put children in groups of 3 and dramatise the passage just read. Group direct one group – one group to show drama and other groups to explain what they liked, could do better etc.
Task:
Label picture of Plop describing physical features and personality.
In Creative writing books, write character descriptions of Plop and then transfer onto worksheet for display.
XX:XX to support – 5 sentences.
XX: At least 6 sentences
XX: At least 10 sentences. | Strand 3 – Group interaction. | Read and Respond

pg 50 |
| Friday | XX Group – to spell topic words correctly.

XX / XX – To spell words containing the short e sound. | XX group: Topic words

XX/ XX group: Words containing the short e sound. | | |

* All initials removed to preserve confidentiality.

Year 3 Literacy. Week 4				Term: Autumn 2009
	Learning Objectives	Activities	Speaking and Listening/ICT	Resources and texts
Monday	Strand 9 – select and use range of technical and descriptive vocabulary	Drama Re-read pg 27-28. Put children in groups of 3 and dramatise the passage just read. Group direct one group – one group to show drama and other groups to explain what they liked, could do better etc. Task: Label picture of Plop describing physical features and personality. In Creative writing books, write character descriptions of Plop and then transfer onto worksheet for display. XX:XX *to support – 5 sentences. XX: At least 6 sentences XX: At least 10 sentences.	Strand 3 – Group interaction.	Read and Respond pg 50
Tuesday	Strand 2 - to introduce reciprocal strategy of clarifying 30 mins	Look at pg 39/40 –Dark is Necessary Highlight unknown words, use Power Point to intro Clarifying Clara (unknown words).Ch imagine they were teacher – how would the teacher help you to work it out. Task In pairs choose words from differentiated passages XX – 3 words, XX – 5 words,	Strand 1 – Speaking Strand 3 – Group interaction. Think, Pair, Share	Class reader – The Owl who was Afraid of the Dark
Wednesday	Strand 2 - to introduce reciprocal strategy of summarising 30 mins	Look at pg 47/48– Dark is Necessary. Recap summarising Susie, using PowerPoint. Model summarising Task In pairs, write a summary of pg 49/50	Strand 1 – Speaking Strand 3 – Group interaction. Think, Pair, Share	Class reader – The Owl who was Afraid of the Dark

		Plenary Share ideas		
Thursday	Strand 11 – sentence structure and punctuation	Look at pg 54-57–Dark is Fascinating. On IWB children to annotate screen, adding correct punctuation to passage. Task Differentiated activities. XX – insert cap letters and full stops XX/XX – as above and speech marks	IWB – copy of text from pg 47 to 48.	Class reader – The Owl who was Afraid of the Dark Pc sheets
Friday	AF group: adding ing XX/XX group: Words containing the short i sound.	AF Group – adding ing XX/XX – To spell words containing the short i sound.		

* All initials removed to preserve confidentiality.

Appendix O. Bookmarks used as prompts during lessons in Study 2 and Study 3 (adapted from Oczkus (2003)).

PREDICTING

With Mystic Mike

- Predict what is going to happen next, or what each section is going to be about.
- Find clues from:
-

The front and back cover, if it is a book

The title and headings

The illustrations

Your own knowledge

What has happened so far?

- **Be prepared to explain why you made that prediction.**

QUESTIONING

With Dennis the Detective

- Think about questions you can ask others as you read.
- Reread the section, looking for parts that you could turn into questions.
- Ask questions that begin with

Who?

What?

Where?

When?

Why?

How?

What if?

- Ask one main idea question. Make sure you can find the answer in the text.
- Ask one 'between the lines' question. Explain how you used clues from the text to form your answer.

CLARIFYING
With Clara
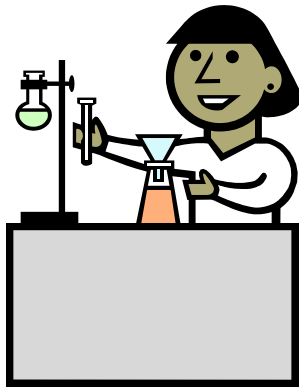
- Think about confusing parts or difficult words as you read.
- Highlight or underline words or sentences you don't understand.

For difficult words

- Check the parts of the words you know (prefixes, suffixes, root words, digraphs)
- Try blending the sounds of the words together
- Think about where you have seen the word before
- Think of another word that looks like this word
- Read on to find clues
- Try another word in the sentence to see if it makes sense
- Use a dictionary
-

For confusing sentences, paragraphs or pages:

- Reread the parts you don't understand
- Read on to look for clues
- Think about what you know about the topic
- Talk to a friend about what it means
-

SUMMARISING
With Susie

After reading:

- Look quickly through the reading and illustrations for main ideas.
- Reread, or skim and scan by running a finger and your eyes down a text to review it.
- Use your own words to summarise.
- Make sure you summarise important events or information in order
- Use words such as:

first

next

then

finally

- For fiction, use story words such as:

setting

characters

problem

key events

ending

Appendix P. Examples of the texts used to encourage inferencing in Study 2 and Study 3.

Text 1

After a long day at the shops, we went home. Dad pulled up in the drive, put the hand brake on and we all got out.

When Mum opened her bag, she frowned and looked over at dad. He felt in his pockets but they were empty. First we tried the living room window. That was shut. Then we walked round the back but the kitchen window was closed, too.

In the end, we went to the shed, got the ladder and leaned it against the back wall. It was Mum who had to climb up. We were too big to get through that window. As she went in, she had to be careful not to crack the basin – but she did it and soon the front door was open.

Text 2

We sat down at an empty table and looked at the menu.
'Do you want soup to start with, John?'
'Yes please, Dad,' he replied.
'What about you, Sharon?'
'Fruit juice, please.'
'And you, Mavis?'
'Fruit juice, thank-you, Dear.'
'Right, now let's think about the main course. There's fish, roast lamb and steak pie.'
'Remember, the doctor said John mustn't eat meat,' said Mum, '………………and I'll have the fish too, I think.'
'I'll have lamb' said Sharon.
The waiter came to the table with his pad and pencil.
'Can I take your order, Sir?'
'Yes, we'll have two soups and two fruit juices to start with and one lamb, one steak pie and two fish to follow.'

Appendix Q. A timetable of the implementation of Reciprocal Teaching in Study 2.

		22/09 1 hour	23/09 1 hour	29/09 30mins	30/09 30 mins	6/10 1 hour	13/10	Half-term	4/11 1 hour	11/11
September 2009 T1 Assessments		RT Introducing the four strategies The owl who was afraid of the dark					trip		RT Poetry text	service
	18/11 1 hour	25/11 1 hour	2/12 30 mins	9/12 30 mins	16/12		Christmas holidays	20/01 1 hour	27/01 1 hour	
	RT History text	RT Group reading text	RT Group reading text	RT Group reading text	play			RT Group reading text	RT Group reading text	
February 2010 T2 Assessments	3/02 30 mins	10/0 30 mins	Half-term	24/02 30 mins	3/03 30 mins	10/03 I hour	Easter holidays	23/03 1 hour	21/04 1 hour	28/04 1 hour	12/05 1 hour
	Visualising			Visualising	Visualising	RTV The Dancing Frog		RTV The Bunnyip	RTV Inference text	RTV I Inference text	RTV Inference text
	19/05 1 hour	Half-term	9/06 1 hour	17/06 1 hour						
	RTV Inference text		RTV Inference text	RTV Inference text						
May 2010 T3 Assessments										

Appendix R. Think-aloud and strategy interview responses at four time points in Study 2 for pupil 2.10 (A.).

Time 1

Think-aloud responses

When asked what he was thinking about A. said initially: "Um I don't know um I'm not thinking of anything." Later he said "When I read I don't usually think of anything. The only time I think of things usually is when I'm looking at pictures and thinking about what's happening there and what's happening next."

When questioned about what was happening he summarised primarily in the words of the text in the first instance, but after the second section, he paraphrased briefly, but in his own words:

"It's about cats defending themselves". The word 'defending' did not appear in the passage. After the third section, when asked to elaborate on his comment that it was about "their tongues", he continued: "They're rough and they have little suckers so when milk touches it, it can't get off, it can't move so when the cat swallows it goes away into their stomachs." Again, he used vocabulary which was not in the text ('suckers' instead of 'bumps') and in his paraphrasing included an inference that the milk would be going into the cat's stomach. A. was thinking when he read as he was clearly making sense of the passage, but he was not aware of his thought processes and did not question the text.

Strategy Interview responses

When interviewed, A. said that the passage did remind him of his own cat, and he was able to explain that thinking about an animal he owned and the things he had done with the animal would help him to know if the "story" was true or not. He was able to make a good prediction about what might come next –"their fur" and why – "because it hadn't talked about their fur yet". When asked if any questions came into his head, he said no, confirming the think-aloud analysis that questioning was not a strategy A. used. A. was able to say which part of the text he thought was the most important, but by saying: "I think the importantest (*sic*) part is the bit when they are telling animals to back away so they don't get killed by like dogs and hunters", A. showed he was not thinking about the meaning of the passage as a whole i.e. that cats, lions and tigers are alike in many ways. This was

demonstrated further when he was asked for a summary and he did not go beyond saying "cats, lions and tigers", even when prompted.

A. said the passage was "easy" because "there was not a lot of hard words", showing an emphasis on decoding as the measure of a difficult text. He reiterated this when asked what he would do to solve a problem and he replied: "If it's like a fact book and I didn't know what it meant I'd go to the like index and it would show me the words that I won't understand. And it'll tell me what they mean." When asked directly if he had problems saying the words or understanding the ideas, he did reply "understanding the ideas" and despite an earlier emphasis on decoding and the meaning of individual words, A. went on to say that he knew he had understand something really well because: "If I've understanded (*sic*) it really well it means I've learned like more things about these things I like to look about and…and learn more facts about."

A. did use visual imagery, but only reported it in answer to a direct question. He was able to talk about another book where he had made pictures in his head to help him understand.

Time 2

Think-aloud responses

The think-aloud showed A. is using a variety of strategies. He began with an inference (albeit an incorrect one) that Saudi Arabia is "a very poor country", and then summarised, but largely using the words in the text. After the second paragraph he showed that he was inferring the meaning of a word and relating what he was reading to something outside the text, as he spontaneously commented: "Oh I know what 'raise animals' means, it means like when you find one you look after it" (The question of what does 'raise animals' mean was in the initial concept questions.) He also commented that "every now and then an oasis appears", referring to the part of the passage that says oases spring up. Again, he was using his own words to paraphrase.

Strategy interview responses

A. wondered if this "story" was true when asked if it reminded him of anything, he made a good prediction about what the author might tell us next. When asked if he still wondered about anything, he said he now knew this was true, but he was still puzzled as to whether an oasis could really appear in the way the text said. He knew that the fact that they make wells was important, but he thought the rain fell on the people as they went up onto the roof, which is not what the text says.

A. thought this passage was "easy or middle". He thought the words were easy, but understanding the ideas was middle. He mentioned using Clarifying Clara to understand difficult words, and that he did this by "seeing other words which sound like it". He mentioned re-reading as a method of trying to understand. A. made pictures in his head of people looking around the desert for food and water, and he said he usually made pictures in his head when reading books at home. When asked how he would help someone else to do this he mentioned that the pictures in his head were moving ones. His summary repeated the detail about the size of Saudi Arabia, but added that it has a big desert. However, he had omitted any reference to the oases or the tribes.

Time 3

Think-aloud responses

When it came to talking about what was going on in his head all A.'s comments were questions. Although a lack of variety in strategies can indicate poor comprehension, or an inactive reading of the text, A.'s questions demonstrated that he was thinking hard about what he had been reading. For example, he asked:

"Is it smell that helps them [the dogs] figure out that an earthquake is coming?"

"Why did they [the dogs] go out of control?"

"What are the gases called?" and

"Why did it [the project] take three years?"

Strategy interview responses

When asked if it reminded him about anything, A. said "No, because I never knew animals could sense earthquakes", which was the most important part of the text. This showed he had developed an awareness of what was important, and that he was aware that he had learnt something from the text.

A. again made a plausible prediction and asked a very good question – acknowledging that this was one he has just thought of: "How long does it take for an earthquake to get to where it's supposed to go..........from where it starts, how long does it take to get to here?" The thing which A. still wondered about was, "When will another earthquake happen again?" which was central to the meaning of the passage.

A. again reported making pictures in his head when reading, and described one in great detail using more descriptive information than when asked about pictures at Time 2, and which went beyond the words used in the text.

One of them animals was sensing something and they were moving and going crazy and   breaking things down…..like…um…..there was a horse and I thought of a horse and there was like a   sack of hay and then it sensed it and went 'neighhhhhhhhhhhhhhhh' and it smashed right through it and it knocked the driver off.

A. thought this passage was "easy", and did not have any problems with the words or the ideas. This was reflected in the accuracy and comprehension scores, but it does not help understand if A. is more concerned with decoding or ideas at this stage. When asked for a summary he was able to give the main idea: "It's about earthquakes and animals knowing when an earthquake's coming" but he did not include the last paragraph, which explained about how scientists are hoping to use this behaviour.

Time 4

Think-aloud responses
A.'s first comment showed how he thought carefully as he read. After reading: "The sun is about halfway through its ten billion-year-long life cycle" A. said, "I'm wondering if in about five billion years the sun is going to die." The inference was made that as the sun is halfway through its life cycle than it must be going to die in five billion years, since that is half of ten billion.

In the second paragraph about nebulae, A. made a spontaneous comment, "I was thinking that there might be a nebula near our sun." This comment showed he was thinking as he read and he did not need any prompting to do it. A. went on to question if space ships could go through nebulae, and answered himself by saying that matter was dangerous and it could have exploded.

In the next paragraph he made another spontaneous comment, another question: "Does that mean before the sun was born it was a great ball of matter? Oh, so that's what the sun is, a great ball of matter. Ah, the sun is quite hot." He was again answering his own question by reasoning about what he had read.

A. concluded his think-aloud by saying something about what he had learned about the sun, which he read in the first paragraph, which showed he was thinking about what he had learned from the passage as a whole.

Strategy Interview

A. said the passage reminded him that he knew the sun was going to burn out, but that he had not known it would be in five billion years. He knew that referring to background knowledge enabled him to consolidate his knowledge and he was learning a little bit more each time. He made a good prediction about the passage going on to tell us about how the sun is going to change and he wondered if there might be another sun to take the place of ours when it dies. He said he sometimes has questions in his head when he is reading, but the think-aloud showed that he was questioning the passage and his own understanding throughout. However, his only strategy for comprehension repair was to re-read. He said he did not make pictures in his head for this passage but that he does when reading Harry Potter. He was able to explain that to be good at making pictures you need to imagine a world and then put yourself in it. His summary was succinct, and again included his new-found knowledge about the age of the sun.

Appendix S. Think-aloud and strategy interview responses at four time points in Study 2 for pupil 2.3 (B.)

Time 1

Think-aloud responses

B. was unable to comment on her thinking at the second stop point and at the other three she gave short summaries, sometimes in the words of the text and sometimes in her own words. She was able to make inferences and she made one reference to background knowledge when she said she had not heard anyone say 'busy as a beaver'.

Strategy interview responses

B. demonstrated that she is unfamiliar with the structure of expository texts as her prediction was based on narrative text: "The busy beaver has a big family and they go in the house together."

She was not able to think of any questions about the text, and her ideas about which parts of the text were the most important contained some events but no explanation. When asked about whether the text was easy, middle or difficult, B replied that it was: "a bit difficult 'cos it's like I've only read a few books with words in it like these and its quite tricky for me." This showed that she concerned with reading at the word level. She did make pictures in her mind, but they were described in the words of the text.

Time 2

Think-aloud responses

When asked to say what the passage made her think, B replied that it made her feel hungry, as the octopus was eating. However, this was a rather general response and did not connect what she already knew with what was happening in the passage in a way which would have increased her understanding. Further, she went on to say that it was "how the animals capture people and eat them" which made her hungry, which was an incorrect summary of the text (which explained that the octopus does not normally attack people, only in science fiction movies). At the second stop point B. said she was not thinking about anything, which was also her response at the subsequent stop points, even when prompted. Therefore, she was only able to respond to the text once.

Strategy Interview

When asked to make a prediction about what might happen next B. used the same strategy as she did at Time 1, which was to offer a narrative text explanation: "and the octopus lived

316

happily ever after?" She did not have any questions about the text, nor did she wonder about anything when she had finished reading. B. was able to say the first paragraph was important as it provided information about "the setting", although this was "a long passage" and '"got like quite hard words." B. was unable to explain how she would know if she had understood something, but when asked what someone should do if they did not understand she said: "You should use Clarifying Clara, Mystic Mike, Dennis the Detective and …what's that other lady?...Summarising Susie." This response showed an awareness of the characters, but did not show any understanding of what they do, or how the strategies are used. B. did have clear ideas about visualising however, and described her image of an octopus, which went beyond the words in the text: "It was purple and it had circles on it like that (she made a circle with her hand) and they were quite big and then it sticks like that (she made a kissing sound)."

Time 3

Think-aloud responses

At the first stop point B. attempted to summarise what she had read, albeit incorrectly at first: "Well scientists had made up a new light' (the passage said scientists had found a new type of light)…..it's called laser light and it comes..it's…it's used in medicine, industry and science research." The summary was in the same words as the text however, rather than in her own words. At the second stop point, B. made another incorrect attempt at summarising, but followed it with a correct one "If I got a white light it would go all the colours of the rainbow", but again she used the words of the text and not her own. At the third stop point, she summarised in the words of the text again, but then attempted to connect what she was reading with her background knowledge in order to expand on what she had read. After reading that light only travels in straight lines she said, "...Sometimes like this light is about that big (points to the fluorescent tube light in the classroom and extends her hands wide) and it covers the whole room but that light (points to a single light bulb) over there just does that bit." Finally she made a summary which included information which was not explicit in the text. After reading 'laser beams can be used to carry radio and telephone messages', she said, "…if you're phoning somebody that's where all the power comes from the..the laser light". But then she continued with another incorrect summary, which stemmed from the problematic sentence which was high-lighted in the miscue analysis; by failing to understand the sentence about reattaching the retina, B said,

"…sometimes you can get blinded by laser light 'cos your eyes could get steamed up and then fall out."

Strategy interview responses

This time B. made a better connection with background knowledge, as the passage reminded her of a book her mother had read to her about lights and lights of different colours. She also made a much better prediction, more in keeping with a non-narrative text - it would be about "how they make laser light". She also had questions for the first time – about what laser light really was and why does light flicker. B. also thought questions helped her to understand what she read and she said they did come into her head when she was reading. When asked about what was important in the text, she was able to identify a section ("the middle paragraph") and explain why it was important - "'cos it's got how the lights work and stuff" and, more importantly: "When I read it, it had all the main bits I didn't know about, all these other bits I knew about already." This showed an attempt to reconcile her background knowledge with the new information in the text and was evidence of an active engagement in the text, which had been missing at Times 1 and 2.

When questioned about her understanding, B.'s main concern was still at the word level: "…it had hard words…like tumours….and that one, reeteene (*retina*)." But she also said that her concern was not with saying the words, but "understanding what they meant". As there was a lot of unfamiliar vocabulary in this passage, this was a response which showed comprehension monitoring. B. was also more aware of the possibility that understanding can go awry; when asked about a friend not understanding something she replied: "I would say do you think um it's gone right or wrong here and they'd probably say wrong". This showed that for the first time B. appeared to be acknowledging that there was more to reading than just saying the words and that understanding does not necessarily follow. It is something you need to think about too. However, B did not report making any pictures in her head when she was reading this passage, despite the fact she had done so previously.

Time 4

Think-aloud responses

B. made more detailed responses than at the other time points. At the first stop point, she recognised the existence of good background knowledge, by referring to the fact that she had recently studied Ancient Egypt, that she had written about Tutankhamen, and that she was thinking about all the different things she had done about Egypt. At the second stop

point she said she was wondering what papyrus was, and she used a clarifying strategy by saying that it was like "papaya", but that she did not know what it was.

At the third stop point, B. talked about what she could see in her head: "a long long river and a boat….and people were trying hard and they to like (*sic*) get the boat out of the river and carry it." This was an image which went beyond the words in the text, and which was a moving picture rather than a static one. It also involved a degree of understanding of the effort involved. At this stop point B. also said, "When I say really long hard words like cataracts it like here it hurts (points to her head)." Despite her excellent word reading skills, she was acknowledging that sometimes a real effort was involved to decode. Finally, at stop-point four, B. said: "I was just reading that but I wasn't thinking anything" which was an admission that she knew she had stopped thinking in the final paragraph and reverted to just reading the words. The awareness that she could read without understanding, when she forgot to think, is important. It showed a new awareness that reading is an active process, which requires engagement.

Strategy interview responses
In the interview B. again referred to her background knowledge from having studied Ancient Egypt in class. However, she was better able to explain how this helped; specifically she knew that to have a familiarity with some of the words and their meanings would be a help. She was able to make a prediction consistent with expository test for the first time, saying that the passage might go on to tell us how often the Nile valley gets flooded. B also had a question in her head, about where did they store the food from the surplus harvests, and she said she still wondered about what papyrus was like.

B. knew that the first paragraph was an introduction and that the rest of the passage was about the flooding of the Nile, which was the main idea. When asked about the passage's difficulty, B said it was "in the middle of easy and medium", and that understanding the ideas was more difficult than saying the words. She said she could use a dictionary or ask the person next to her when she was not familiar with something she read. Rereading was also a strategy she said she would use when she did not understand something, and that she would ask someone what it meant if she read it a couple of times and still did not understand it. B. was also able to talk about how pictures helped her: "There's a picture in my mind and because it's in my mind it helps me with the story because it keeps on going on and on." The implication was that if the picture did not keep going on she did not understand the text.

319

Time 1

Think-aloud responses

At the first stop point C. had to be prompted twice for a response. When she did reply she gave a retelling of the first sentence, using the words from the passage. At the second stop point she again had to be prompted with the question 'What do you think it is about?' before replying in the words of the text. At the third stop point she did not need prompting, but once again she replied briefly and did not use her own words. It should be noted here that the think-aloud instructions allow for the child to be prompted once, and C. was prompted more than this. However, the researcher felt that to give her confidence it was important that C. was able to give a response at each point. Her responses were coded, and therefore her score for this time point should be considered to be higher than expected, rather than lower.

Strategy interview responses

C. was unable to think of anything of which the passage reminded her. She was able to make a prediction, albeit an incorrect one ("the whales might eat the fish"), and possibly based on a narrative text structure rather than that of expository text. She did not have any questions in her head, although she did say when asked if there was anything she still wondered about, "They lay eggs", suggesting this may have been novel to her. C. did not know which parts of the passage were important, not whether she had had any problems with understanding the passage. When prompted (again in the interests of maintaining her confidence), C. did say when asked about a good way to work out a word she did not understand, that she would go back and read it again. However she was unclear about how reading it again would help her. She was able to describe a picture she had in her mind, of a blue whale with water coming out of the top. This went beyond the words of the text as the term 'blue whale' was not mentioned.

Time 2

Think-aloud responses

C. had four times as many responses at Time 2 than at Time 1 (16 comments compared to 4). At the first stop point she talked about the passage reminding her of where she lived, in a

village, and then said they were thinking of moving to closer to the school, but that her mum could never live in a city or a town and that she could only live in a quiet place. C. was personally identifying with the text, and that was something she continued to do at the next stop point, where she was prompted by a section saying that you can go for walks in the country, to talk about a favourite walk she did with her Mum whilst her brother and Dad were at football.

Finally, after the last paragraph, she asked: "Are suburbs villages?" and "Are they like villages what are quite like quiet?" This shows that C. had not only recognised that her comprehension had broken down (in not knowing the meaning of 'suburbs') but she had made a good attempt at clarifying.

Strategy interview responses

When asked what the passage reminded her about, C. repeated her think-aloud observations about where she lived and how it was quiet there. She was also able to make a prediction more in line with expository text - "It could be about how busy cars are and how fast they go through towns and villages". Her question was about the meaning of the word 'suburbs' and she had a variety of ways to find out about it, she could use a dictionary, read on, or "remember something you've learned", all strategies suggested by Clarifying Clara. She also thought that questions helped you to remember things which were important, but her strategies for remembering what was important were to write it down or keep saying it in her head. These strategies are more relevant to memorising in general and not applicable to remembering the important parts of a passage, without an effort at summarising or selecting the main idea first.

C. thought the passage was "easy" as "there weren't many tricky words", which suggested her focus was still at the word level. However she did say that understanding the ideas was important. Her visualisation, although brief, was not restricted to words from the text ("the woods...and lots of cars when I said city") since the word 'cars' did not appear in the passage. As at Time 1 C.'s pictures served to integrate her reading of the text with her background knowledge.

Time 3

Think-aloud responses

C. made a summary in her own words at the first stop point, observing that "beavers work very hard". At the second stop point she said (without being asked what she was thinking

about): "I have a picture in my head of a beaver blocking a river…the beaver was building .um.. a house and he was blocking all the water." C. had gone beyond the words of the text to describe what she could see, which was a moving picture of what the beaver was doing.

At the third stop point she made a short summary in her own words, which involved inference. C. referred to the beavers "swimming under the water to go inside its home". The text only told us the lodge was under the water and did not say that the beaver had to swim to get to it. The same strategy was used at the fourth and final stop when she again used her own words to say that the animals had to find another place to stay because of the beavers' actions.

Strategy interview responses

This time, the passage did not make C. think of anything she already knew. This contrasted with all the associations she made at Time 2, but probably reflected the fact that she is unfamiliar with beavers, whereas she knew a lot about living in the country and had her own opinions about that. She made a prediction consistent with expository text – that we might go on to learn about beavers swim, and she had a question in her head about the text: "Why are beavers such good builders?" C. knew that some parts of the passage were more important than others and she said she made pictures in her head to remember what was important. She thought the passage was "middle" as it had two hard words in it, but she continued to believe that reading the words was not the only factor, and that understanding the ideas was more important. She knew that if she had a problem with understanding she needed to "read back", and that visualising was a way to correct breakdowns in comprehension.

Time 4

Think-aloud responses

C.'s responses were rather brief. The passage made her think about people living in cities and other places but she did not make the personal connections she had made at Time 2. At the second stop point she made a short summary in the words of the text, whilst at stop point 3 she visualised "a farmer feeding the animals and no cars." Finally, she made the same observation as did at this point at Time 2, that a suburb "might be a village or another name for it."

Strategy interview responses

C. repeated her assertion from the think-aloud that the passage did not remind her of anything. Her prediction was consistent with the structure of expository text, and she had the same question – "What is a suburb?" She was able to identify the important parts of the passage, "It like tells you what suburbs and cities are and what happens there", and said that some people might not know what cities are. She thought it was "medium" but she thought that "understanding what things mean was the hardest". This seemed to imply that it was not saying the words which was difficult, but their meanings. This was a step up form a straight forward decoding emphasis, but it may not have been as wide as the understanding the ideas focus which she had at Times 2 and 3. C. said she used pictures "to see what is happening", but this time she did not suggest she was using this as a monitoring device. Her summary however contained a key theme – "It was about people and why they live in the kinds of places they live in."

Appendix U. Think-aloud and strategy interview responses at four time points in Study 2 for pupil 2.9 (D)

Time 1

Think- aloud responses

At the first stop point, D made a personal connection with the text by saying it reminded him of the ride at Alton Towers, called 'Air'. However this was just an association and it did not further his knowledge about the text. At the second point he did not have any thoughts, despite the fact he had been reading about something in direct contrast to his background belief that the wind could only move light things. This section of the text told him that strong winds can move heavy things. At the final stop point he made another general association, that 'weigh' made him think of "the weighs you make cakes with" (i.e. scales).

Strategy interview responses

When asked how thinking about the Air ride at Alton Towers had helped him to understand the passage, D said it had given him a picture in his head, but he could not elaborate on why that might help. He was not able to make a prediction about what else the author might go on to tell us, other then "wind" when prompted, which did not go beyond what was already in the text. He did not have any questions in his head and the only thing he wondered about at the end was "my cat 'cos he's really fat………and he's always yawning", which was not connected to the text. D.'s response to a question about any problems in understanding the text was "split them up" showing a focus at the word level, and he had not had any problems anyway because "the words are just like the words I read back in year 2."

Time 2

Think-aloud

At the first think-aloud opportunity D. said he was thinking about a farm where they were growing vegetables and flowers. This was an image connecting his background knowledge with the text, since there was only a reference to spring being the season when new life begins, and that it is when tulips come. He went on to say that he had had trouble with the word 'tulips': "I got it wrong the first time but then I got it right the second time"; this showed he was monitoring his comprehension. At the second stop point he made a summary in the words of the text, whilst at the third point he said he was thinking about "a nut and a squirrel", both words from the text, but he went on to elaborate by saying, "I thought of a squirrel 'cos I have ...we have... some sort of birds and we put nuts but bird

325

food and the squirrels eat it. And hang upside down." This was different to just making a connection with a ride called ' Air' as he did at Time 1, since this time background knowledge was enabling him to visualise an event from the text, which fostered his comprehension whilst doing so. D. also said: "When I was here it says 'summer…fall[18]', and I've never heard of that season……. Spring, summer, autumn, winter……it's the one where the leaves fall off…autumn." This was evidence of successfully clarifying a word of which he did not know the meaning.

Strategy interview responses

When asked about background knowledge D. was able to explain that thinking about the squirrels in his garden had helped him as it made a picture in his head, but he did not have any thoughts about a prediction. He did have a question about the word 'fall', "When I came to the word 'fall' I said in my head, what does that mean?" D. was seeking clarification of word that his comprehension monitoring had identified as problematic. He was also able to identify an important part of the passage - the dates- which were important to him as they represented new knowledge, as was evident from the concept questions. He had much clearer ideas about comprehension monitoring, suggesting that if someone in his class had problems understanding they could: "Try and sound it out...um... yeah...try and ask them to try and sound it out and to think…to think…like……and to help them something like a picture in their head would help them…Um… I would say, read it again." Although D has suggested sounding out the word he also knew you needed to understand what the word meant and that making a picture could help to clarify what you were reading. Also, re-reading implied a focus beyond just the word level.

D.'s pictures went beyond the words of the passage. He could see "a beetle", although only the words 'insect' and 'bug' are used. This showed he was making a model that was beyond the text level as it was incorporating his background knowledge.

Time 3

Think-aloud

At the first stop point it was clear that D. was visualising from the text. After reading: 'House cats, lions and tigers are part of the same family. When animals are part of the same family they are alike in many ways. House cats are like lions and tigers in many ways too. When

---

[18] The word 'fall' for autumn appeared in the passage as the QRI was published in the USA. The passages were adapted by the researcher to use in the UK, but this term was not replaced.

kittens are first born they drink milk form their mothers. Lions and tigers drink milk from their m others too. When kittens are born they have claws just like big cats. Claws are used by lions, tigers and kittens to help them keep away their enemies', D. said: "There's like a forest, grass growing up and a tiger head poking then and there's a cat going nrrrr…..and there's a monkey swinging from a tree above the thing, the tiger." This was a visualisation which went beyond the words of the text and information about the background (the forest and the grass) as well as what was happening – the cat defending itself by growling.

At the second stop point, having read that cats hunt for animals by stalking them, D. said, "I've got a cat and three rabbits and um my sister lets the cats ...umm... rabbits, run in the garden and when the cat's too close my rabbit's thumping with the tail." D. was making an association with what he knew about animal behaviour from personal experience, but he related it closely to the text in that he realised the rabbit was defending itself from the cat. Then he showed he was visualising again, he said he was thinking of, "A picture with a broken neck, like that (he gestured a snapping action)."

The next paragraph talked about cats puffing themselves up when they are afraid, which prompted D. to comment that;

> 'Cos my cat's really fat, when he puffs up his belly goes up and it makes him look really big like that (gestures a large round shape) 'cos when he ..um...sits up he can only sit up for a couple of seconds then he jumps up like that far (gestures with his hand a foot off the ground)'cos he's so fat.

Again he was relating what he had observed of his own cat to what he was reading in the passage. He continued in this theme, relating what the passage told us about cat's tongues with his own experience:

> I've got two things now again now...um...it remembered me of the rough tongue 'cos when he licks me… what he does is he sits like this (gestures with his head tucked into his neck) goes round with his neck and he cleans his fur and then if he wants to clean his tail he grabs it with one paw and he brings it up.

D. gave a very clear picture of what he was seeing in his mind, which went much further that the words in the text, which only said that cats 'clean themselves with their tongue'.

When asked about what the passage reminded him of, and why thinking about that helped, D said: "My cat…'cos... when it says they have...um...rough tongues I was like, what type of rough? Do they have ground on their tongues? …um... or it could be spiky and it would be easier to know if you actually had a cat and you know." Besides showing that he understood how background knowledge helped he was also giving an insight into how he clarified a word, since he initially thought that 'rough' must be referring to "rough ground" but then thinking about his own cat he realised it must have been 'rough' in the sense of "spiky". This was particularly interesting in the light of research on poor comprehenders (Cain, 2006; Pinkerton 2010, Pinkerton & Nation, 2010) which has shown that such readers have more difficulty in disinhibiting meanings that are not relevant compared to controls. D. shows that he can now recognise a meaning that is not needed in this context and reject it in favour of the meaning that makes sense.

Time 4

Think-aloud responses

D. again demonstrated his strong visual sense, when he said he was thinking of: "A forest. A forest with lots and lots of trees…..It's like a kind of painting." The words 'forest' and 'trees' did not appear in the text , only 'plants and animals', so D. was going beyond the literal meaning of the text and developing his own situation model, enhanced by his visualisation. His visualisation continues at stop point two, when he could see, "The school lake and someone standing on one of the lily pads and it's really (unintelligible)". D. was using his personal experience, of having seen the lily pads mentioned in the text on the school lake, but he was now introducing a novel image of someone standing there on the lily pad. The fact, given in the text that a lily pad could be strong enough to stand on, has been transferred to an environment with which he was familiar to produce an image of something he has never seen before.

The next paragraph, about pine trees, reminded D. about a hundred foot Christmas tree in had in his old back garden, but this was a more general associative image and did not add greatly to his understanding. He continued in this theme when the final paragraph (about cacti) reminded him of a school trip he went on in Year 3 when they saw cacti in a giant greenhouse, and that when he was three or four he had touched a cactus and got pricked. Again these were general associations to personal experience but they had not

enhanced the meaning of the text in the way that the image of someone standing on the lily pad had.

Strategy interview responses

D. remembered the images he had talked about in the think-aloud: "A trip, pricking my finger, a Christmas tree in my garden, the school lake and lily pads." This level of recall showed that the visual images he had created served to enhance his memory of the things he was reminded about by the text. He also appreciated that these pictures had a comprehension fostering aspect: "If you read it's like a bundle of words but if you think about it you can get a picture in your head and then you get it."

D. was unable to think of any questions he had in his head whilst reading. As there were several words he mispronounced and some at which he had several attempts, it would appear that he was not monitoring his comprehension as thoroughly as he was at Time 3. He thought the most important part was that you can step on a lily pad, as that's "really impressive". We have seen how connecting this idea with an image from his own personal experience had served to make this fact so memorable. When asked if the passage was easy, middle or difficult, D. said it was middle as some of the words were quite hard, and that saying the words was the hardest bit. The focus appeared to have shifted back to the word level somewhat. However, he then went on to say that he would need to reread if he did not understand which implied a concern with meaning at a wider level. D. made a good summary of the passage, showing his understanding did go beyond the text level.

Appendix V. Permission letter for Study 3.

Institute of Education
University of London
20 Bedford Way
London
WC1H OAL
12/09/11

Dear Parent/s,
This letter is to introduce myself and to explain about some research which I will be carrying out in Year 3.
My name is Frances Hampson-Jones. I am currently studying for a University of London PhD at the Institute of Education. I am undertaking a research project concerning interventions to improve reading comprehension.
I am writing to ask permission for your son/daughter to take part in this research. This will only involve the kinds of activities which children already do in school as part of the English curriculum.
I am interested in observing how a class teacher can improve reading comprehension, and how the children respond to that teaching. Your child will continue to be taught by their normal class teacher, but I will be involved in observing the class, and talking to the children about what they have learned. I would like to begin with some interviews and assessments of reading ability. There will be two teaching phases during the year, and children's reading will be assessed after each one. These assessments will be very useful to the class teacher and to your child's teacher/s in the following academic year.
The interviews and some observations will include the use of a tape recorder. Your child will not be identified on any tape recording. Please be assured that all research is strictly confidential. No names will be kept on any computer system and names will not be used when reporting findings.
 Research undertaken by students at the Institute is subject to regulation by an Ethics Committee and is conducted using guidelines from the British Psychological Society (see www.bps.org.uk for further details), of which I am a graduate member.
The findings of the research will be made available to the school, and to anyone who is interested, after I have completed the study. The findings (using anonymised data) will also be written up for submission to a peer reviewed journal.

You have the right to withdraw your child from this research project at any time.

Mrs …….. and Mrs………. have kindly given me their full support.

I am very much looking forward to starting this project in September 2011.
Yours faithfully,

Frances Hampson-Jones

If you do not wish your child to take part in this research project, please notify the Head Teacher in writing, by Wednesday 21st September 2011.

Appendix W. An example of a lesson plan/script for Study 3.

WALT: to use strategies as we read.

Whole class- 10 mins
Ask for a summary of last week's section of The Iron Man.

(The Iron Man challenged the dragon to an ordeal by fire. The Iron Man lay on a bed of fire, but the dragon had to lie on the sun. They both survived but the dragon was horribly changed by the ordeal.)

Ask for a summary and for feedback on it. Emphasise why summarising is important.

Summarising what you are reading in your head is a good way to check you are understanding what you are reading, so it is a good idea to stop every now and again when you're reading and think 'What was that all about?' and make a short summary in your head.

Ask someone to describe the picture in his or her head of the dragon when he returned to earth.

Visualising is another good way to check you understand what you read. If you can't see what is happening in your mind it might be a good idea to read that bit again and try and picture it in your head.

Group work-10 mins

Read the next extract (extract nine) as a group, and highlight any words or groups of words that need clarifying. Discuss the words in the group.

Whole class-5 mins

Ask for words that needed clarifying and get the meanings. Ask how they worked it out.

Did they:

      Sound it out

      Re-read

      Read on to find clues

      Try another word

Group work-10 mins

As a group, think of questions you could ask the other groups about this extract. Thinking about questions as you read helps you to think more deeply about what you are reading and what it means.

Try to think of a hard, between the lines question, where the answer is in the text if you think hard about it, and an 'outside' question, where you would have to use your background knowledge, or look it up, to find the answer.

The 'teachers' should write down the best question for each type (on sheets).

Whole class-10 mins

Ask each group for their between the lines question (write them on the board).

Ask other groups for their answer. Decide if it was a 'between the lines question' (identify others as clarifying, predicting, or literal (easy) questions if necessary).

Ask each group for their 'outside' question (write them on the board)

Ask the other groups where they could look for the answer to the questions (an adult, a book, the internet etc.)

Individual work-5 mins

Prediction

Write down what they think will happen next – the end of the story of The Iron Man.

Whole class- 10 mins

Now we will see if those predictions will come true.

Read the end of the story (maybe get children on the carpet to listen).

(Last paragraph p.57 to the end)

Appendix X. A timetable of the implementation of Reciprocal Teaching in Study 3.

	Half-term	1/11 30 mins	3/11 30 mins	8/11 30 mins	10/11 30 mins	15/11 30 mins	17/11 30 mins	22/11 30 mins	29/11 30 mins	30/11 30 mins	6/12 30 mins	13/12	Christmas holidays	5/01 1 hour	10/01 1 hour
Sept/October 2011 T1 Assessments		RT Rules for working as a group Introducing the four strategies				RT Dinosaur text						Play rehearsal		RT Inference texts	
	12/01 1 hour	27/01 1 hour		24/01 1 hour											
	RT Inference texts														

	21/02 1 hour	28/02 1 hour	6/03 1 hour	13/03 1 hour	20/03 I hour	27/03 1 hour	Easter holidays	24/04 1 hour	1/05 1 hour	8/05 1 hour
Feb 2012 T2 Assessments	RTV 30 mins Visualising 30 mins Iron Man	RTV 30 mins Visualising 30 mins Iron Man	RT 15 mins Visualising 45 mins Iron Man	RTV 15 mins Visualising 45 mins Iron Man	RTV 30 mins Visualising 30 mins Iron Man	RTV Iron Man		RTV Iron Man	RTV Iron Man	RTV Iron Man
	15/05 1 hour RTV Iron Man									
May 2012 T3Assessments										

Appendix Y. Sample Gant charts showing how group work was carried out in Study 3.

BFG	Clarifying	Teacher	WRITTEN QUESTIONS	WRITTEN ANSWERS
CHARLIE		WRITTEN QUESTIONS	WRITTEN ANSWERS	Teacher
FOX		Teaching assistant	WRITTEN QUESTIONS	WRITTEN ANSWERS
CROCODILE		WRITTEN QUESTIONS	WRITTEN ANSWERS	
JAMES			Teaching assistant to tell BFG and Fox to swap after 25 mins	
DANNY			and help groups were necessary	
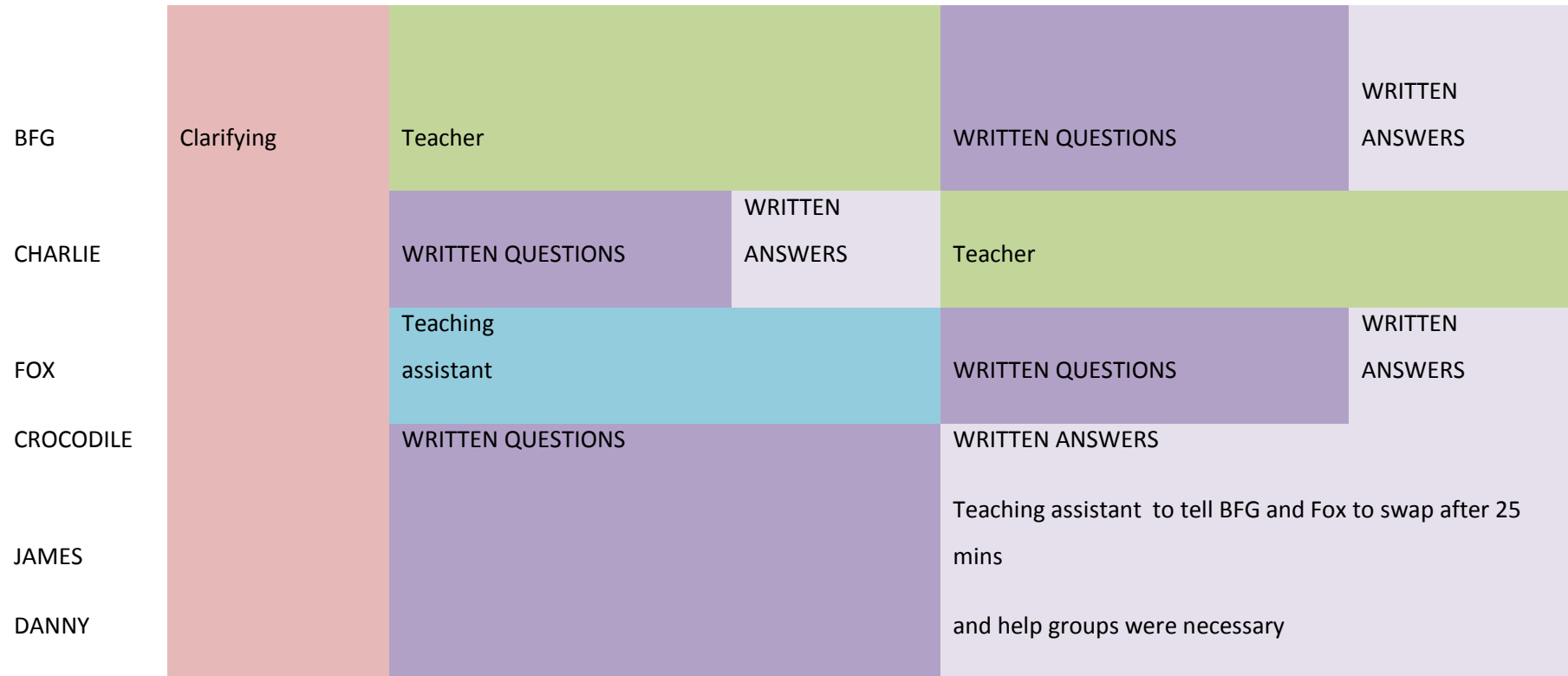
CROCODILE	Questioning	Teacher	WRITTEN QUESTIONS	WRITTEN ANSWERS
JAMES		WRITTEN QUESTIONS	WRITTEN ANSWERS	Teacher
DANNY		Teaching assistant	WRITTEN QUESTIONS	WRITTEN ANSWERS
BFG		WRITTEN QUESTIONS	WRITTEN ANSWERS	
CHARLIE			Teaching assistant to tell Crocodile and Danny to	
FOX			swap after 25 minutes and help groups as necessary	