UNIVERSITY OF ADELAIDE LIBRARY
24.3.95

Topical Inferences as Strategies to Improve Comprehension

Suzanne Pope
Department of Education
University of Adelaide
June 1994

Awarded 1995

Acknowledgements

I would like to thank the Department of Education at the University of Adelaide for the short-term release from teaching duties required to carry out the experimental work for this thesis.

Thanks to participating staff and students at Madison Park and Salisbury Primary Schools, and especially at St Joseph's Hectorville Primary School. Special thanks to Sybil Wakefield for her cooperation and encouragement.

I am grateful to Geoff Russell for his support and assistance with word processing.

Particular thanks go to Dr Jack Rowell, my supervisor, for his willingness to discuss and comment on my work, helpful suggestions, but especially the unfailing support and encouragement without which this thesis may never have been completed.

Disclaimer

I certify that this thesis does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any university; and that to the best of my knowledge and belief it does not contain material previously published or written by another person except where due reference is made in the text.

I consent to this thesis being made available for photocopying and loan if applicable and if accepted for the award of the degree.

Signed:

Abstract

The ability to learn from expository texts is an important skill. Yet research indicates that this skill is often poorly developed in upper primary school children, and that traditional comprehension instruction does not explicitly teach strategies for efficiently processing instructional text.

This study aimed to design and implement a training programme to improve the ability to learn from expository texts. Children were taught strategies which involved deeper processing of headings and subheadings, paragraph summarisation and comprehension monitoring.

The theoretical foundations of the study included models of reading by Just and Carpenter (1980; 1987) Kintsch and van Dijk (1978) and van Dijk and Kintsch (1983), as well as Markman (1981; 1985) on comprehension monitoring, and Jacobs and Paris (1987) and Borkowski and Kurtz (1987) on metacognitive knowledge and skill.

Palincsar (for example, Palincsar and Brown, 1984) has successfully used strategies such as those above to improve the comprehension of relatively small groups of below average comprehenders in grade 7. This study extends the work of Palincsar in several ways. Younger children were taught, and the group included readers of varying ability levels, including good readers. The group was larger, involving a whole class of 25 children. Teaching materials and student responses were written rather than an oral dialogue.

One intact, mixed ability grade 5 class was taught daily for 3 weeks. After training this group showed significantly improved comprehension relative to a control group, both immediately and after ten weeks. Class differences on two

of the three transfer tests approached, but did not quite reach significance. The programme was suitable for use with a mixed ability class, since there was no interaction between treatment and reading ability. A significant discriminant function was found, with testable educational implications, to discriminate students who improved from those who did not.

This study demonstrates a wider application of the type of approach used by Palincsar than has previously been investigated. The written format ensures a high proportion of academically engaged time, even in large classes, and could make the programme easier for teachers to implement, although this suggestion remains to be experimentally tested.

Contents

1 Theoretical Models of Reading 13

1.1 Introduction 13

1.2 General assumptions 15

1.3 Stages of processing 20

1.3.1 Get next input 20

1.3.2 Word encoding and lexical access 22

1.3.3 Case role assignment (semantic analysis) 41

1.3.4 Interclause integration 50

1.3.5 Sentence wrap-up 68

1.3.6 Macrostructure formation 73

1.4 The role of knowledge in processing text 95

1.4.1 Subject knowledge 95

1.4.2 Metacognitive knowledge 106

1.5 Conclusion 130

2 Skill Differences 134

2.1 Introduction 134

2.2 Processing efficiency and working memory capacity 136

2.2.1 Efficiency of word encoding and semantic memory access . 136
2.2.2 Working memory capacity and comprehension 147
2.3 Integration 162
2.4 Macrostructure formation 179
2.4.1 Assessing the topical coherence of paragraphs 183
2.4.2 Determining the importance of ideas in texts 197
2.5 Metacognition 215
2.5.1 Metacognitive Knowledge and its Application 217
2.5.2 Motivational influences on performance 232
2.5.3 Comprehension Monitoring 237
2.6 Conclusion 284

3 Instructional Programmes **288**

3.1 Introduction 288
3.2 Encouraging integration 302
3.3 Macrostructure formation 312
3.3.1 Using explicit text features and conventional text structures 314
3.3.2 Generating macrorelevant information 322
3.3.3 Self-questioning about main ideas 329
3.3.4 Teaching summarisation rules 344
3.4 Improving metacognitive awareness 352
3.5 Composite packages: Palincsar 367
3.6 Composite packages: Paris 384
3.7 Conclusion 400

4 Design of Materials and Method 415

4.1 Design and Selection of Materials 415

4.1.1 Reading booklet 416

4.1.2 Work booklet to accompany teaching texts 418

4.1.3 Comprehension tests and practice passages 421

4.1.4 Transfer tests 424

4.1.5 Standardised test of comprehension 426

4.1.6 Survey of Attitudes to Reading 427

4.2 Pilot testing of test materials 429

4.2.1 Comprehension tests 430

4.2.2 Transfer tests 433

4.2.3 Reliability of tests 436

4.3 The main experiment 439

4.3.1 Subjects 439

4.3.2 Pretests 442

4.3.3 The teaching programme 444

4.3.4 Immediate posttests 451

4.3.5 Delayed posttests 453

5 Results and Initial Comment 454

5.1 Introduction 454

5.2 Pretest class comparisons 458

5.3 Overall effectiveness of training: comprehension tests 461

5.4 Overall effectiveness of the training programme: transfer tests . . 471

5.4.1 Error detection tests 472

5.4.2 Cloze tests 480
5.4.3 ACER Primary Reading Survey 482
5.5 Treatment-ability interactions 487
5.6 Variables which predict comprehension gains 492
5.7 Additional comments 503
5.7.1 Ethnic differences 503
5.7.2 Attitudes to reading 506

6 Discussion and Conclusions **510**
6.1 Introduction 510
6.2 Comprehension tests 512
6.3 Transfer tests 516
6.4 Discriminating students who improve from those who do not . . . 522
6.5 Suggested changes and further research 527

UNIVERSITY OF ADELAIDE LIBRARY

Introduction and Overview

The motivation for this study arose from my experience as a high school teacher and my observation that many students appeared to have difficulty understanding expository texts. For example, after reading a selection in class many students could not state succinctly what the text was about and some could not even recall a smattering of details. When researching projects, many students would copy verbatim from books, rather than stating the main points in their own words.

The intention in this study was to investigate a teaching programme to improve the ability to learn from texts among children in the upper primary grades, in preparation for work they will be expected to do in high school. The study was not intended as a remedial intervention for poor readers only, but rather was aimed at normal, intact classes containing students of varying reading ability levels. A further aim was to develop materials which could potentially be used by classroom teachers without extensive in-service training, since such training is impractical, especially in the case of rural teachers in Australia.

From the examination of theoretical and empirical work outlined below, a training programme was developed in which a central strategy involved the generation of topical inferences. In other words, students were taught to infer the main idea of each paragraph. This strategy was taught within a metacognitive framework to inform students of the value of this, and other strategies taught, and to encourage them to make self-directed use of the strategies.

Several researchers have confirmed the personal observation of students' reading problems noted above. For example, Berkowitz (1986) notes that it is a com-

mon observation among teachers and researchers alike that middle grade students have difficulty understanding and remembering what they read in content area textbooks, and B. Taylor (1986) adds that middle grade students form almost no mental gist after reading such material. Armbruster, Anderson, Armstrong, Wise and Janish (1991) report findings which show that U.S. high school students have difficulty learning from expository texts.

In spite of the difficulty and importance of being able to learn from text, traditional comprehension instruction did little to improve this ability. Durkin (1978-9) found that American primary school teachers spent very little time on direct instruction of comprehension. This is not to say that time was not devoted to reading, or even reading lessons. However, teachers tended not to give explanations of cognitive skills involved in comprehension, relying instead on repeated exposure to comprehension questions and skill exercises.

Dole, Duffy, Roehler and Pearson (1991) refer to repeated exposure to skill exercises as the drill-and-practice method, and Pearson (1984, p. 226) describes the rationale of the method as follows: *"The prevailing wisdom regarding comprehension instruction seems to be that if students get enough exposure to a skill or kind of question they will eventually improve at it."* According to Pearson, only students who are already moderately skilled are in a position to benefit from this practice. Those students who are lacking basic skills will remain ignorant and continue to make the same mistakes.

An extensive body of research indicates that important comprehension skills can be explicitly explained and demonstrated to children so as to produce improved understanding and retention of expository texts, as well as maintenance and transfer of the trained skills. Many skills could be chosen for such instruction,

since reading is a complex and multi-faceted process. Dole et al (1991) suggest three broad criteria for determining which skills should be taught:

1. the skills should be consistent with a cognitive model of reading;

2. they should differentiate good and poor readers;

3. they should be amenable to instruction.

These three questions provide the framework for organising the chapters to follow. More specifically, the first chapter will examine cognitive models of reading, the second chapter will examine relevant model-based differences between more and less advanced readers, and the third chapter will examine the extent to which important skills identified in the previous two chapters area amenable to instruction and succeed in improving comprehension. A training programme will be developed on the basis of this theoretical and empirical review.

The models of reading developed by Just and Carpenter (1980; 1987), Kintsch and van Dijk (1978) and van Dijk and Kintsch (1983) have been chosen as the central focus of chapter one because they offer a detailed analysis of comprehension skills within an overall model. The work of Just and Carpenter, based on the analysis of eye fixation data, includes discussion of lower-level processes, such as decoding and lexical access, and their relationship to overall processing efficiency and higher-level comprehension performance. On the other hand, the work of Kintsch and van Dijk was originally based on the analysis of recall protocols, and their model is particularly strong in the area of macrostructure formation. In the course of chapter one, these models will be discussed in relation to other theories and empirical findings. The intention is to present a detailed model of reading, firstly to isolate skills which could potentially be the basis of a training

programme, and secondly to provide a framework within which to examine, in subsequent chapters, why some students benefit from instruction while others do not.

One aspect of comprehension performance which is acknowledged but not developed within these models is that of metacognition. Thus, models of metacognition such as those developed by Jacobs and Paris (1987), Flavell (1981) and Borkowski and Kurtz (1987) will be examined and related to the models of reading.

On the basis of chapter one, a set of important skills and attributes within a cognitive model of reading will be isolated for further examination in chapter two, including:

1. processing efficiency and working memory capacity;
2. interclause integration, including the use of prior knowledge to construct inferences;
3. macrostructure formation, including the ability to identify or infer information central to the gist of the passage;
4. metacognitive characteristics, including goal setting, strategy selection, comprehension monitoring, and motivational characteristics such as attributional style.

An examination of differences between more and less skilled readers in these four areas will be the subject of chapter two. The empirical work reviewed in that chapter will show that better readers demonstrate superior skills in each of these four areas. However, it will be argued that the areas most profitably investigated

as a means to improve comprehension are prior knowledge activation and inference generation, macrostructure formation, and metacognitive knowledge and skills.

Training programmes focussed on these areas will be reviewed in chapter three to examine the degree to which these skills are amenable to instruction and training in the skills succeeds in improving comprehension. On the basis of this chapter it will be argued that training focussed on macrostructure formation is most likely to achieve the main aim of this study, namely, improving the ability to learn from expository text. Studies such as those of Stevens (1988), Kurtz and Borkowski (1987), Borkowski, Weyhing and Carr (1988) and Salomon, Globerson and Guterman (1989) show that, following training involving elements of macrostructure formation, students are more able to summarise the main idea of expository paragraphs, to answer main idea questions, and to perform well on standardised reading tests.

Apart from macrostrategies, an important aspect of learning from texts is being able to assess the state of one's comprehension (Baker, 1984b). Studies reviewed in chapter two will show that children (for example, Markman, 1979) and even adults (for example, Baker, 1985b) can claim to understand a text which contains explicit contradictions. As Epstein, Glenberg and Bradley (1984, p. 355) note: *"If a reader cannot detect a contradiction of a main idea (inserted into the text by an experimenter), then it is unlikely that the reader has understood the text as the author intended."* Therefore, as part of the training it is important to teach children to monitor their processing and to become aware of difficulties in constructing meaning, which could signal an unknown word, a misunderstanding, a lapse in concentration, or incomplete and poorly explained information in a text.

However, *what* is taught is only one aspect of a training programme; equally

important is *how* skills are taught. Firstly, the cognitive processes involved need to be explained to avoid the problems of traditional comprehension instruction mentioned previously. This approach has been referred to as the direct instruction model (for example, Book, Duffy, Roehler, Meloth and Vavrus, 1985), and involves explicit explanation and modelling, as well as guided and then independent practice with feedback.

In addition, though, training needs to include a metacognitive component if skills are to be maintained over time and transferred to other comprehension tasks. Paris and Oka (1986a, p. 104) note that: *"children must be convinced that the strategies are important, worth some extra effort, and instrumental for enhancing performance. These beliefs motivate students to use the strategies without prompting or help."* Thus, students need to be informed of the value and rationale for the strategies they are being taught, that is, when and why they are useful, to develop what Paris, Lipson and Wixson (1983) refer to as conditional knowledge. Students, in the course of training, also have to learn to regulate their own strategy use by gradually being given more and more responsibility for guiding and monitoring their own processing in the manner that has been modelled for them by the instructor. This amounts to the informed, self-control training which Brown, Campione and Day (1981) and Brown and Palincsar (1982) see as crucial to the maintenance and transfer of skills. Students who cannot carry out this executive function for themselves are unlikely to make unprompted use of the trained strategies.

The Reciprocal Teaching programme developed and tested by Palincsar (for example, Brown and Palincsar, 1982; Palincsar and Brown, 1984; Palincsar, 1987) incorporates the skills and methodology described above. As is set out in chap-

ter three, this programme has been very effective in bringing about a sustained improvement in reading comprehension. The present study builds on the work of Palincsar, but aims to extend it in the following ways:

1. All of her experimental work reported in detail in chapter three involved grade 7 students, although Palincsar and Brown (1986) note that children as young as grade 1 can benefit from the programme if texts are read aloud to them. Children in grade 5 were chosen for study because they are nearing the end of primary school when the ability to learn from expository texts is becoming more important, but at the same time they are less skilled than children in grade 7 in certain areas. For example, they are less able to rate the importance of different idea units in a text (Brown and Smiley, 1977; 1978). While grade 7 students are approaching an adult level in this skill, it is only beginning to develop in grade 5 children. Garner, Alexander, Slater, Hare, Smith and Reis (1986) show that grade 7 students are able to interrelate ideas within a paragraph, and to understand that unity of meaning is the basis of paragraphs, to a significantly greater extent than children in grade 5. Since the synthesis of summary statements from the separate idea units within a paragraph is an important component in the proposed training programme, one question to be addressed is whether grade 5 children can benefit from the programme to a similar extent as the grade 7 children in the Palincsar studies.

   Other studies teaching macrostrategies to grade 5 children suggest that their level of cognitive development is sufficient to be able to benefit from such training (for example, Adams et al, 1982; Taylor, 1982; Armbruster et al, 1987; Kurtz and Borkowski, 1987). However, only Brown and Palincsar

(1982) and Palincsar and Brown (1984) have investigated the maintenance of trained strategies over a period longer than one month, so a further aim will be to examine not only whether the children improve their comprehension, but also whether gains are maintained in the longer term, and whether new skills are transferred to other comprehension tasks.

2. In most of the Palincsar studies, students were selected for inclusion according to criteria of adequate decoding and poor comprehension. In some cases, a wider range of students was studied, for example, intact reading classes in Palincsar (1987) and intact science classes in Palincsar and Brown (1986). However, in the latter case, no information is given on the actual range of students involved, and in both cases only average gains are reported, so it is not clear how the intervention affected students of various ability levels. In Palincsar and Brown (1984), one student reading at close to grade level made impressive gains as a result of training. Hence, a second main question to be addressed is whether the proposed training programme will benefit children of varying levels of reading ability, as found in normal, intact primary school classes.

   On the basis of studies teaching elements of macrostructure formation to mixed ability classes (Taylor, 1982; Armbruster et al, 1987; Berkowitz, 1986; Baumann, 1984; Davey and McBride, 1986), it was anticipated that children of all reading ability levels in a normal classroom could improve their ability to learn from expository text following a composite training programme such as that used by Palincsar.

3. In Palincsar and Brown (1984), most of the students made impressive gains

in comprehension performance, although there was one student who was a notable exception. From the data given, there is no apparent reason why this student failed to benefit from training. A third question to be addressed is whether any factors can be isolated which influence the degree to which students benefit from the programme.

A discriminant analysis will be used to attempt to find a cluster of variables which discriminates those students who improve as a result of training from those who do not. Variables to be included in the analysis will be those identified as being important within a cognitive model of reading in chapter one, and shown to distinguish more and less advanced readers in chapter two. They include decoding speed and accuracy, vocabulary knowledge, subject knowledge, attitudes to reading, and a variety of comprehension pretest measures (comprehension questions, standardised reading test, cloze test, error detection test).

4. Groups of up to 16 children have been taught according to the Reciprocal Teaching method. However, if intact classes are being taught, they may contain 25 or more students. It is then a question of how the teacher can apply this method and maintain the engagement of such large numbers of children. Palincsar, Brown and Martin (1987) indicate that a class can be subdivided into groups and taught by peer tutoring. However, experienced teachers are needed to train the tutors, and this can be a problem. Palincsar (1987) notes that the teachers in this study were instructed in the method for one full day and two further sessions, but three of the six teachers still required quite extensive coaching when they applied the method to their

classes because they found it difficult.

A fourth question to be addressed, then, is whether teaching material can be produced in written form, as in the Informed Strategies for Learning package produced for teachers by Paris and Oka (1986b). Paris, Wasik and van der Westhuizen (1987, p. 38) comment that: *"Until interventions are more available in a format that is easily used by educators, experimental research is unlikely to be adopted in classroom settings."* The written materials produced for this study should make it easier for an instructor to plan and guide instruction without, however, leading to mechanical application, which would negate the metacognitive component. The aim is to provide a framework of the cognitive and metacognitive skills to be taught, while still allowing students to discuss their thinking and to gradually assume more responsibility for applying strategies.

A further question relates to whether students can master the skills being taught by individually writing responses and then discussing them as a class, rather than conducting an entirely oral dialogue, as in most of the Palincsar studies. Since each student is involved in responding to each segment of text, it should be easier to maintain the engagement of large groups. Palincsar (1985) reports only small gains in comprehension when children wrote responses and received very limited feedback. However, in Palincsar and Brown (1986), children in intact science classes wrote and then discussed responses. In this case, there was a marked improvement in the ability to comprehend the science texts studied during training, although transfer of training to novel texts and comprehension tasks was not assessed. Therefore, the use of the written format has not been fully

investigated.

Chapter four describes the design of test materials and their pilot testing prior to the main experiment. The chapter also describes the design and implementation of the teaching programme, including the strategies to be taught to the children.

Chapter five reports the results of the experiment, including the effect of training on immediate posttest comprehension tests, and comprehension tests administered after a delay of ten weeks. In addition, the effect of training on three transfer tests is assessed, these being a standardised reading test, a cloze test and an error detection test. Possible interactions between treatment and reading ability are investigated. This chapter also reports the results of a discriminant analysis to determine which combination of variables, if any, distinguishes those students who improve as a result of training from those who do not.

Chapter six presents the discussion and conclusions arising from the findings. Possible reasons for the findings are discussed in terms of the theoretical and empirical work reviewed in the first three chapters. Certain changes to the teaching programme are suggested, and possible questions for further research are outlined.

As indicated in chapter six, some weaknesses in the present training programme and its implementation became apparent. However, the outcome was sufficiently positive to suggest that the programme could provide the basis for materials to be made available to teachers. The teaching in this study was a modification of the successful Palincsar programme, which is an integrated package based on repeated practice of a small set of powerful strategies, rather than the short-term practice of a wide range of skills. These strategies have been shown to be effective in improving the understanding of expository texts, and they are soundly based on theoretical and empirical work. Equally important is the man-

ner of teaching, organised such that responsibility for applying the strategies is transferred gradually from the teacher to the students.

As noted previously, Paris et al (1987) consider that research findings will not be implemented in the classroom unless interventions are available in an easy to use format. Even when teachers are instructed on what they should do, they may still find it difficult to translate the instructions into appropriate materials and methodology, as found by Beck et al (1986). In the Beck et al study, teachers were taught how to recast prescribed basal text skills as strategies, and how to make explicit statements about the skills being taught. However, several teachers found the instructions difficult to implement and did not adhere to the intended procedure. In a study by Raphael and Wonnacott (1985) one group of teachers was provided with in-service training but no prepared materials. While the teachers were capable of generating their own materials, they noted that this activity was very time consuming and that it would be a great advantage to have teaching materials provided. If the present programme, with the suggested changes, is made available to teachers, it is a question for future research to determine if it can easily be used by teachers, and in the manner intended.

The ultimate aim would be to produce an effective package for use by teachers. Such a package, perhaps supported by videotapes modelling its implementation, could be particularly useful for teachers working in rural or isolated Australian schools, with little opportunity to attend in-service courses or to share ideas with other teachers.

UNIVERSITY OF ADELAIDE LIBRARY

Chapter 1

Theoretical Models of Reading

1.1 Introduction

As noted in the introductory chapter, Dole, Duffy, Roehler and Pearson (1991) propose as one criterion for a reading programme that it should be consistent with a cognitive view of reading. Hence, this chapter will examine models of reading, with the intention of developing a detailed theoretical framework within which to discuss approaches to comprehension instruction.

The model of reading developed by Just and Carpenter (1980) has been chosen as the central focus of this chapter for several reasons. It is a relatively complete model, incorporating all aspects of comprehension from word encoding through to the integration of ideas into larger units of meaning. As will be demonstrated, it is also well supported by empirical evidence. Where relevant, the Just and Carpenter position will be compared to other models; often, but not always, these positions are compatible. However, one aspect of the Just and Carpenter model is not well developed, and that is the formation of a macrostructure. Here

the work of van Dijk and Kintsch (1983) in particular will be used to show how skilled readers abstract main ideas to form an interpretation of what a text is about at a global level.

The first two main sections of this chapter will examine the general processing assumptions underlying the Just and Carpenter (1980) model and then describe aspects of processing from word encoding through to macrostructure formation.

The third main section will examine how a reader's knowledge influences the processing referred to above. In particular, the structure of the knowledge base required to support rapid and context-sensitive on-line processing will be discussed, together with empirical findings showing the effect of high or low subject knowledge on comprehension. In addition, theorising on metacognitive knowledge will be examined and related to the model of reading to show how reading goals, strategy selection and comprehension monitoring influence performance.

As a concluding comment, a complete model of reading is presented in this chapter, even though the central focus is on higher-order comprehension processes. The reason for doing so is that comprehension takes place within the constraints of the total processing system. Such an overall picture is necessary to examine which students benefit from a training programme, and to suggest reasons as to why some students benefit more or less than others.

1.2 General assumptions

Just and Carpenter (1980) based their model of reading on the analysis of gaze durations as university students read extracts from *Scientific American*. They noted systematic variations in gaze durations on different parts of the text. However, in order to be able to use eye tracking data to infer the processes carried out during reading, two assumptions must be made, which Just and Carpenter (1980, p. 330) refer to as the immediacy assumption and the eye-mind assumption.

> *"The first, called the immediacy assumption, is that a reader tries to interpret each content word of a text as it is encountered, even at the expense of making guesses that sometimes turn out to be wrong. Interpretation refers to processing at several levels such as encoding the word, choosing one meaning of it, assigning it to its referent, and determining its status in the sentence and in the discourse. The immediacy assumption posits that the interpretations at all levels of processing are not deferred; they occur as soon as possible.*
>
> *The second assumption, the eye-mind assumption, is that the eye remains fixated on a word as long as the word is being processed. So the time it takes to process a newly fixated word is directly indicated by the gaze duration."*

In terms of these two assumptions, the Just and Carpenter view contrasts with other models, such as that of Kintsch and van Dijk (1978), where a clause or short sentence is entered into a short-term memory buffer before being processed. However, in his later theorising (van Dijk and Kintsch, 1983; Kintsch, 1988), Kintsch too accepted that processing is carried out as soon as possible.

It should be noted that although Just and Carpenter refer to the 'immediacy' assumption, not all processing can be carried out immediately, for two reasons (Carpenter and Just, 1989). First of all, the text may withhold essential information, so processing must be postponed of necessity rather than by choice. Second, even when processing is initiated immediately, readers still use the context which follows a piece of text to update their initial interpretation. This updating is carried out as each new piece of information is encountered, but also at the end of the sentence. Thus, readers process information as soon as possible, but not necessarily immediately.

Within the Just and Carpenter (1980) model there are two mechanisms by which processing is carried out, namely, fast and automatic spreading activation and a largely serial production system. Spreading activation can operate in parallel with productions, but operates only on semantic associations. For example, encoding the word *dog* can lead to increased activation of closely associated concepts in the lexicon, such as *cat*. On the other hand, Carpenter and Just (1989, p. 35) point out that productions consist of condition-action pairs, where the action is carried out only when the specified condition is satisfied.

> *"...procedural knowledge consists of a set of autonomous units called productions. Each production consists of a condition-action contingency that specifies what mental action is to be taken when a given information condition arises in working memory. Typically, the mental action of a production changes the contents of working memory, thereby enabling another production. For example, encoding the word the might enable a production that would add to working memory the knowledge that a noun phrase has been encountered. The*

knowledge that a noun phrase has been encountered might, in turn, enable another production that relates the new noun phrase to the rest of the syntactic structure of the sentence being read."

All types of productions have access to information active in working memory. As one production fires, activation is transferred from the condition to the action element, which is then active in working memory and can serve as the condition for other productions. Although some processing within the production system will of necessity be serial because enabling conditions only gradually become available, there is also the possibility of productions firing in parallel. Information active in working memory can serve as the enabling condition for more than one production at the same time (Carpenter and Just, 1989).

Within this production system there is no need for a superordinate controlling programme to determine the order in which mental actions are to be carried out (Just and Carpenter, 1980). The system is self-sequencing, since each action taken potentially becomes an enabling condition for further productions to fire. As will be described below, the reader's goals can function as a type of control over processing. While these goals can influence the manner and depth to which a text is processed, they are not essential for sequencing productions.

Given the self-sequencing nature of the production system, there is no fixed order in which productions must fire (Just and Carpenter, 1980). For example, for the sentence *John pounded the nail with a ...*, the intermediate processing products active in working memory could trigger the production assigning the case role of instrument to the next word before the meaning of that word has been determined. Thus, the production system allows for interactions between stages of processing, that is, normally later stages can be carried out before,

and influence, normally earlier stages if their enabling conditions are present in working memory earlier than is normally the case.

All the actions specified within the production system are carried out in working memory, where the products of the action are also temporarily stored. The elements active in working memory are those whose activation level has reached a certain threshold; once this activation level declines, the element is lost from working memory. The nature of working memory is clearly an important aspect of the whole processing system, one which Daneman and Carpenter (1980, p. 450-1) describe as follows:

> *"Working memory is assumed to have processing as well as storage functions; it serves as the site for executing processes and for storing the products of these processes ... Information can become part of working memory through several routes; it may be perceptually encoded from the text; it may be sufficiently activated so that it is retrieved from long-term memory; finally, it may be the output of a comprehension process. Information can also be lost from working memory, since its capacity is assumed to be limited. Information may be lost through decay or displacement. Decay occurs if the activation of information subsides to a subthreshold level with time. Displacement occurs if additional structures are encoded, activated or constructed until the capacity is exceeded. An important aspect of information loss is the assumption that processes and structures compete for a shared limited capacity."*

Thus, within this view, working memory capacity cannot be measured in terms of chunks of static storage. Rather, storage capacity depends on how efficiently

the various processing operations are carried out. For example, if a person is reading about a very familiar topic, the concepts referred to and the productions needed to integrate them are both likely to have high base levels of activation due to frequency of use. There is, therefore, relatively little time and effort required to raise these elements to threshold level. In other words, processing of the text is efficient, leaving more remaining capacity for storage. As Carpenter and Just (1989) point out, working memory capacity is not static, and can only be measured in relation to particular processing demands in particular domains.

1.3 Stages of processing

From the analysis of their eye fixation data, Just and Carpenter (1980) identify five stages of processing, these being:

1. get next input
2. word encoding and lexical access
3. case role assignment (semantic analysis)
4. interclause integration
5. sentence wrap-up

Although the word 'stage' is used here, it does not necessarily imply a fixed sequence of actions, as has already been indicated. As was mentioned in the introduction, another aspect of processing will also be discussed, namely the formation of a macrostructure as conceptualised by van Dijk and Kintsch (1983).

1.3.1 Get next input

In most instances, the first stage in the interpretation of a text segment is to get the next input. In this phase, the eyes move one or two words to the right once the goal conditions set for the processing of the previous text have been met. Just and Carpenter (1980) refer to this movement of the eyes as a saccade, which takes about 30 msec to execute.

The reader's goals are one important determinant of how deeply a text is processed, and therefore how long a word is fixated before the next saccade.

Just and Carpenter (1980, p. 350) describe the importance of these goals in the following way:

> *"The reader's goals are perhaps the most important determinant of the reading process. A reader who skims a passage for the main point reads differently than someone who is trying to memorise a passage, or another person who is reading for entertainment. Goals can be represented in several aspects of the theory, but the main way is to require that each goal is satisfied or at least attempted before proceeding on to the next word, clause, or sentence."*

As the eyes move along a line of text, most words are fixated at least once. However, some short function words are not fixated. Just and Carpenter (1987) suggest that these words are still processed in the following way. In a phrase such as *it is possible to*, when the end of the word *possible* is fixated, the word *to* is still in foveal vision. Since *to* is a short word, it can be encoded without a further fixation. Since it is a common word, it can be processed with little effort in parallel with *possible*. However, if the word following *possible* were longer or less familiar, an additional fixation would be required.

1.3.2 Word encoding and lexical access

Although Just and Carpenter (1980) present word encoding and lexical access as one stage within their model, the two processes can be separated. Indeed, Just and Carpenter (1987) clearly make such a separation. In general terms, they consider that the visual encoding process results in what they call a word percept, that is, an internal representation of the visual characteristics of a word, but not its meaning. Once a word percept has been encoded, it activates its associated word concept(s), that is, the meaning(s) associated with that percept. This latter process is often referred to as lexical access.

Word encoding

Just and Carpenter (1987) assume that each letter is encoded in turn in the process of forming a word percept. They propose a production to form word percepts as follows: *If you see a letter in a given serial position, then increase the activation level of word percepts that contain that letter in that position.* So, for example, the first letter in the word *man* would activate word percepts such as *man, met, mob, mat*, and so on. The second letter would activate its own set of word percepts with a central *a*. Some words in the first set, such as *man* and *mat* would receive additional activation during this second process. Only *man* would receive such additional activation again as the third letter activates word percepts.

From the regression analysis of their eye fixation data, Just and Carpenter (1980) report that each additional syllable added an average of 52 msec to the gaze duration on a word. In relation to individual letters, each additional letter in a word increased gaze durations by an average of 30 msec in Just and Carpenter

(1980) and 40 msec in Carpenter and Daneman (1981).

Some other studies also indicate an increase in word encoding time with increasing word length. For example, Gough (1972) found that pronunciation was initiated 78 msec faster for 3-letter than for 10-letter words. The process of verifying whether a letter string constituted a real word took 35 msec longer for 6-letter than for 4-letter words. Using eye fixation data, Haberlandt and Graesser (1985) report that gaze durations increased by an average of 15 msec for each additional letter in a word.

However, if word encoding could proceed only in a serial, letter-by-letter manner, as proposed by Gough (1972), it would be impossible to account for contextual effects on word recognition. Just and Carpenter (1987) acknowledge that such effects do occur. At a localised level, the speed of recognising clusters of letters is influenced by the familiarity of the cluster. For example, words can be vocalised faster than pseudowords (Perfetti and Hogaboam, 1975), and letters can be perceived more accurately in orthographically regular strings compared to strings which violate English orthography (Rumelhart and McClelland, 1981). Rumelhart and McClelland also indicate that high frequency words are perceived more accurately at a given exposure level than low frequency words, and that high frequency words can be perceived at lower levels of exposure. Findings such as these led Rumelhart and McClelland (1981, p. 39) to propose a more interactive rather than strictly serial process of word percept formation:

> *"...the process of letter identification need not be complete, exhaustive or accessible to report before partial results of letter processing begin to interact with our knowledge about words. Perhaps what word knowledge does is allow us to reinforce partial, preattentive activations*

of letters that are consistent with words in our vocabulary. Suppose the presentation of a display produces activations of letters consistent with the display and that these in turn produce activations of words consistent with the letters. These activations, in turn, produce feedback to the letters, reinforcing the activations of letters consistent with the activated word."

Rumelhart and McClelland envisage a system where each unit, such as a letter or a word, is represented by a node which is connected to other nodes. Within this model, and that of Just and Carpenter (1980), all nodes have a particular resting level of activation. What this means is that a frequently accessed node, on the basis of being activated many times in the past, requires only a small amount of activation to bring it to the threshold of recognition, while a less frequently used node requires a larger amount of activation, since its resting level is lower. When one node is activated during the encoding of letter features, this activation spreads to its neighbours. Activation can move in both directions, for example, from letters to words or words to letters. Very familiar words will more rapidly accumulate sufficient activation both to send feedback to the letter level and to reach recognition threshold. This model also explains why letter clusters are recognised more quickly when they form words rather than pseudowords, since more feedback flows back to the letter level from recognisable words.

The wider semantic context in which a word occurs can also influence the speed of its encoding. A word in a highly constrained context is fixated for a shorter time than the same word in a less constrained context (McConkie and Zola, 1981). A target noun is named more quickly in a congruous sentence than in a neutral, uninformative context (Stanovich and West, 1983). Similarly, a target

noun (for example, *floor*) is named more quickly when preceded by a restrictive verb (*swept*) rather than a general verb (*cleaned*) (Schustack, Ehrlich and Rayner, 1987). Several studies have also shown that a word is vocalised faster when it occurs in a story rather than in isolation (Biemiller, 1977/8; West and Stanovich, 1978; Perfetti, Goldman and Hogaboam, 1979).

In order to accommodate these wider contextual effects, the Rumelhart and McClelland (1981) model described above needs to be slightly extended to allow activation to flow not only from the word percept to the letter level, but also from the word concept level. When a target word is preceded by a constraining context, processing of that context is likely to activate associated concepts, as in the case of *swept - floor*. Once letter-level analysis is initiated, but prior to its completion, activation can flow from concepts to the word percept and letter levels, reinforcing partially complete analyses and speeding the encoding process.

However, it would appear that such facilitation takes times. Context has a larger effect on the encoding of difficult (for example, *avalanche*) than easy (*snow*) words (Stanovich and West, 1983), and a larger effect on longer and low frequency words (Perfetti, Goldman and Hogaboam, 1979). It may be that when the encoding of a word is rapid, due to relatively simple visual analysis and feedback from a familiar word percept with high base level activation, there is not sufficient time for these wider contextual effects to influence the encoding process. Such speculation is supported by the finding that poor readers receive relatively greater facilitation from context than good readers (for example, Perfetti, Goldman and Hogaboam, 1979; Perfetti and Roth, 1981). Among the poor readers, visual analysis of graphic information is relatively slow, hence contextual facilitation has time to influence encoding to a greater extent than among good

readers.

So far, evidence has been presented to support Just and Carpenter's assumption of letter-by-letter processing in the formation of word percepts, but allowing for contextual effects to influence this process. One more point will be discussed before concluding.

LaBerge and Samuels (1974) propose that skilled readers do not necessarily have to analyse letters to form word percepts. With practice and many exposures to particular words, graphic input can directly and automatically activate word percepts without prior analysis at the lower level. LaBerge and Samuels refer to this holistic recognition of words as 'unitising'.

In support of this proposal, Terry, Samuels and LaBerge (1976) showed that word recognition latencies for adults were not affected by word length under normal viewing conditions, suggesting holistic recognition. However, when these same high frequency words, three to six letters in length, were shown in mirror image, latencies increased with word length, suggesting that these unfamiliar conditions forced readers to use letter-by-letter analysis.

Samuels, Miller and Eisenberg (1979) trained adults to recognise one set of words in mirror image over several sessions. Other words were shown in only one of the sessions. While word length influenced recognition latencies for both sets of words, the effect declined over the course of the experiment for repeated words, but not for new words. The authors conclude that with the repeated words, subjects were learning to recognise units larger than single letters.

Samuels, LaBerge and Bremer (1978) also suggest that unitising is the result of extensive practice. Word length did not affect the word recognition latencies of college students, while there was an almost linear increase with word length

among grade 2 students. Latencies for children in grades 4 and 6 were between these two, suggesting that the children were not yet recognising words holistically like the adults.

It should be noted, though, that the words used in these studies were relatively short, high frequency words, likely to be very familiar to the adult subjects. The context in which Just and Carpenter (1980) found increases in gaze durations due to increases in word length were rather different; subjects were reading extracts from *Scientific American*, hence words were likely to be more diverse, longer, and of lower familiarity and normative frequency. While it remains a possibility that the recognition of common words can be unitised, it is unlikely that this process can play an important role in more complex reading tasks, a view which is expressed by Perfetti (1984, p. 46) as follows:

> *"It has become clear that simple ideas such as 'holistic' word identification are inadequate. The identification of words is mediated by the perception of letters. Cues of word shape and word length are of some significance, but clearly carry only a small portion of the burden of word identification compared to letter perception."*

This proposal receives support from the finding that poor readers are more affected than good readers by print degradation, which makes letter identification more difficult, but does not change word shape or length (Perfetti and Roth, 1981). If skilled readers are distinguished by the ability to recognise letters rapidly and on the basis of very few featural cues, then it follows that they would be less affected by print degradation than the less skilled reader, who requires more cues to identify graphic information.

In conclusion, the evidence presented in this section supports the Just and Carpenter view that word percepts are formed largely on the basis of letter identification, although some contextual facilitation can occur. While the model is interactive, it is weakly interactive, similar to the views of Perfetti and Roth (1981, p. 27) when they propose:

> *"... a model of reading whose processes are at once interactive and asymmetrical. Top-down and bottom-up data are not used in strictly reciprocal ways. An important consequence of this asymmetry is that so-called bottom-up processes can carry on reasonably well without top-down processes, but not vice versa. No matter how helpful top-down processes are, they are neither definitive nor essential."*

In order to have sufficient working memory capacity to carry out the integrative processes essential to the formation of a coherent text representation, readers must clearly have good visual analysis skills according to the model of reading set out in this section.

Lexical access

Once a word percept has been encoded, it activates its associated word concept(s), that is, the meaning(s) associated with that percept. Following the model of Collins and Loftus (1975), these concepts are represented as nodes connected by labelled links to form a semantic network. Just and Carpenter (1980, p. 338) describe this system as follows:

> *"Perceptual encoding productions use the visual features as conditions; their action is to activate the representation of the word. Once the representation of the word has been sufficiently activated, its corresponding concept is accessed and inserted into working memory. The concept serves as a pointer to a more complete representation of the meaning, which consists of a small semantic network realised as a set of productions. The major nodes of the network are the possible meanings of the word, the semantic and syntactic properties of the meanings, and information about the contexts in which they usually occur."*

By allowing a concept to serve as a pointer to a small semantic network, Just and Carpenter are acknowledging that conceptual meanings do not consist of single definitions, but arise from the relationships between concepts, a point which is emphasised by Till, Mross and Kintsch (1988, p. 284):

> *"It is probably not the case that fully developed word meanings reside somewhere in the lexicon waiting to be selected and inserted into a discourse representation. Rather, the full contextual meaning of a word is likely to be constructed in the specific discourse context.*

Only such a generative contextual view of word meaning can do justice to the enormous flexibility and adaptability of human language."

Each concept in memory is assumed to have a particular base level of activation corresponding to its frequency of use (Just and Carpenter, 1980). Thus, a concept which is frequently accessed will have a relatively high base level, and therefore will require little additional activation to reach threshold, at which point it is active in working memory and potentially available for integration into the text representation being constructed.

The normal base level of a concept can be raised in two ways. The first of these is repetition. When a concept is accessed, its activation level is increased but then gradually decays. However, the level stabilises slightly above the original base level. The second way is through contextual facilitation. When a concept is accessed, activation spreads to other, closely associated concepts, access to which is thus facilitated because of their increased activation level.

Just and Carpenter (1980) use the analysis of their eye fixation data to support these assumptions. They report that gaze durations on words increased by an average of 53msec for each log unit decrease in relative word frequency, supporting the notion of differential base levels of activation. Gaze durations were on average 49 msec shorter on the second appearance of a word than on the first, supporting the notion of increased base level activation following access to a concept.

As an example of contextual facilitation, Schustack, Ehrlich and Rayner (1987) found that gaze durations were shorter on a word such as *picture* when it was preceded by a specific verb such as *hung*, rather than a general verb such as *put*. Within the Just and Carpenter framework, activation spreads from a specific verb to a small set of possible objects. Although *picture* does not reach threshold as

a result of this priming, subsequent access to the concept is facilitated. No such facilitation occurs with the much more general verb *put*. The number of concepts potentially associated with such a general verb is too large for any one of them to receive substantial activation.

Just and Carpenter (1987) note that contextual facilitation should not be confused with conscious prediction. Concepts activated by spreading activation rarely reach threshold, and therefore do not reach consciousness or become part of the text representation unless they receive further activation from other sources.

The factors influencing lexical access discussed so far indicate why it is that immediate interpretation is usually correct. Interpretation is biased in favour of commonly used meanings, repeated meanings, and meanings primed by the context. These same factors also influence the interpretation of ambiguous words. In some cases, the two meanings of an ambiguous word have a similar frequency of occurrence, hence the two word concepts have similar base levels of activation. Ambiguous words in this category are referred to as 'equibiased'. In such cases, a word percept is likely to activate both word concepts associated with it. For example, the word *mint* activated both the meaning of a candy and a building where money is produced (Till, Mross and Kintsch, 1988). However, within 400 msec one meaning was selected and the other deactivated. Similarly, Kintsch and Mross (1985) show that when equibiased words such as *iron* were read in a sentence, both meanings were still activated 150 msec after presentation of the word. However, one meaning had been selected by the time readers moved on to the next word.

While in some cases more than one meaning is activated for an ambiguous word, there are also cases where only one meaning is activated. Only one word

concept is likely to be activated when either one meaning of an equibiased word is strongly primed by the preceding context, or when a word is non-equibiased, that is, one meaning occurs more frequently and hence its word concept has a higher base level of activation. Carpenter and Daneman (1981, p. 1478) summarise this point as follows: *"It (their model) postulates that whether one or more meanings are retrieved depends upon the context and the relative frequencies of the interpretations."* Rayner and Frazier (1989) make a similar point when they propose that integration processes operate on the first word concept that reaches a threshold of activation which, as indicated, is influenced by both the relative frequency of the two word concepts and their degree of contextual priming.

Duffy, Morris and Rayner (1988) empirically demonstrate this point by analysing gaze durations on ambiguous words in different contexts. In the case of equibiased words, only one meaning was activated when a disambiguating context preceded the word and one meaning was primed by this context. In the case of non-equibiased words, on the other hand, only one meaning was activated when the context followed the word and context did not alter the differential in base level activation of the two word concepts. Similarly, Seidenberg, Tanenhaus, Leiman and Bienkowski (1982) found that both meanings of an equibiased word such as *straw* were activated in a neutral context, but only one meaning was activated in the context of a farmer.

While both frequency of occurrence and conceptual priming influence word concept activation, Carpenter and Daneman (1981) show that relative frequency effects are generally stronger than priming effects. These authors used non-equibiased homographs, that is, words with the same spelling, but different pronunciations and meanings. When primed by a preceding context, the higher

frequency meaning of ambiguous words was pronounced during oral reading on 87% of occasions, compared to only 38% of occasions for the lower frequency meanings when they were primed. Some very low frequency meanings were never pronounced in spite of contextual priming, although the selection of these meanings could be increased to some extent by strengthening the degree of priming. Tabossi (1988a) also indicates that a highly constraining context is required to override the effect of frequency of occurrence. In this study, one meaning of an equibiased word was activated, rather than both meanings, only when the context was highly constraining.

In the case of any word, whether ambiguous or not, there is evidence to show that contextually relevant attributes of a concept are activated, even when these attributes are not explicitly mentioned in the text. For example, Barclay, Bransford, Franks, McCarrell and Nitsch (1974) found that *heavy* was encoded as an attribute of *piano* in the context of lifting, while *nice sounding* was encoded in the context of tuning. Tabossi (1988b) showed that *fat* was activated as an attribute of *butter* in the context of a diet, but not in the context of melting the butter. Tabossi and Johnson-Laird (1980) demonstrated that *hard* was activated as an attribute of *diamond* in the context of cutting glass, but not in the context of dispersing light. In a study by McKoon and Ratcliff (1988), *red* was activated as a characteristic of *tomato* in the context of a still life painting, while *round* was activated in the context of rolling the tomato along the floor. These authors also showed, by testing after every three paragraphs and hence delaying the test, that contextually relevant aspects of meaning were not only activated, but also encoded into the memory representation.

Just and Carpenter (1987) acknowledge that such elaborations can occur, but are not obligatory; their production will depend on the depth of processing undertaken by the reader, which in turn depends on the processing capacity available and the reader's goals. While word concepts are activated and integrated immediately, elaborations such as those described above may not be available for integration into the text representation immediately. Till, Mross and Kintsch (1988) suggest that elaborated meanings take longer to reach a threshold of activation than the word concepts themselves, and may not be available for integration until the sentence wrap-up operations. These authors presented sentences containing an ambiguous word, as for example: *"The townspeople were amazed to find that all the buildings had collapsed except the mint"*. It took up to 1000 msec after the presentation of the final word for elaborations to become available, in this case the topic inference of an earthquake. Of this elaboration, Till et al (1988, p. 293) note that *"subjects make elaborative inferences if they have enough time to do so. The important finding is that this kind of inferencing is not immediate but appears to be part of the sentence wrap-up processing."*

It should be noted, though, that Till et al are discussing topical inferences, that is, inferences regarding the overall topic of the sentence, which is more general than elaborating the meaning of the word *mint*. These authors did not actually investigate whether the meaning of *mint* was elaborated as *a strong, solid building*, that is, a building strong enough to withstand an earthquake. Therefore, what Till et al refer to as word sense elaborations are of a more general nature than the inferred conceptual attributes described in the preceding paragraphs.

On a similar point, Just and Carpenter (1980) note that although the meanings of a word are accessed while that word is being fixated, this initial meaning

may need to be refined following the integration of subsequent words, or during the sentence wrap-up phase. For example, the general meaning of the word *large* is familiar and readily accessed, but its precise meaning depends on its context. The concept of *large* takes on different shades of meaning when it is associated with *insect* as opposed to *building*, so the careful reader updates its meaning after processing the associated noun.

Just and Carpenter (1987, p. 79) note that children and poor readers take longer to access word meanings, but that the contribution which speed of lexical access makes to comprehension difficulties is not clear.

> *"Some theorists believe that the lexical access process is a major source of individual differences in reading comprehension performance. If readers take a long time to access word meanings, then other processes will be allotted less time, and storage limitations may cause the loss of some information from other levels before it can be used. There is fairly convincing evidence that slowness in lexical access is a contributing factor to poor reading comprehension, but the size of the contribution is currently unclear."*

Although the quantitative contribution of lexical access to comprehension problems may not be clear, there are several ways in which access to a well developed network of word concepts may be important. Most obviously, word concepts have to be available to construct meaning from a text; comprehension will be impaired, or at least slowed, when a text contains unfamiliar words. Secondly, if the available word concepts are well practised, they will be readily accessible and require a minimum amount of processing time, as indicated above by Just

and Carpenter (1987). Finally, if word concepts in memory are embedded in a network of associated concepts, readers can not only benefit from the facilitative effects of contextual priming through spreading activation, but can also construct a richer text representation through the elaborated meanings generated from this network.

A study by Stahl, Hare, Sinatra and Gregory (1991) illustrates the importance of lexical access in terms of breadth of vocabulary knowledge rather than speed of access per se. Subjects in grade 10 read a lengthy passage describing the celebration of a baseball player's retirement. In a regression analysis of the number of idea units recalled, vocabulary knowledge explained a significant 15.1% of variance when entered before baseball knowledge, and 13.4% when entered after baseball knowledge. Baseball knowledge did not explain a significant percentage of variance in total recall, although it made a significant contribution to the explanation of variance in recall of the retiring player's career achievements, leading Stahl et al (1991, p. 504) to conclude that: *"... domain knowledge and vocabulary knowledge affect what is comprehended as well as how much. A text may be difficult for a reader with low vocabulary knowledge but high domain knowledge if the comprehension task involves microprocessing, but not if the reader need only get a general idea of what the text is about."*

A study by Beck, Perfetti and McKeown (1982) illustrates how expanded vocabulary knowledge and increased speed of access can improve comprehension. Children in grade 4 were taught to pronounce and define words in several semantic groups, such as moods, eating, and types of people. To increase speed of access, the children engaged in competitions to match words and definitions with maximum speed and accuracy, and competed against the clock to establish

personal best times for this matching. To develop an interconnected memory representation, the children constructed relationships between the new words, for example, by explaining how a *crook* and an *accomplice* fitted together, whether a *hermit* could be an *accomplice*, whether an *accomplice* would rob a bank alone, and so on. The achievements of this programme, including expanded vocabulary knowledge, speed of access to this knowledge, and its representation in an interconnected network, succeeded in improving reading comprehension. This success suggests the important features a vocabulary instruction programme should have, but also supports the model of reading presented in this section, as well as the hypothesised structure of the knowledge base to be presented later in the chapter.

Recovery from inconsistent interpretations

By processing each word as it is encountered, the reader greatly reduces the burden which would be placed on working memory if all processing were carried out at the end of the clause or sentence. Initial interpretations are usually correct due to the use of converging evidence. However, there are times when these interpretations prove to be incorrect in the light of subsequent information, and so the reader needs to have available error recovery heuristics to determine a more appropriate interpretation.

To investigate error recovery heuristics, Carpenter and Daneman (1981) wrote short passages which primed appropriate or inappropriate meanings of ambiguous words. For example, the homograph *bass*, meaning either a fish or a low pitch in music, was preceded by a description of a lakeside resort which was likely to prime the 'fish' meaning. In one version of the passage, the resort was the venue for a fishing competition, and hence the primed meaning was appropriate, while

in the other version the resort was the venue for a rock concert, where performers included a bass guitarist.

Carpenter and Daneman used eye fixation data to analyse readers' detection of and recovery from inconsistent interpretations of ambiguous words. Gaze durations on the first disambiguating word, for example, bass *guitarist*, were longer when the word was inconsistent with the reader's initial oral reading. If the disambiguating information was contained in a whole phrase, gaze durations on the whole phrase were longer when it contradicted the initial reading of the ambiguous word. These longer gaze durations suggest that readers interpret and attempt to integrate words immediately, but have difficulty with the integration in passages priming inconsistent meanings.

Readers were over four times more likely to show regressive eye movements if there was inconsistency between their initial oral reading and disambiguating information, suggesting that they were attempting to recover from the error which had been detected. Regressions were mostly initiated after fixating the disambiguating word, but some regressions were initiated at the end of the disambiguating phrase, the end of the sentence, or even the end of the passage.

Based on this eye fixation data and oral corrections, Carpenter and Daneman (1981) conclude that recovery processes are often initiated as soon as inconsistent information is encountered. However, sometimes readers read on and use the end of the sentence or even the end of the passage to resolve any remaining inconsistencies. The authors infer three main error recovery heuristics from their data, namely, try to reinterpret the first word that appears inconsistent, check previous words that caused processing difficulty, and read on in the hope that subsequent information will explicate that apparent inconsistency. When the level of incon-

sistency is not too great, some readers elaborate disambiguating information to make it consistent with the their initial interpretation of the ambiguous word. In this case, however, they have changed the writer's intended meaning.

Although there was clearly evidence that readers detected and attempted to recover from inconsistent interpretations, not all readers were successful. Carpenter and Daneman (1981) asked probe questions immediately after each passage to determine readers' final interpretations of the ambiguous words. In one passage, for example, the target sentence read *It was the lead in the inquiry that enabled him to solve the mystery.* The probe question asked *What enabled the detective to solve the mystery?.*

On the basis of oral corrections after the first reading of the ambiguous words, readers recovered from initially incorrect interpretations in 36% of instances. In a further 22% of instances readers showed, by correctly answering probe questions, that they had recovered from the inconsistencies by the end of the passage. However, in 19% of instances readers gave ambiguous answers which did not indicate clearly the nature of their final interpretation, and in 23% of instances the inconsistencies had clearly not been resolved. Thus, even the university students who served as subjects in this study unequivocally recovered from errors in only 58% of instances.

The failure of almost half of these tertiary students to recover from inconsistent interpretations could have detrimental effects on their comprehension. On the one hand, they may have failed to construct a coherent representation of the text. On the other hand, they may have resolved the difficulty in a manner not intended by the writer, which may also pose a problem when the task involves learning from expository texts. Efficient comprehension monitoring and error

recovery, therefore, are important skills for readers to develop.

1.3.3 Case role assignment (semantic analysis)

The third processing stage is referred to by Just and Carpenter (1980) as case role assignment, although Just and Carpenter (1987) discuss the same processes under the heading of semantic analysis. This stage involves the first of the integrative phases, namely, determining the conceptual relations between words in a clause and assigning roles such as agent, recipient, object, location, time, instrument, manner, action and state. Therefore, this stage can also be referred to as intraclause integration, resulting in the formation of individual propositions.

Many authors, including Just and Carpenter, assume that the clause is an important processing unit, that is, information is chunked into clausal units during processing. For example, Haberlandt and Graesser (1985) report reliable increases in reading time at the end of clauses, suggesting that integrative processing is carried out at clause boundaries. Chang (1980) and Clark and Sengul (1979) report that the clause immediately prior to words currently being processed is still active in working memory, also suggesting that information is chunked into processing units at clause boundaries.

Just and Carpenter (1987) give an example of the sentence *John killed the snake with a knife on the verandah* analysed into its constituent roles and expressed in simple propositional form: (Action:IS kill), (Agent: IS John), (Instrument: IS knife), (Object: IS snake), (Location: IS verandah). This analysis is independent of the grammatical form in which the sentence was expressed. The case roles would be the same if the sentence had been written in passive form, since case roles attempt to capture conceptual rather than grammatical relations.

Just and Carpenter (1987) follow van Dijk and Kintsch (1983) in suggesting that concepts are organised into complex propositions showing semantic relation-

ships. Thus, the sentence *John inadvertently gave the old book to Peter in the library* would be represented as follows:

PREDICATE:	give (a,b,c)	
	Modifier:	inadvertently
ARGUMENTS:		
Agent:	John = a	
Object:	book = b	
	Modifier:	old (b)
Goal:	Peter = c	
CIRCUMSTANCES:		
Time:	yesterday	
Place:	in the library	

For on-line semantic analysis to be able to take place, it has to proceed in a strategic fashion, based on incomplete information. Van Dijk and Kintsch (1983, p. 14) suggest that a propositional scheme is used in this strategic fashion:

> *"Such a schema is a strategical unit: It allows a fast analysis of surface structures into a relatively fixed and simple configuration. Thus, a first occurrence grammatical subject noun, or a pronoun, when denoting a person, will be strategically assigned the agent position in such a schema, even before the rest of the clause has been analysed."*

Like Just and Carpenter (1987), van Dijk and Kintsch assume that propositions are constructed strategically from incoming information. The latter authors do not discuss these strategies in detail, but they do note that surface structure

features such as word order, morphological structure and syntactic categories are useful cues for filling the slots of the propositional schema. Just and Carpenter (1987) give examples of cues used by readers to assign case roles, including word order, type of verb, particles and prepositions.

Word order may be a useful cue. In a sentence such as *The postman gave the woman the dog*, the expected order of participants would be agent, goal, object. However, this strategy would not work in a passive construction, or in the permissible reordering *The postman gave the dog to the woman.*

The type of verb used is itself a cue to the number and nature of roles likely to be involved. For example, a verb like *give* has three participants associated with it, while a verb like *die* has only one. However, since the verb is not normally placed at the beginning of the sentence, this type of information only becomes useful later in the analysis of a clause.

Particles can be a useful cue to case role assignment. Just and Carpenter (1987) give the following examples to show how a particle changes the case role assigned to *hammer*:

Wayne put the picture with the hammer. (location)

Wayne put up the picture with the hammer. (instrument)

Wayne put up with the picture with the hammer. (modifier)

Some prepositions can provide useful cues. For example, *above* can be associated with only very few case roles and therefore provides useful cues. However, some words have functions in addition to being prepositions, making the role of the participant which follows less predictable. For example, *before* can be a preposition followed by a location, as in *He stood before the judge*, or a conjunction followed by the agent of a new clause, as in *He stood before the judge spoke.*

None of the strategies above is sufficient on its own to correctly assign case roles. A further source of information, namely, general knowledge of the referents, is required to efficiently carry out the process. Just and Carpenter (1987) nicely illustrate this point with the examples: *Jake ate the sausage with a friend/with a fork.* General knowledge dictates that *friend* must fulfill the role of co-agent rather than instrument, while *fork* is interpreted as an instrument. None of the other cues mentioned so far could determine the case roles in these examples.

Just and Carpenter (1987, p. 176) summarise the use of cues to carry out semantic analysis as follows:

> *"The semantic analysis processes that use these cues must operate according to several principles. First, most semantic role assignments have multiple cues, so information from several cues generally must be combined, and their information value must be weighed in proportion to the reliability of the cue. Second, semantic analysis is performed as each word of a clause is encountered by the reader, so the order in which the cues are encountered will be a major determinant of the semantic analysis process. Third, many extrasentential information sources - such as the reader's knowledge of actions and objects in the world and her knowledge of the preceding text - will be brought to bear on the semantic analysis process."*

Case roles are assigned within clauses, so the reader must have ways of determining clause boundaries. Just and Carpenter (1987) suggest that syntactic analysis is a useful tool for chunking the information presented sequentially in a sentence. They put forward a number of strategies for the use of syntactic cues, including the use of function words, affixes and punctuation.

Readers may use function words to determine where a new constituent begins. For example, a determiner such as *the* is a cue to begin a new noun phrase, a conjunction such as *because* is a cue to begin a new clause, and a preposition such as *with* is a cue to begin a new phrase. Along similar lines, van Dijk and Kintsch (1983) in their model incorporate the parsing strategies suggested by Clark and Clark (1977). As an example, in order to segment sentences into constituent parts Clark and Clark (1977, p. 59) propose the following general strategy: *"Whenever you find a function word, begin a constituent larger than one word."* More specific subcategories of this strategy suggest, for example, to begin a new noun phrase whenever a determiner or quantifier is found, or to begin a new clause whenever a relative pronoun or subordinating conjunction is found. However, the same word can fulfill different syntactic functions. For example, *that* is a determiner in *that yellow windmill*, but a relative pronoun in *the money that she earned.* So, while the function words are useful cues to chunking words into constituents, the process is made more difficult by the fact that the same function word can play different syntactic roles.

Affixes can be useful cues to the syntactic function of a particular word. For example, words ending in *-ly* are likely to be adverbs, those ending in *-ed* are likely to be verbs, those ending in *-ive* are likely to be adjectives, and endings such as *-tion, -ness, -ment* are likely to signal nouns. However, the same ending can also be found on words from different classes, as for example in *larger* (adjective), *writer* (noun) and *deter* (verb).

Punctuation is a more unambiguous cue to syntactic analysis. Commas and full stops are cues to wrap up the current clause and to start a new constituent.

However, syntactic cues alone cannot determine the function of words in a

sentence and how they should be chunked. For example, in the sentence *I saw the dog with three legs*, it is the reader's general knowledge which indicates that the final phrase modifies the noun *dog*, while in the sentence *I saw the dog with binoculars* the final phrase modifies the verb *saw*.

Just and Carpenter (1987) propose that syntactic analysis is carried out using multiple converging cues, including semantic ones. The same cue can serve as the enabling condition for more than one production to fire in parallel, and a variety of cues can trigger productions which add activation to one particular interpretation, as Just and Carpenter (1987, p. 154) explain:

> *"Each production increments the activation level of the interpretation it thinks is the correct one, and the interpretation whose activation level reaches a standard threshold first becomes the accepted interpretation. Collaboration occurs when two or more processes converge on establishing a given interpretation by jointly incrementing its activation level to the standard threshold when neither process alone may have effectively done so. For example, syntactic and semantic productions can collaborate on establishing the nature of a noun phrase. The same mechanism (conjoint incrementation of an activation level) also integrates information from different sources within the syntactic level, as occurs when several different syntactic cues converge on the same interpretation."*

At a general level, the syntactic strategies form as few constituents as possible. In other words, the reader attempts to incorporate each successive word into the preceding clause until the words no longer fit. For example, in the sentence *Mary*

dropped the cup and the saucer accidentally landed on the rug, a first parsing is likely to interpret *the cup and the saucer* as the object of dropped, and *accidentally* as a modifier of *dropped.* However, when the second verb *landed* is encountered, *the saucer* has to be reassigned to start a new clause, and *accidentally* modifies *landed* (Just and Carpenter, 1987). Similarly, in the sentence *When farmers are ploughing the soil is most fertile*, the clause boundary is first likely to be placed after *soil*, until the second verb *is* suggests that the boundary should be placed after *ploughing* (Just and Carpenter, 1980).

The above aspects of the Just and Carpenter model are supported by a number of studies using the eye tracking methodology. For example, Frazier and Rayner (1982) hypothesised that readers try to attach each incoming word to the clause currently being processed. The authors constructed sentences where this parsing strategy was appropriate, as well as slightly altered versions of these sentences where it was not appropriate. For example, in the sentence: *Since Jay always jogs a mile (this) seems like a short distance to him*, the strategy is appropriate when *this* is included and *a mile* can be included in the first clause, but not appropriate when *this* is deleted and *a mile* has to be reparsed into the second clause. In this latter case, reading times per character were longer, as predicted, than in the former case, suggesting that readers used the hypothesised strategy and then had to reparse. Analysis of gaze durations suggested that readers immediately detected a parsing problem when one occurred. Recovery from incorrect parsing did not necessarily involve regressive eye movements, but where these did occur, the most common pattern was to move directly from the disambiguating region (*seems*) or the end of the sentence to the ambiguous region (*a mile*). Thus, the reader's mental representation of the sentence was involved in recovery, either as

the basis for reinterpretation without regressions, or by directing the eyes to the region requiring reanalysis.

Van Dijk and Kintsch (1983) again refer to the work of Clark and Clark (1977) with regard to possible parsing strategies. Like Frazier and Rayner (1982), Clark and Clark (1977, p. 65) suggest as one strategy: *"Try to attach each new word to the constituent that came just before."* So, readers attempt to include each successive word into the clause currently being processed, until a word is encountered which can no longer be assigned a role in that clause. In a sentence such as *The dog bit the fox that was rabid*, this strategy suggests that a new clause is required when the word *that* is encountered, and that *that was rabid* should be attached to the immediately preceding element of *fox*.

Rayner, Carlson and Frazier (1983) provide evidence of a second preferred parsing strategy, namely to try to parse the simplest possible syntactic constituents. Thus, according to this strategy readers would in the first instance attempt to form active rather than passive sentences. Rayner et al used sentences such as: *The performer sent the flowers was very pleased/The performer sent the flowers and was very pleased with herself.* Reading times per character were longer in the first than the second example, suggesting that readers attempted to parse both examples in the simplest way as active sentences, but then had to reparse the first example as a passive construction on encountering *was very pleased.* Again, analysis of gaze durations and regressive eye movements supported this interpretation.

Frazier and Rayner's results support their two hypothesised parsing strategies, as well as Just and Carpenter's immediacy assumption. This position is also consistent with the work of van Dijk and Kintsch (1983) and Clark and Clark

(1977). Although like all heuristics these parsing strategies can be wrong, very often they do work. They permit on-line intraclause integration in the absence of complete information, and thus reduce the burden on working memory capacity which would exist if information had to be buffered before analysis.

1.3.4 Interclause integration

The fourth stage of the Just and Carpenter model involves integrating concepts with previously processed information beyond the current clause. Some of the ways in which this integration can take place include searching for argument repetition or semantic entailment, constructing bridging inferences, and finding referents for pronouns and other anaphoric expressions.

Argument repetition and semantic entailment

The simplest form of integration involves searching for argument repetition, as in the sequence *The child put the flowers in a vase. The flowers brightened the room.* Kintsch and van Dijk (1978) in their original model used argument repetition as a basis for integration and establishing coherence. Even though these authors acknowledged that it was neither necessary nor sufficient, they did nevertheless propose that checking for argument repetition was a useful first step in establishing coherence. They suggested that incoming information was checked against previously processed information still active in the short-term memory buffer. If at least one shared argument was located between new and old information, the new information was accepted as coherent, whereas if none was found an inference procedure was initiated to produce overlapping arguments.

There is some evidence to show that argument overlap between propositions plays a role in establishing cohesion, as for example in the study by McKoon and Ratcliff (1980). However, several studies also show that argument overlap is not sufficient. Manelis and Yekovich (1976) examined recall of what they called related and unrelated paragraphs, as in the following examples:

1. *Arnold lunged at Norman. Norman called the doctor. The doctor arrived.*

2. Ellen pushed Judith. Judith obeyed the nurse. The nurse coughed.

Both paragraphs are coherent if argument overlap is the criterion. However, in the related paragraph (1) the verbs contribute to a possible scenario, whereas in (2) they do not. Related paragraphs were recalled significantly better than unrelated paragraphs.

Haberlandt and Bingham (1978) used materials similar to those of Manelis and Yekovich to show the importance of verbs in the construction of causally related sequences of events. They compared pairs of paragraphs such as the following:

1. *Brian punched George. George called the doctor. The doctor arrived.*

2. *Brian punched George. George liked the doctor. The doctor arrived.*

As in the case of the Manelis and Yekovich (1976) examples each sentence in a paragraph was related to the previous one by argument overlap, but only the first paragraph could be constructed into a coherent scenario with a consistent theme. Readers took longer to read the second paragraph than the first, with the difference largest for the final sentence. Haberlandt and Bingham (1978) suggest that the increased processing time was due to difficulties with integration, thus showing again that argument overlap is not sufficient for establishing coherence.

In view of findings such as those above, van Dijk and Kintsch (1983) take a broader view of integration to establish coherence than the one taken in the 1978 model. Argument overlap is now seen as only one indicator of continuity of topic, and other indicators also need to be taken into account. Van Dijk and

Kintsch (1983, p. 150) note that: *"Local coherence does not stand alone ... linear coherence needs macrocontrol in the form of a theme, topic or point, as they are theoretically constructed as macrostructures."* The second set of sentences in the above studies were lacking an overall topic or theme, and hence were not coherent in spite of argument overlap.

Van Dijk and Kintsch (1983) propose that referents in a text must be organised in relation to some central object(s) or person(s) acting as the topic of the discourse, predicates must be organised in relation to macroactions or events, sequences of actions or events must have global goals or motivations, and there must be some unity of time and place in the events described. These authors point out that readers draw on a range of general knowledge about concepts and their relationships in order to integrate information in the coherent manner suggested.

Just and Carpenter (1987) note that a reader's knowledge of synonyms, part-whole relationships and category-exemplar relationships can be used to find points of attachment between incoming and previously processed information. The following example was mentioned earlier: *The child put the flowers in the vase. The flowers brightened the room.* If the above sequence continued *The girl knew her parents would be pleased,* the reader's knowledge suggests that *the child* and *the girl* are being used coreferentially. In other possible follow-on sentences, referring expressions may signal links to old information by means of category-exemplar relations, for example, *The roses smelled beautiful,* or part-whole relations, for example, *Unfortunately the petals were starting to fall.* Although the relations involved vary, in each case the nouns suggest a point of attachment to previously processed information. In each case there is the coherence of theme and the continuity of person and object required by van Dijk and Kintsch (1983).

Carpenter and Just (1977) discuss verb entailment as a cue to integration. The reader's knowledge about a verb such as *murder* suggests an agent is required, unlike in the case of a verb like *die*. Carpenter and Just analysed eye movement data on sentence pairs such as the following: *It was dark and stormy the night the millionaire was murdered/died. The killer left no clues for the police to trace.* Not only were gaze durations on the word *killer* significantly longer when the preceding verb was *died* rather than *murdered*, but gaze durations were also longer at the end of the sentence when the verb was *died*, suggesting that increased difficulty with integrating the word *killer* during sentence wrap-up was responsible for a large part of the increased processing time. The verb *murdered* suggests a point of attachment for the word *killer* in a way that *died* does not.

Just and Carpenter (1987) report a similar finding when a verb implies an instrument, as in the sequence *The maid diligently swept/cleaned the floor until it was spotless. The broom had been worn down by her excessive zeal.* Gaze durations on the word *broom* were significantly shorter when the verb was *swept* rather than *cleaned*, again suggesting that the implication of the verb assisted integration.

Inferences

The previous examples already require some degree of inferencing to establish relationships between old and new information. However, in some cases more extensive reference to general knowledge is needed to establish cohesive links between successive clauses. For example, in the sequence *Mary fell down the stairs. John ran for the doctor*, there is no explicit point of attachment for the second sentence. Integration can only take place following substantial inferencing

to the effect that falling down stairs can lead to injuries, which may need to be treated by a doctor. Once these inferences are made, argument repetition or semantic entailment provide points of attachment for each successive proposition.

Haviland and Clark (1974) refer to inferences such as those described above as bridging inferences, since they create a bridge between old and new information. Bridging inferences are backward inferences, made once the reader can find no explicit point of attachment for incoming information. Thus, in the previous example a bridging inference would be initiated from *doctor* back to previous information in order to construct an explicit connection.

Haviland and Clark (1974) constructed sentence pairs which in one version were explicitly connected by argument repetition, while in the other a bridging inference was required. Reading times were longer for the second sentence when an inference was required, suggesting that readers were indeed constructing the inference to integrate the two sentences as they read. The same result was found even when the explicit condition used coreferential expressions (*became absolutely smashed/got very drunk*) rather than argument repetition. Integrating such expressions still took less time than constructing bridging inferences.

Miller (1980, expt 3) also demonstrated that bridging inferences required for integration are made while reading by examining the speed with which statements could be verified when they were either implicit or explicit in the stimulus material. The following sentence pair is an example of the implicit condition: *A burning cigarette was carelessly discarded. The fire destroyed many acres of virgin forest.* The explicit version stated that the cigarette had started a fire. The test sentence to be verified was: *The cigarette started a fire.* There was no significant difference in response times between the implicit and explicit conditions, suggest-

ing that readers had inferred the connection between the cigarette and the fire while reading.

While studies such as these support the view that careful readers construct bridging inferences to integrate successive segments of text which are not explicitly connected, the picture is less clear with regard to elaborative inferences. As the name suggests, such inferences elaborate on information in a text, but are not necessary to connect new with old information. Elaborative inferences are sometimes referred to as forward inferences since they do not refer back to previous information in the way that bridging inferences do.

Miller (1980, expt 1) and Johnson, Bransford and Soloman (1973) argue that elaborative instrument inferences are made during reading. On reading *he was pounding a nail*, subjects were assumed to infer the probable instrument as a hammer on the basis of a high false alarm recognition rate for a test item containing that instrument (Johnson et al, 1973). However, Corbett and Dosher (1978) and Singer (1979a, 1979b) dispute this conclusion, arguing that the inferences were made at the time of testing. Singer (1979b) showed that for sentence pairs such as *The boy cleared the snow (from the stairs/with a shovel). The shovel was heavy*, readers took longer to read the second sentence when the instrument was not mentioned in the first sentence, indicating that *shovel* was inferred as a bridging inference while reading the second sentence, not as an elaborative inference while reading the first.

In the case of implicit consequences/antecedents, Brewer (1977) found that elaborative inferences were made. Following sentences such as *The safecracker put a match to the fuse*, subjects were more likely to produce *lit the fuse* in their free recall than the original *put a match to the fuse*, indicating that they had

made an elaborative inference. However, Singer and Ferreira (1983) did not find evidence of elaborative causal inferences. Following a sentence such as *A rock flew through the window*, an elaborative inference would state that the window broke. However, following a stimulus such as *Bob threw his report into the fire. The ashes floated up the chimney*, it would be a bridging inference that stated the report had burned. After reading, subjects answered questions such as *Did the window break? Did the report burn?* Questions based on bridging inferences were answered significantly faster than those based on elaborative inferences, suggesting that bridging inferences were made during reading, while elaborative inferences were not.

It appears that the question to be answered is not whether readers do or do not make inferences, but under what conditions they do so. Some studies have shown that elaborative instrument inferences may be made, but only when there is a highly probable instrument suggested by the text. For example, Garrod and Sanford (1981) show that a sentence such as *The car had recently been overhauled* was read just as rapidly after a sentence mentioning only *driving* as it was after a similar sentence mentioning *taking the car*. The authors therefore conclude that *car* was inferred as a highly probable instrument of *driving* while reading the first sentence.

O'Brien, Shank, Meyers and Rayner (1988) also show that elaborative inferences may be made, but only in highly constrained conditions. These authors used paragraphs such as the following:

> *"All the mugger wanted to do was steal the woman's money. But when she screamed, he stabbed/assaulted her with his knife/weapon in an attempt to quiet her. He threw the knife into the bushes and ran*

away."

Thus, there were four versions of the text, varying in degree of explicitness (*knife/weapon*) and degree of constraint (*stabbed/assaulted*). Gaze durations on the target word *knife* in the the final sentence were comparable in the explicit conditions and the highly constrained implicit condition, suggesting that following the word *stabbed*, readers inferred that *weapon* referred to a knife. However, it appears that no such elaborative inference was made following the less constraining verb *assaulted*, hence a bridging inference was required in the final sentence to connect *weapon* and *knife*, and processing time was increased as a result.

McKoon and Ratcliff (1981), in a carefully designed study, confirmed that instruments must be highly probable before they are inferred. These authors used paragraphs in which, for example, the first sentence mentioned a number of tools, including a hammer. The fifth and final sentence then stated either that boards were glued together or pounded together with nails. Subjects then verified whether or not words had occurred in the paragraph. *Hammer* was verified significantly faster following *pounded together* rather than *glued together.* According to McKoon and Ratcliff, this result occurred because subjects inferred that the pounding was being done with the hammer referred to earlier in the paragraph. No such inference was made on reading *glued together.* McKoon and Ratcliff also show that no such inference was made when the instrument was less probable, for example, a mallet, or when the first sentence stated that the hammer was broken. Thus, rapid verification of *hammer* in the test did not depend only on a pre-experimental association between *pound, hammer* and *nail*, but rather depended on the coherently constructed meaning of the paragraph, involving a highly probably instrument.

Finally, Goetz (1979) also carried out a carefully designed study, which showed that a number of factors influence whether or not a particular inference will be made. He wrote passages of which one could be summarised as follows: Joan was flying to San Francisco to start a new job. She had just enough time to catch her plane. As she was packing, she remembered that her son wanted her to ring before she left. She arrived at the airport half an hour late and had to wait for the next plane.

There were eight versions of the text, half of which explicitly stated that Joan was late because she phoned her son, and half of which did not. Within each of these groups, the texts also varied in terms of the salience of the phone call and the importance of missing the plane. In the highly salient versions, Joan had promised to ring her son, while in the less salient versions, the son had only wanted her to ring. In the highly important versions, the plane missed by Joan crashed, killing all passengers, whereas in the less important versions, missing the plane resulted only in Joan having to kill time at the airport.

After reading one version of the passage, subjects were asked: *Why was Joan late?*. More correct answers were given when the information was explicit. However, of greater interest here is the fact that the percentage of correct answers when the information was implicit varied according to condition. The target inference was more likely to be made when the premise on which it was based was salient and when the consequence of the action was important. In other words, if Joan had promised to talk to her son it was more likely that she had actually done so, and if the consequence of the delay was that her life had been saved it was more likely that readers would construct a reason for this lucky escape.

The number of inferences which could potentially be constructed from a text

is very large, but clearly this number has to be restricted if processing is to be carried out on-line within the limitations of working memory capacity. As the empirical evidence and the Just and Carpenter (1987) model indicate, inferences tend to be restricted to bridging inferences required to establish coherent relations between text segments during integration, and elaborative inferences which are strongly suggested by a highly constraining context.

This is also the conclusion reached by McKoon and Ratcliff (1992) in what they call their minimalist hypothesis. These authors consider that many readers engage in minimalist processing in situations such as reading a magazine on the plane, reading a newspaper over breakfast, or reading texts without specific instructions in psychology experiments. In such cases, McKoon and Ratcliff conclude that readers do not automatically construct inferences to fully represent the context described in the text. Rather, they construct only those inferences required to establish local coherence, as well as elaborative inferences based on information which is quickly and easily accessible. Of such elaborative inferences, McKoon and Ratcliff (1992, p. 441) make the following comment:

> *"Inferences based on general knowledge have been demonstrated in the encoding of such inferences as elaborations about 'what will happen next' in a story if what will happen next is very predictable, the encoding of inferences about aspects of the meaning of words if they are highly typical aspects, the encoding of inferences about instances of categories if the instances are highly typical, and so on."*

McKoon and Ratcliff make it clear that they are discussing minimalist processing, not the strategic, goal-directed processing readers often use when learning new information from texts. Under such circumstances, readers are likely

to engage in more extensive inferencing, including global inferences. As already indicated, Just and Carpenter (1980) also acknowledge that the reader's goals influence processing through the criteria of coherence required for the attainment of those goals. Readers intending to read a text only superficially are likely to have less stringent criteria of coherence, hence will require fewer bridging inferences to satisfy their criteria than readers who wish to construct a detailed text representation. It may also be that thoughtful readers construct further elaborative inferences while reflecting back on the text but, as noted, the number of such inferences is limited during the process of reading.

Anaphoric expressions

Apart from bridging inferences, another important component of interclause integration is assigning referents to anaphoric expressions, including pronouns, proverbs, and many other expressions, such as: *We will meet there/then.* In all these cases, a previous reference in the text has to be found to which the anaphoric expression can be attached. For example, a previous reference to place or time has to be found to determine what *there* or *then* refer to. Most of the empirical work in this area has focused on pronoun assignment, although similar principles apply to the resolution of other types of anaphoric expressions.

Within the Just and Carpenter (1980) framework, anaphoric resolution is attempted immediately, but may need to be completed or updated during subsequent processing. The authors use the following example to illustrate this point: *Although he spoke softly, yesterday's speaker could hear the little boy's question.* The pronoun *he* is initially assigned the referent of *yesterday's speaker*, but this assignment has to be changed when *little boy* is encountered.

Just and Carpenter (1987) suggest a number of cues that can be used to assist in determining the referent of a pronoun:

1. the gender and number of possible referents.

2. the noun which plays the same grammatical role as the pronoun. Frederiksen (1981) nicely illustrates the operation of this strategy with examples such as the following:

   a) *Modern advertising does not, as a rule, seek to demonstrate the superior quality of the product.*

   b) The superior quality of the product is not, as a rule, what modern advertising seeks to demonstrate.

   c) It plays up to the desire of Americans to conform, to be like the Jones.

   Sentence c) was read more rapidly when it followed sentence a) since the pronoun in c) could be assigned a referent from a) according to the strategy of matching grammatical roles, which was not appropriate following b).

3. the noun which is foregrounded. Carpenter and Just (1977) report that a noun which is foregrounded as new information in a cleft or pseudocleft construction is more likely to be chosen as the referent for a subsequent pronoun. Thus, in an example such as: *It was the woman in jeans who helped the girl after the accident. She ...*, the pronoun is more likely to be assigned the referent of *woman* because this noun is foregrounded.

4. the noun most recently mentioned.

5. knowledge about possible referents. Just and Carpenter (1987) use the example *The Pope and the rabbi met for lunch; he was accompanied by his wife* to indicate how pronoun assignment is guided by general knowledge. The Pope is known to be unmarried, so *he* must refer to the rabbi. It should be noted, though, that this assignment could not be made until the end of the sentence.

Knowledge about referents is likely to be particularly important when pronouns are ambiguous, that is, they have more than one possible antecedent which is compatible in gender and number. Cues such as recency and and syntactic structure on their own are of limited usefulness in these cases. For example, Garvey, Caramazza and Yates (1975) give the following examples of pronoun assignment based on knowledge of referents:

> *Rosemary trusted the secretary because she was efficient/she was a good administrator/she was gullible.*

In the first example, *she* is likely to be assigned to the referent *secretary*, since being efficient is a good reason for being trusted. In the third example, on the other hand, *she* is likely to be assigned to the referent *Rosemary*, since being gullible is not a reason why the secretary would be trusted, so it must be Rosemary who is gullible because she misplaced her trust. In the second example, *she* could refer to either the secretary or Rosemary; perhaps the secretary is trusted because she is a good administrator, or Rosemary is a good administrator because she trusts her staff.

In cases such as the sentences above, or the rabbi and Pope sentence mentioned previously, it is unlikely that pronoun assignment can be completed before the

end of the sentence, a point which Just and Carpenter (1987, p. 208) take into account when they describe pronoun assignment in the following way:

> *"The search for the referent of a pronoun is initiated as soon as the pronoun is encountered in the text. The search operates by using the various heuristic cues discussed above to activate the potential antecedents. That is, activation is directed from each cue source to all the elements of the referential representation that satisfy the constraint imposed by the cue. The amount of activation each candidate receives is a function of three factors: (1) the number of cues that support its candidacy; (2) the relative importance of each of the cues (for example, a gender cue has more weight than a recency cue); and (3) the strength of the supporting evidence accompanying the cue in the current context.*
>
> *The candidate whose activation level is the first to reach threshold is chosen as the pronoun's antecedent in the referential representation. If no candidate's activation level reaches threshold by the time some deadline has expired, then the referential assignment is postponed until the end of the clause or sentence, when the search is reinstated with the benefit of any additional cues encountered in the latter part of the sentence."*

Just and Carpenter (1987) note that this process applies not only to pronouns, but to any expression which refers back to previous information. For example, the same process could apply to proverbs, as in determining the meaning of *did* in the sequence *Mary laughed hysterically when she heard the news and so did John.*

It should be noted that in the view quoted above, Just and Carpenter (1987) are not putting forward a strong version of the immediacy assumption. Their position is thus not undermined by showing that pronoun assignment, or indeed the assignment of referents in general, is not always completed while the eye is fixating the pronoun or other anaphoric expression.

One study examining the locus of pronoun assignment was carried out by Ehrlich and Rayner (1983), who measured fixation times in the vicinity of an unambiguous pronoun, that is, one for which there was only one possible referent in the preceding text. This referent was either in the sentence previous to the pronoun and immediately adjacent to it, or two sentences removed.

Fixation durations increased sharply at the point of encoding the pronoun. In the near condition, fixation durations then decreased, but in the far condition the raised durations continued over the two subsequent fixations, suggesting that the assignment of a referent for the pronoun was carried out while the pronoun was being fixated in the near condition, but in the far condition processing continued beyond this point. This conclusion was also supported by the observation that the length of saccades, that is, the number of characters covered by a single eye movement, was shorter following pronoun encoding in the far condition compared to the near condition.

As Ehrlich and Rayner (1983) point out, pronoun assignment is not like lexical access in that referents for a pronoun cannot be looked up in an internal lexicon, but have to be searched for in the integrated representation of previous sections of text. Hence pronoun assignment is likely to be a more time consuming process than lexical access, especially when the referent is far removed from the pronoun in the surface structure of the text, or when the pronoun is ambiguous and there is

insufficient preceding information to select one antecedent rather than the other. However, this view does not contradict the Just and Carpenter (1987) view.

Ehrlich (1983, p. 260) presents very similar findings and concludes that:

> *"There is no single place in the text at which pronoun assignment always occurs. Rather, the effect of increasing the distance between the pronoun and the antecedent is to shift the locus of assignment to a later point. Thus it seems that readers do not remain at a particular fixation until they have recovered the antecedent, but they keep reading, perhaps pausing, when the antecedent has been accessed, causing a momentary increase in processing."*

Unlike Ehrlich and Rayner, Corbett and Chang (1983) used pronouns with two possible antecedents. After each sentence, subjects were asked to verify as quickly as possible whether a certain word, including one of the two possible antecedents, had occurred in that sentence. From differences in response times Corbett and Chang concluded that both possible antecedents were activated when the ambiguous pronoun was encoded. The authors propose that pronoun assignment was initiated as soon as the pronoun was encoded, but then continued in parallel with subsequent processing.

In such cases, a final assignment may not be made until the sentence wrap-up phase. Van Dijk and Kintsch (1983) refer to the work of Frederiksen (1981), which indicates that readers take longer to read a sentence containing an ambiguous pronoun. Van Dijk and Kintsch endorse the view expressed by Frederiksen (1981, p. 379):

> *"...readers appear to retrieve all of the alternative referents that*

are available for a pronoun (i.e., nouns that agree in number and gender) and then select from among them the referent that fits the semantic constraints of the sentence in which it occurred."

Van Dijk and Kintsch suggest that alternative pronoun referents are considered within a clause, but once a major syntactic boundary is reached, one candidate is selected as the referent and others are deactivated. Kintsch (1988) gives an example of such processing. He proposes that in the sequence: *The lawyer discussed the case with the judge. He said: I shall send the defendant to prison*, both possible referents of *he* are considered, and therefore two propositions are constructed for the last clause, namely SEND(LAWYER, DEFENDANT, PRISON) and SEND(JUDGE, DEFENDANT, PRISON). The latter, but not the former proposition will activate additional propositions in the knowledge net associated with the sentencing power of judges. Therefore the activation level of *judge* as a potential referent will be increased and this referent is likely to be selected, but possibly not before sentence wrap-up.

This position is consistent with the model of pronoun assignment put forward by Just and Carpenter (1987). If in the case of an ambiguous pronoun both potential referents receive activation from several cue sources, such as number, gender, recency, and so on, and both are activated as a result, then a final selection may have to be deferred until further evidence accumulates from the rest of the sentence.

In summary, readers use many sources of information in order to coherently integrate each successive proposition with previously processed text. Readers use cues from the text to suggest how such links might be established in the text representation but, as indicated, readers also need to draw heavily on general

knowledge. In both the Just and Carpenter (1980) and van Dijk and Kintsch (1983) models, all the aspects of integration discussed are attempted immediately, although a final integration may not be completed until the sentence wrap-up phase.

1.3.5 Sentence wrap-up

The final stage in the Just and Carpenter model is sentence wrap-up. This stage does not involve a different type of processing, but is defined by its occurrence at the end of sentences. Since a sentence embodies one thought, the sentence boundary is a useful place to revise or conclude processing in the light of more complete information. The types of processes carried out include assigning referents which may previously not have been clear, constructing inferences where necessary to connect clauses, and resolving any remaining inconsistencies, for example, as in determining the meaning of ambiguous words.

Just and Carpenter (1980) report that gaze durations were on average 71 msec longer on the final word of a sentence. Daneman and Carpenter (1983) compared gaze durations on the same word when it occurred either in the middle or the end of a sentence, as in the following example:

He found a bat that was very large and brown ...

He found a bat. It was very large and brown ...

Gaze durations on the target noun, in the above case *bat*, were an average of 207 msec longer when the word occurred at the end rather than in the middle of the sentence.

Another indication that additional processing is carried out at the end of sentences comes from examples where integration is deliberately made difficult. It was previously mentioned that verb entailment provides a powerful cue to the relationships that need to be established during integration, as in the example *... the millionaire was murdered. The killer ....* When the verb was changed to

died, processing time was increased by about 500 msec, which was distributed between the word to be integrated, that is, *killer*, and the end of the sentence (Just and Carpenter, 1980). It appears, therefore, that integration is attempted immediately, but is not concluded until the end of the sentence.

Mitchell and Green (1978) also report that viewing time of the visual stimulus increases, not only at the end of sentences, but also at the end of clauses. This increase was particularly marked when material was more conceptually difficult, leading the authors to suggest that the additional time is required to integrate new information with that already encoded in the mental representation of the text. Thus, the longer and more conceptually difficult the sentence, the greater the increase in end-of-sentence viewing time due to the increased demands of integrative processing.

According to Just and Carpenter (1980), the increased burden placed on working memory by the end-of-sentence integrative processes may help to explain why verbatim memory declines sharply after a sentence boundary. Once exact word and word order information has been used for lexical access, syntactic analysis and case role assignment, it is unlikely to participate in the productions involved in conceptual integration and therefore begins to decay. On the other hand, the conceptual information involved in the integrative processes maintains a high level of activation. As the increased processing demands of sentence wrap-up decrease working memory storage capacity, information with the lowest activation level is likely to be displaced from working memory, thereby also decreasing its chances of being incorporated into the long-term text representation and being available for subsequent recall.

Just and Carpenter (1980) also report that gaze durations increased by an

average of 157 msec on the final word of a paragraph. It is not clear whether the integrative processes carried out at this point are similar to the ones at the end of sentences, that is, finding points of attachment for all concepts in the sentence, or whether larger segments of text, such as the text roles, are being interrelated and linked to the topic of discourse. Just and Carpenter do not elaborate on this point.

How often and where wrap-up processes are carried out varies depending on the goals of the reader and the type of text being read (Just and Carpenter, 1980). The reader who skims a relatively simple passage may perform wrap-up processes only at the end of paragraphs, while the reader attempting to fully comprehend a legal contract may carry out these processes at the end of each clause.

Just and Carpenter (1980) note that the reader's goals influence the criteria adopted for processing, including the wrap-up phase. Readers determined to understand a text fully are likely to have more stringent criteria, and hence are likely to spend more time on each wrap-up phase to ensure a complete integration of concepts, possibly also including elaborated meanings in their text representation.

Kintsch (van Dijk and Kintsch, 1983; Kintsch, 1988) proposes a similar function for the sentence wrap-up phase. During this processing pronoun assignment may need to be completed, as in the example given previously of a judge and a lawyer discussing the sentencing of a defendant (Kintsch, 1988). Any word sense elaborations may be integrated into the text representation at this time (Till, Mross and Kintsch, 1988).

However, some authors ascribe a more extensive role to end of sentence processing. For example, Haberlandt and Graesser (1989, p. 187) devised a buffer-

integrate-purge model, which they describe as follows:

> *"During the buffering phase of a cycle, the words within a clause accumulate within a temporary buffer until the end of a clause or sentence is reached. During the integration phase, at the end of a clause or sentence, the new information in the clause is linked with the previous passage context and thus is passed to the text representation. Integration processes typically involve such operations as search for antecedents and/or inferences. They are relatively resource-intensive and time-consuming, at least compared with buffering operations. At the conclusion of a cycle during the purge phase, the buffer is cleared of the content in the current clause."*

Haberlandt and Graesser examined gaze durations as university students read passages from *Readers Digest* and *Scientific American.* Not only did gaze durations on boundary words (end of clause or sentence) increase as the cumulative number of new argument nouns in the clause or sentence increased, but gaze durations in non-boundary words increased as well. The boundary increase was assumed to be due to the greater difficulty of integrating the larger number of new arguments, while the non-boundary increase was assumed to be due to the need to buffer an increased number of new arguments.

However, throughout this chapter evidence has been cited to indicate that many aspects of processing are initiated immediately, even if they are not completed immediately. For example, Frazier and Rayner (1982) showed that readers have preferred strategies to parse sentences on-line, and that they immediately become aware of parsing problems when the strategies are inappropriate. They

do not wait until the end of the clause or sentence to parse and reparse. Carpenter and Daneman (1981) and Daneman and Carpenter (1983) showed that readers immediately become aware of, and often immediately try to correct, contextually inappropriate interpretations of ambiguous words. Just and Carpenter (1980) showed that integration is made easier, and processing time is reduced, when verbs with appropriate semantic entailments are used, as in the example *"...when the millionaire the murdered. The killer ..."*. Gaze durations on the word *killer* were shorter following the the verb *murdered* than following *died*, which would not be the case if integration was delayed until the end of the clause or sentence.

None of these findings can be accommodated within a model that insists on all integration being delayed until a clause or sentence boundary. A weak version of Just and Carpenter's immediacy assumption, where integration is attempted immediately but may not be concluded until the wrap-up phase, takes better account of a range of empirical findings.

1.3.6 Macrostructure formation

The previous section on interclause integration discussed strategies used by readers to relate each successive proposition to already processed text to form a coherent representation. However, some authors have suggested that skilled readers establish more than this local coherence; they also condense the text into a macrostructure which establishes its global coherence (van Dijk and Kintsch, 1983). Just and Carpenter (1987, p. 251) make some reference to such a thematic summary when they state that *"... a text can provide a high-level abstraction, like a summary statement of a theme, that encompasses a number of statements and relates them to each other. If such an abstraction is not provided by the text and the reader needs one, then he must construct it himself."* However, these authors do not discuss such summary statements in as much detail as van Dijk and Kintsch (1983).

The nature of macrostructures and their effect on processing

At the highest, most condensed level, the macrostructure consists of a statement of the overall theme or topic of a passage. At the lower level, Kintsch and van Dijk (1978) propose that the macrostructure results from the deletion of details and the construction of superordinate statements to subsume related groups of propositions in the text. Van Dijk (1977) also points out that some macropropositions are stated explicitly in the text, and hence need only to be identified rather than constructed by the reader. Along similar lines, Just and Carpenter (1980) consider that in scientific text experienced readers search for and process more deeply information related to topics, definitions, causes and consequences. While such information is part of the text representation constructed directly from ex-

plicitly stated information, it is also the basis for a summary or gist retelling of a passage.

Meyer in her work (for example, Meyer, Brandt and Bluth, 1980) makes a similar point; good readers use their knowledge of conventional text structures to search for high-level, important information, which receives additional processing. Again, information such as the statement of a problem and its possible solutions, one of the text structures studied by Meyer, could be an effective summary of the text.

What these examples illustrate is that there is some overlap between the macrostructure of a text and the textbase or microstructure; some propositions can be part of both. In this discussion of more global aspects of comprehension, the macrostructure will be referred to as a condensation of a text into its most important elements, regardless of whether these elements are constructed by the reader or selected from explicit statements in the text. However, macropropositions must be coherent statements, not merely lists of key points (van Dijk and Kintsch, 1983).

Kintsch and van Dijk (1978, p. 365-6) note that coherence in a discourse cannot be established only at a local level by connecting each propositions to those already processed. *"Relating propositions in a local manner is not sufficient. There must be a global constraint that establishes a meaningful whole, characterized in terms of a discourse topic."* This point was illustrated in the section on argument repetition; argument repetition was not sufficient to connect sets of sentences lacking an overall theme.

The importance of a thematic focus for coherent text processing and recall is shown in several studies. For example, Bransford and Johnson (1972) indicate

how difficult it is to comprehend and recall a passage in which clues as to the overall theme are deliberately withheld. These authors began a passage about washing clothes as follows:

> *"The procedure is actually quite simple. First you arrange things into different groups depending on their make-up. Of course one pile may be sufficient depending on how much there is to do."*

When no title was supplied to give a clue to the theme of the passage, readers rated the comprehensibility of the passage as low and recalled it poorly. Dooling and Mullet (1973) report similar findings from a passage about Christopher Columbus, in which the ships were referred to as *three sturdy sisters* and the ocean as a *vast calmness.* Without a title to guide interpretation, it was difficult for readers to comprehend the text. Similarly, in Bransford and McCarrell (1974), subjects found it difficult to recall sentences such as *The haystack was important because the cloth ripped* without the aid of thematic cues such as *parachute.*

Schwarz and Flammer (1981) indicate that providing a thematic title, which is a clue to macrostructure formation, increases recall from passages. However, a title did not improve recall of very disorganised passages, where sentences were presented in random order, unless subjects were given twice as much reading time. With extra time and a thematic title readers were presumably able to restructure the disorganised text into a more coherent whole, improving recall as a result.

Schallert (1976) shows that different thematic titles influence not only the quantity recalled, but also the overall interpretation of the text. In other words, the title influences the construction of the main ideas of the text. A passage containing ambiguous sentences such as the following was used: *Our most valuable*

pitchers might crack in this heat. If only we had more fans, we would all feel better I'm sure. The title was either *Worries of a baseball team manager* or *Worries of a glassware factory manager.* In both free recall and multiple choice questions subjects showed that they had interpreted the passage in a manner consistent with the thematic title.

While in these studies the texts were deliberately constructed for the purpose and were somewhat artificial, it is quite possible that a careless or immature reader who did not process the thematic information available in a more natural text would be in much the same position as subjects reading the Bransford and Johnson text without a title. Without some form of macrocontrol it could be difficult to establish precise referents and to construct relationships between them.

Van Dijk and Kintsch (1983, p. 195) consider that the formation of a macrostructure is extremely important for the performance of a range of complex tasks.

> *"When we understand a text, we no longer have access to all previous sentences we have read, and the same holds for the understanding of events and actions, of which the multiple details can only be partially retrieved. If, however, large sequences of semantic structures of this kind can be subsumed by macrorules under a few hierarchically structured macropropositions, and if such macrostructures are a sufficient basis for the further understanding of the discourse, the events, or the actions, then the extremely complex task of keeping some order in the vast amounts of semantic detail can be managed. That is, we have a) a common semantic basis to connect (sometimes incoherent) sequences; b) a relatively simple semantic structure which may*

be kept in short-term memory; c) a device for the hierarchical organization of the sequences in episodic memory; d) an important cue for the actualization of knowledge for longer stretches of discourse, events or actions; e) powerful retrieval cues for the sometimes necessary reactivation of semantic details; and f) an explicit construction defining the essential, most important or relevant semantic properties of a discourse or episode. This information is needed for later cognitive tasks (learning, recall, action planning), as well as for monitoring and controlling further discourse processing."

While the role of a macrostructure as described above by van Dijk and Kintsch seems plausible, the question is whether there is any empirical evidence to support it. Some authors have hypothesised that signalling in a text, that is, additions such as headings and topic sentences to highlight the global structure, should make it easier for readers to form a macrostructure, which in turn should influence what is learned from the text. These authors assume that the construction of the global topic structure of an expository text is important. For example, Lorch and Lorch (1985, p. 137) express this importance as follows: *"Because most texts are organized around a small set of topics, an important component of text processing involves identifying a text's major topics and their interrelations."*

Lorch, Lorch and Inman (1993) signalled the structure of their experimental texts through the use of topic overviews and summaries, headings, and typographical indicators of subsection divisions. The effects of signalling included recall of a larger number of topics, greater clustering of subordinate information around topics in recall, and a closer correlation of the order in which information was presented in the text and in recall. However, readers did not recall more

detailed information about each topic as a result of signalling. Similar results are reported by Lorch and Lorch (1985, expt. 2) using topic sentences at the beginning of each paragraph as signals to make the topic structure more salient.

It appears, then, that signalling does increase recall of the topic structure of a text, and that the encoding of this global topic structure does produce a more organised recall, as proposed by van Dijk and Kintsch (1983). However, in these studies there was no evidence that the topic structure provided retrieval cues for the reactivation of semantic details, another of the van Dijk and Kintsch proposals.

Loman and Mayer (1983) examined in more detail how signalling and understanding of the global relations in a text influence not only recall but also problem solving. They hypothesised that in order to generate high quality solutions to a problem described in a text, an integrated and coherent understanding of the overall chain of cause and effect leading to the problem was necessary. Therefore, if signalling increased this understanding, it should also lead to higher quality problem solving.

Loman and Mayer used preview sentences to emphasise the structure of the passage, headings to identify each of the major conceptual parts, and logical connectives such as *because of this* and *as a result of this* to signal important causal relationships. As a result of receiving a text with this signalling, grade 10 students recalled more conceptual information, that is, idea units in the cause and effect chain, and generated more high quality solutions to the problem described in the text. However, they did not recall more detailed information, as measured by multiple choice questions, or improve their verbatim recognition, as measured by a sentence verification task.

This finding highlights an important aspect of the van Dijk and Kintsch (1983) view cited earlier, namely, that constructing the most important elements of a text is essential for later cognitive tasks, such as creative problem solving. Thus, while not all aspects of the van Dijk and Kintsch view are supported by the studies referred to here, many important aspects thereof are.

Another question to be addressed is whether readers normally construct a macrostructure or general topic structure as they read. There is some evidence to suggest that adult readers do establish the global structure as they read. In Guindon and Kintsch (1984), for example, subjects read paragraphs with or without summary statements. They then verified a list of words, indicating whether or not these words had appeared in the paragraph.

Following paragraphs with explicit summary statements, both response times and error rates were lower for the acceptance of macrowords than for words taken from micropropositions. However, when the summary statement was not presented, it took longer to reject the macrowords than words both thematically related and unrelated to the micropropositions. There was also a mean false alarm rate of 65%. In other words, in only one third of responses were the macrowords correctly rejected, and subjects took a long time to do so. When these words were incorrectly accepted in the other two thirds of responses, response times were significantly shorter for macrowords than for both thematically related and unrelated microwords. Thus, it was difficult to reject macrowords even when they had not been presented in the paragraphs, suggesting that readers inferred a summary statement when it was not explicitly presented in the text.

Lorch, Lorch and Matthews (1985) also conclude that readers make use of macrorelevant information in the form of topic sentences and relate such state-

ments to previous paragraph topics. These authors investigated reading times for topic and non-topic sentences in a variety of conditions. The experimental texts described six attributes of two entities in twelve paragraphs. One text, for example, discussed features of two fictional countries. One version of the text was organised by attributes, so if the countries are represented by upper case letters and the attributes by lower case letters, the order of the paragraphs was Aa, Ba, Ab, Bb, and so on. In the disorganised version there was a shift of both attributes and countries with each paragraph, for example, Aa, Bb, Ac, Bd, and so on. A change of country but not attributes was considered a minor shift, while a change of both was considered a major shift. Lorch et al assumed that macroprocessing would be more difficult following a major shift because it would be more difficult to relate the current topic to previous paragraph topics. If macroprocessing is done spontaneously while reading, this increased difficulty should be reflected in increased reading time for topic sentences.

In all conditions topic sentences showed longer reading times than non-topic (detail) sentences. In the organised version, reading times for topic sentences were longer following a major shift than a minor shift (this comparison is not relevant in the disorganised version, since all shifts were major). In a second experiment, Lorch et al examined reading times for the organised texts either with an introductory paragraph outlining the structure of the passage or a general and uninformative introduction. The presence or absence of this summary statement did not affect reading times for non-topic sentences, but the difference in reading times between topic and non-topic sentences was smaller when the introductory paragraph was informative rather than uninformative, suggesting that the overview paragraph made it easier for readers to establish the macrostructure

of the passage as they were reading.

While the above studies suggest that readers do process, and if necessary infer macrorelevant information, McKoon and Ratcliff (1992) point out that such processing does not occur automatically, but depends instead on particular reading goals and strategic processing. In the absence of such goals, these authors claim that global inferences are not made, and that global information is used to guide processing only when local coherence cannot be established, that is, when an integration problem is encountered. Experiments carried out by McKoon and Ratcliff supported these conclusions. However, the authors used a relatively simple action narrative without specific learning instructions (expts. 3 and 4), which may have encouraged their subjects to use the minimalist processing discussed by these authors. This type of narrative would normally be associated with reading for pleasure. Lorch, Lorch and Klusewitz (1993) found that college students made a clear distinction between reading as a form of study with a clear purpose in mind, and light reading for entertainment or to pass time. The reading processes reported for these two broad categories of reading were quite different, with the latter, not surprisingly, being much more superficial.

In contrast to the text used by McKoon and Ratcliff, the studies discussed above by Loman and Mayer (1983) and Lorch and his colleagues used detailed expository texts which subjects would more normally associate with study, perhaps encouraging them to engage in more than minimalist processing.

What can be concluded from these studies is that experienced readers can form a macrostructure while reading, given that they have the goal of learning from the text and understanding the most important concepts and relationships described. However, in the absence of such goals, readers are unlikely to devote

special processing to macrorelevant information. The importance of reading goals will be discussed again towards the end of this section, as well as under the heading of metacognition.

Forming a macrostructure

In their original model, Kintsch and van Dijk (1978) assumed that information was buffered before being processed. The reader was therefore presumed to be in a position to apply a set of macrorules to the group of propositions being held in the buffer at any one time. The most general rules presented by Kintsch and van Dijk (1978, p. 366) to condense a text to its essential points include the following:

1. "*Deletion: Each proposition that is neither a direct nor an indirect interpretation condition of a subsequent proposition may be deleted.*

2. *Generalization: Each sequence of propositions may be substituted by the general proposition denoting an immediate superset.*

3. *Construction: Each sequence of propositions may be substituted by a proposition denoting a global fact of which the facts denoted by the microstructure propositions are normal conditions, components or consequences.*"

Van Dijk (1977) refers to a further rule of invention or selection, which involves selecting macropropositions explicitly stated in the text, for example, in the form of topic sentences. Thus, a particular microproposition may be deleted from the macrostructure, retained as a macroproposition, or be subsumed in a generalisation or construction. As examples of the last two possibilities, a sequence of propositions listing several small domestic animals sharing a girl's home can

be generalised as *She had several pets*, while a description of the chopping and cooking of vegetables and pasta can be constructed as *He prepared a meal.*

However, the rules as set out above need to operate on complete discourse segments, making such rules quite appropriate for the task of summarising a text, but not appropriate for ongoing text comprehension based on the immediacy assumption. Since van Dijk and Kintsch (1983, p. 15-6) adopted this assumption, they propose that readers use more flexible heuristics, or macrostrategies, to form a macrostructure while reading.

> *"In contrast to the abstract macrorules as defined in our earlier work, macrostrategies are flexible and have a heuristic character. A language user need not wait until the end of a paragraph, chapter, or whole discourse before being able to infer what the text or the text fragment is about, globally speaking. In other words, it is plausible that with a minimum of textual information from the first propositions, the language user makes guesses about such a topic. These guesses will be sustained by various kinds of information, such as titles, thematic words, thematic first sentences, knowledge about possible ensuing global events or actions, and information from the context. Again we see that an expedient strategy will operate on many kinds of information, which individually are incomplete or insufficient to make the relevant hypothetical assumption."*

Many sources of information provide clues to macrostructure formation. Within a text, the most explicit and complete clue is an abstract or summary, which presents the main points of the discourse in an already coherent form.

Other useful, but less complete clues are headings and subheadings. They can guide the generation of subsequent macropropositions, although they are not complete macropropositions in themselves. Topic sentences at the beginning of paragraphs can signal macrorelevant information.

As already mentioned, Lorch, Lorch and Matthews (1985) showed that topic sentences receive longer processing than non-topic sentences, and that the provision of an introductory overview or abstract facilitates the processing of topic sentences. With regard to initial mention as a source of information, Kieras (1980) showed that a statement of the text's theme has a greater effect on readers' overall interpretation when the thematic statement is placed at the beginning rather than in the middle of the text.

Another important source of information according to van Dijk and Kintsch (1983) is knowledge of conventional text structures, the schemata of which are referred to as superstructures. Van Dijk and Kintsch (1983, p. 16) describe the role of these superstructures in the formation of macropropositions as follows:

> *"Language users manipulate superstructures in a strategic way. They will try to activate a relevant superstructure from semantic memory as soon as the context or the type of text suggest a first one. From then on, the schema may be used as a powerful top-down processing device for the assignment of relevant superstructure categories (global functions) to each macroproposition - or sequences of macropropositions - and will at the same time provide some general constraints upon the possible local and global meanings of the textbase."*

Examples of the texts which have clear and conventional structures, and hence for which readers may have available superstructures, include simple stories and

psychological reports. Kintsch and Yarborough (1982) set out what they refer to as rhetorical schemata, including the forms of argument, comparison and contrast, definition, and procedural description, which are, like narrative superstructures, used to organise information from the text. Hence, a text organised according to a familiar rhetorical schema should be easier to comprehend because the schema facilitates macrostructure formation. This claim is supported by results in the Kintsch and Yarborough study, where subjects read texts with either good or bad rhetorical form, the latter created by reordering paragraphs. Subjects gave significantly better responses to macrostructure questions (What was the text about? What were the main ideas?) after reading texts with good rather than poor rhetorical form. However, this form had no significant effect on cloze test performance, which Kintsch and Yarborough see as a measure of microprocessing only. It appears that the effect of rhetorical form was limited to macroprocessing.

Meyer has also investigated how readers use the top-level structure of a text to determine main ideas. Meyer, Brandt and Bluth (1980, p. 78) describe this process as follows in relation to a passage about the problem of oil spills:

> *"Each new proposition is related back to the problem of oil spills from supertankers continually selected for retention in the short-term buffer for interpreting the new propositions. Previous knowledge about problems keeps the reader searching for causal relationships among descriptive information about the problem and what caused it. Also prior experience with problems leads readers to anticipate and search for solutions, solutions which must satisfy most of the previously stated causes of the problem. Thus, the problem and its causes are retrieved continuously from long-term storage to the short-term buffer for re-*

lating to the subsequent propositions in the passage. This additional processing of these superordinate propositions and their interrelationships increases the depth with which they are processed and the ease with which they can be retrieved."

Knowing that the text in this example has a problem-solution structure, and setting out to instantiate this general structure with specific information from the text, should assist the reader in forming a macrostructure.

Just and Carpenter (1980) present a similar approach to what they call text roles. They assume that text segments play one of a limited number of roles in scientific texts, these being topic, subtopic, definition, cause, consequence, expansion and detail or example. They hypothesised that experienced readers would be looking for important text segments which contain the key points of a passage. Analysis of gaze durations while reading showed that readers did indeed spend most time on topics, definitions, causes and consequences, and least time on details. The probability of recall was also highest for the former text roles. Just and Carpenter (1980, p. 344) explain this finding as follows:

"Text roles that are usually more important to the text and to the reader's goals, such as topics or definitions, are integrated differently than less important units, such as details. The more central units will initiate more retrievals of relevant previous knowledge of the domain (schematic knowledge) and retrievals of information acquired from the text, but no longer resident in the working memory. In addition, more relations will be computed between the semantically central propositions and previous information because centrality inherently entails

relations with many other units. By contrast, details are often less important to the reader's goals and to the text. Moreover when a detail is to be integrated, the process is simpler because details are often concrete instantiations of an immediately preceding statement (at least in these scientific texts), so they can be quickly appended to information still present in the working memory. Thus, higher level units will take more time to integrate because their integration is usually essential to the reader's goals, and because integration of higher level units involves more relations to be computed and more retrievals to be made."

As mentioned in relation to Meyer, the reader who has knowledge of how a particular text type is conventionally organised can use that knowledge to search for the semantically central information which constitutes the macrostructure.

However, there are two main limitations on the use of text structure to guide macroprocessing. The first is that natural texts rarely occur in a 'pure' form. For example, Just and Carpenter (1987, p. 248) state: *"Sometimes the different types of texts are divided by no more than a fine line."* Kintsch and Yarborough (1982, p. 834) make the same point in more detail when they comment as follows:

"Most essays are composed of several different, often complexly related rhetorical units. Thus, a definition might be embedded in a procedural description, which in turn may be part of a comparison and contrast scheme. There seems to be no such thing as an 'essay grammar'; instead, a writer can use the various rhetorical units with great freedom. Nevertheless, recognizing such units for what they are

whenever they are encountered as parts of a larger text may still be very helpful in controlling and guiding the macroprocesses."

The second limitation is that knowledge of text structure is unlikely to be useful in the absence of adequate subject knowledge. For example, although a text may signal a problem-solution structure, the reader has to have sufficient knowledge of the topic to understand how the problem and proposed solutions relate to each other in order to be able to instantiate the general structure. Roller (1990, p. 86) supports this conclusion when she states:

"Text structure variables operate differently depending on the extent of the reader's knowledge of the topic of the text. Because structure variables highlight or make explicit the relations between ideas in the text, they exert their influence only in moderately unfamiliar text. In this situation, readers know enough about the topic to isolate concepts and to use structural cues to construct the relations between those concepts. However, when the topic covers very familiar topics, structural cues become redundant, because readers already understand the relations between concepts. Structural cues are also unhelpful in text about extremely unfamiliar topics, because readers have so much difficulty isolating the basic concepts that they cannot begin to sort out the relations between them."

While many sources of information, such as headings, topic sentences, initial mention, and knowledge of text structures can all be useful in constructing a macrostructure, in all cases the reader needs to have sufficient subject knowledge to identify and make use of this potentially helpful information.

Factors influencing macrostructure formation

It has already been mentioned that knowledge about the subject of a passage is one important factor influencing the nature and extensiveness of the macrostructure formed. Other factors to be discussed in this section include the working memory capacity available for macroprocessing, and the reader's goals and perceived task demands.

Van Dijk and Kintsch (1983, p. 201-2) draw attention to the importance of subject knowledge in the formation of a macrostructure when they suggest that: *"As language users will bring to bear different sets of knowledge, beliefs, etc, on the interpretation, it is likely that different macrostructures may be derived."* The authors consider that such individual variation will be found mainly in the type of macrostructures generated, but since macropropositions also guide local processing, some of this variation would also be expected to carry over into the nature of the textbase constructed.

An example of variations in knowledge leading to different interpretations as to what a text is about can be found in Anderson, Reynolds, Schallert and Goetz (1977). In this study, music and physical education students read two ambiguous passages. One could be interpreted as an escape either from prison or a wrestling hold, while the other could be interpreted as playing either music or cards with friends. Anderson et al hypothesised that music students would interpret the passages as a prison break and a musical evening respectively. The physical education students, on the other hand, were expected to produce interpretations of a wrestling match and a card game. The results supported these hypotheses, indicating that background knowledge influenced the construction of a theme for the texts.

Afflerbach (1990) examined not only the final output of main ideas, but also the process of generating topic statements. He analysed think-aloud protocols collected from doctoral students in Chemistry and Anthropology as they read passages varying in familiarity relative to their subject specialisation. These readers used different strategies to generate main ideas depending on how familiar the topic of the text was. When the topic was very familiar, main ideas were either constructed automatically, with the reader having no insight into or conscious awareness of the construction process, or the reader used what Afflerbach calls the Initial Hypothesis Strategy. In this case the reader confidently generated a reasonable hypothesis as to the main idea of a paragraph soon after beginning reading and then monitored the accuracy of the hypothesis while reading on.

When the topic was unfamiliar, readers used either the Draft and Revision or the Listing Strategy. In the former, the reader put forward a tentative main idea statement and then proceeded to progressively revise it, while in the latter, the reader searched for and listed important or related words to form the basis for a main idea statement.

Another factor which can influence macrostructure formation is the working memory capacity available for such macroprocessing. Even among skilled readers, a text which is difficult to process may not allow macropropositions to be generated on first reading. Van Dijk and Kintsch (1983) refer to a study, published in German in 1981 by Schnotz, Ballstaedt and Mandl, in which subjects attempted to summarise a difficult text immediately after reading it. The first attempt resembled a recall protocol rather than a summary, with little differentiation between micro- and macropropositions. On the second attempt, however, subjects produced summaries containing mainly macrorelevant information. In

relation to the difficult material used in this study, van Dijk and Kintsch (1983, p. 53) conclude:

> *"There is at least a suggestion here that in the early stages of processing (i.e., during first readings) readers are concerned mainly with forming a coherent textbase and the local inferences involved in that process; reductive inferences may be postponed for later."*

Presumably in the study by Schnotz et al, local processing was so resource demanding that global processing was, of necessity, deferred. However, in an easier text, sufficient processing resources should be available to skilled readers to process micro- and macropropositions concurrently. Van Dijk and Kintsch also speculate that global coherence strategies may be learned relatively late compared to, for example, proposition formation strategies. Thus, young children may not form macrostructures either because they have not learned the required strategies (and possibly do not have the knowledge to support such strategies), or because the formation of a textbase is still too resource demanding to allow macropropositions to be constructed concurrently.

One factor which may reduce the processing demands involved in producing a macrostructure is a 'considerate' text. For example, Lorch, Lorch and Matthews (1985) showed that processing of topic information was faster when a passage contained an introductory overview of the topic structure. Lorch, Lorch and Inman (1993) showed that topic information was also better recalled when topic structure was signalled by overviews, summaries, headings and indicators of subsection divisions. Similar findings are reported by Lorch and Lorch (1985), using topic sentences as signals, and Loman and Mayer (1983) using preview sentences,

headings and logical connectives to signal the text structure. Presumably the signalling in these studies reduced the processing demands required to construct a macrostructure.

Finally, the goals set by readers and the way they interpret task demands can determine if and how a macrostructure is formed. Kintsch and van Dijk (1978), as an example, contrast the reader who has the goal of constructing the main sequence of events in a story with another reader who uses the story to gather evidence of social attitudes during a particular time period. Due to these different goals, these readers would make different judgements about what is most important in the text. The reader's goals can also determine the depth of processing, as in the case of readers with the goal of skimming newspaper reports and constructing an adequate macrostructure, but an incomplete textbase. Presumably, though, this situation could occur only for texts with a relatively high degree of predictability.

As a further example, Pichert and Anderson (1977) deliberately manipulated the reading goals of their subjects and hypothesised that information in a text would be judged more or less important, depending on the goals of particular readers. One experimental text described two boys skipping school and exploring the home of one of the boys while his parents were out. Some subjects read the text normally, while some were told to read it from the point of view of a burglar, others from the point of view of a home buyer. In a first experiment subjects rated the importance of each idea unit from the point of view of a particular perspective, while in a second experiment another group of subjects recalled the text after reading it from one of the three perspectives.

Correlations between importance ratings given to idea units by the three dif-

ferent groups were very low, indicating that the importance of an idea is not predetermined by the text, but depends on the knowledge and goals the reader brings to the interpretation. In free recall, almost half the idea units rated as high in importance were produced, compared to only one quarter of the idea units rated as low in importance. Thus, the readers' goals, what they intended to gain from the text, influenced the judgement of importance which is central to macrostructure formation.

Graesser, Hoffman and Clark (1980) examined time allocated to different aspects of text processing when university subjects were told to read either in preparation for writing an essay or for answering multiple choice questions. This difference in reading goals influenced macrostructure but not microstructure analyses, leading to the conclusion that: *"When the reader attempts to form a cohesive, well-organized, and recallable representation of a passage, the reader allocates more resources to macrostructure analyses than in conditions where the reader merely wants to comprehend the contents of a passage."* (p. 146)

As noted previously, McKoon and Ratcliff (1992) in presenting their minimalist hypothesis of processing emphasise the importance of the reader's goals. They propose that forming a macrostructure and using this global information to guide processing is a strategic activity which depends on specific reading goals focused on learning from the text.

In terms of such learning, the study by Loman and Mayer (1983) is particularly significant in showing that a coherent representation of the central concepts in a text is important for tasks such as problem solving based on newly acquired information. Forming a macrostructure of the main ideas is therefore an important process in this type of reading. However, readers need to first be aware that

such understanding of main ideas is an important goal when studying expository text, and they need to have available strategies for achieving this goal. This may be a problem for children accustomed to reading mostly simple narratives. They may not have learned to set goals and select strategies for studying expository texts, unlike the adults in the study by Lorch, Lorch and Klusewitz (1993), who reported using quite different strategies for studying as opposed to 'light' reading of novels and magazines. Children may still need to learn to make use of macrorelevant information such as headings and topic sentences, and to infer such information if it is not explicitly stated.

1.4 The role of knowledge in processing text

In the previous sections, reference was made on several occasions to the effect on processing of a reader's goals and level of knowledge about the topic of the text. In these final sections, these aspects will be discussed in more detail. Firstly, subject knowledge will be examined, in particular, how such knowledge is organised in order to enable contextually sensitive interpretations to be made. Secondly, reading goals will be examined under the heading of metacognitive knowledge. Theorising on this subject will be discussed and related to the model of reading presented thus far.

1.4.1 Subject knowledge

In order to construct contextually appropriate word meanings, bridging inferences, macropropositions, and so on, readers must be able to rapidly access a well organised knowledge base. On the one hand, readers have to be able to generate all the necessary options for interpreting a text from this knowledge base, while on the other hand the number of options has to be limited so as not to overload working memory and the processing system. Schema theory was one approach to this construction of correct options; the schema as a hierarchically organised packet of information with slots in which to instantiate the component variables was assumed to guide interpretation and to allow predictions and inferences to be generated. Rumelhart and Ortony (1977, p. 111-2) offer one such view of comprehension:

> *"Schemata are the key units of the comprehension process. Within the general framework presented here, comprehension can be consid-*

ered to consist of selecting schemata and variable bindings that will 'account for' the material to be comprehended, and then verifying that those schemata do indeed account for it. We say that a schema 'accounts for' a situation whenever that situation can be interpreted as an instance of the concept the schema represents. Thus, the bulk of the processing in a schema-based system is directed toward finding those schemata which best account for the totality of the incoming information. ... Thus, in general, the process of comprehension can be regarded rather like the process a scientist goes through in testing a theory; evidence is sought which either tends to confirm it, or which leads to its rejection. Upon finding a theory which, to our satisfaction, accounts for the observations we have made, we feel that we understand the phenomenon in question."

Kintsch (1988) is one author who has been critical of this view of comprehension and knowledge organisation. He notes that if such a system is powerful enough to constrain processing and to limit the number of options generated, then it will not be flexible enough to deal with the variety and subtle differences in the many contexts that humans are capable of comprehending.

Kintsch (1988) proposes instead that the concepts and propositions which constitute a reader's knowledge form an associative net. The concepts and propositions are nodes, connected to related nodes by links with varying strength values. Within this system, larger semantic structures or full contextual word meanings are not predetermined, but are generated as the need arises through the activation of particular subsets of the net. The system is flexible enough to produce context-sensitive interpretations, as explained by Kintsch (1988, p. 165) with regard to

conceptual meanings:

> *"Concepts are not defined in a knowledge net, but their meaning can be constructed from their position in the net. The immediate associates and semantic neighbors of a node constitute its core meaning. Its complete and full meaning, however, can be obtained only by exploring its relation to all other nodes in the net. ... It is not possible to deal with the whole huge knowledge net at once. Instead, at any moment only a tiny fraction of the net can be activated, and only those propositions of the net that are actually activated can affect the meaning of a given concept. Thus, the meaning of a concept is always situation specific and context dependent. It is necessarily incomplete and unstable. Additional nodes could always be added to the activated subnet constituting the momentary meaning of a concept, but at the cost of losing some of the already activated nodes."*

According to Kintsch, input from a text activates a subset of this net, and the integration processes then operate on these activated nodes. These integration processes, linking the new input to the text already processed, will determine which of the activated concepts and propositions are relevant in the particular context. Thus, the system will generate all the options required because a whole subset of the knowledge net is activated, rather than only particular concepts and propositions, but interpretations will also be contextually appropriate because they are constrained, loosely by the text input and more specifically by the integration processes.

Such a view is not inconsistent with the position of Just and Carpenter described earlier, although these authors do not discuss knowledge organisation in

as much detail. As was pointed out in the section on lexical access, for example, Just and Carpenter follow Collins and Loftus (1975) in assuming that concepts are stored as a semantic network. Of concepts inserted into working memory in the course of processing Just and Carpenter (1980, p. 338) state: *"The concept serves as a pointer to a more complete representation of the meaning, which consists of a small semantic network ...."*

Similarly, the approach adopted by Kintsch (1988) is consistent with revised views of conceptual structure and of even apparently simple knowledge structures such as scripts.

With regard to the categorisation of concepts, Medin and Smith (1984) point out that category membership is not usually determined by reference to a definition stating necessary and sufficient conditions. If membership were determined by definitions, then there should not be any unclear cases, such as whether or not *rug* belongs in the *furniture* category. Also, all members should have the same status as category members, since all share the same set of necessary and sufficient properties. However, Rosch (1973) shows that this is not the case; some exemplars are consistently rated as being more typical of a category than others.

To overcome such problems, Rosch and Mervis (1975) put forward a prototype model of category structure, in which categories are networks of interrelated attributes. How typical an exemplar is depends on the extent to which it has attributes overlapping those of other concepts in the category. Thus, there are no necessary and sufficient conditions, and the 'edges' of the category are blurred.

As Medin and Shoben (1988) note, this last quality points to a problem for prototype models; there are no constraints on the attributes which are relevant to categorisation, and it is difficult to determine any basis for the psychological

cohesion of the category. Medin and Shoben argue that concepts should not be classified only in terms of a set of attributes, but also in terms of the relationships they enter into, especially structure-function and cause-effect relations.

This point is expanded by Murphy and Medin (1985, p. 290), who argue that attributes relevant to a particular manner of categorisation are chosen on the basis of more general theories, with a theory described as *"a complex set of relations between concepts, usually with a causal basis"*. Thus, categories are seen as more flexible and context-specific than was the case in the traditional categorisation by definition. Using Murphy and Medin's approach, it is also possible to create ad hoc categories as the need arises, for example, the category of *things to remove in case of fire* or the category of *things to buy for birthday presents.* This view is in keeping with Kintsch's assertion that larger semantic structures are generated through activation of particular subsets of the associative net.

As models of conceptual structure have moved away from a rigid and clearly defined form, so too have notions of scripts and schemata. In early versions of schema theory, as presented by Anderson (1980) for example, the interpretation of a text involved selecting appropriate schemata and instantiating their slots with information from the text or with default values specified by the schemata themselves. Schemata were, thus, relatively fixed and prestored structures.

As a result of the limitations in the original model of schemata, the conception of even apparently simple scripts, the schemata of stereotyped human activities (Schank and Abelson, 1977), has undergone revision. Schank (1982) suggests that scripts do not exist in precompiled form in memory. Rather, general organising structures select appropriate scenes for a particular script, such as going to the dentist, and the scenes are given particular colourations to take account of the

specifics of a given situation. Thus, Schank sees scripts as being compiled on demand and in a context-sensitive manner, in much the same way that Murphy and Medin (1985) see general theories as organising categories. Van Dijk and Kintsch (1983, p. 30) point out the similarity in the evolution of conceptual structures and scripts:

> *"Indeed scripts are only too similar to our familiar categories. Like them they cannot be defined precisely, but are fuzzy by their very nature; they cannot be used in simple, fixed ways but are highly context dependent ... Most of the difficulties in designing knowledge systems derive from this flexibility that concepts, scripts and schemata share. We have seen the shift that took place within psychology in the last decades in the way we think about concepts: From the neat, clean definitions via necessary and sufficient properties we have moved to context dependent feature sets and exemplars. Scripts (and with them schemata and frames) have undergone very much the same evolution, having changed from fixed, prepared structures to context-sensitive entities that are constructed in response to some task demand."*

Given the preceding description of knowledge organisation, it is apparent that not only the amount of knowledge possessed by an individual in terms of the number of concepts and propositions, but also the way this knowledge is interconnected is very important in determining the interpretation given to a text. The reader with a rich and well organised knowledge base is in a better position to generate contextually appropriate word meanings and elaborations, to integrate text elements in coherent ways, including through inferences, and also to

UNIVERSITY OF ADELAIDE LIBRARY

determine important elements and the thematic unity of the whole text in the formation of a macrostructure. These points are demonstrated empirically in several studies.

Lipson (1983) examined reading performance when texts described culturally familiar and unfamiliar events. Children in grades 4 to 6 in Jewish and Catholic schools read three passages. One was culturally neutral, one described a Bar Mitzvah, and one described a First Communion. When analysing data, performance on the neutral passage was used as a covariate for each dependent measure. When reading culturally familiar material, children read more quickly, produced both more explicit and inferred information in free recall, made fewer errors, and answered correctly more questions tapping text-explicit information. Lipson notes that children showed evidence of confusion about culturally unfamiliar passages. Some confused responses involved Jewish people wearing a tortilla, and Catholics eating bread left behind for them by Jesus.

Pearson, Hansen and Gordon (1979) devised a test for children in grade 2 to assess background knowledge about spiders. Questions tapped information related to the passage, but not explicitly stated within it. For example, one knowledge question asked why oil was used to fry an egg, on the assumption that a general knowledge of the lubricating function of oil would help children to understand why spiders did not stick to their own web, as explained in the text. After reading the passage, children answered two sets of questions, one requiring recall of information stated explicitly in the text, the other requiring an elaboration of explicit information based on general knowledge. Overall, high-knowledge students answered more questions correctly. The difference between the high- and low-knowledge groups was significant for implicit questions, but

failed to reach significance for explicit questions, although the difference was in the same direction. The two groups were not significantly different in terms of IQ or reading ability, as measured by the Metropolitan Achievement Test; these students were above average on both tests.

Several studies examining the effect of baseball knowledge on comprehension of texts on the same subject indicate that knowledge increases the ability of readers to determine important ideas and to organise information around those ideas.

Chiesi, Spilich and Voss (1979, expt 3) found that recall of short baseball descriptions was higher among individuals with high knowledge (HK) than those with low knowledge (LK). That this difference was due to level of baseball knowledge rather than an overall memory advantage is shown by the fact that recall of descriptions of city scenes was higher among LK than HK readers. More importantly, though, Chiesi et al (expt 5) suggest a reason why a higher level of subject knowledge leads to a higher recall of text. In this experiment, subjects were asked to recall a set of target sentences which were preceded either by 0, 1 or 2 context sentences to form a logical sequence of baseball related information. Recall of sentences presented with 0 context sentences did not vary with level of knowledge, but the recall of HK individuals increased with increasing numbers of context sentences, while that of LK individuals decreased. For the LK individuals, increasing the number of context sentences had a similar effect to increasing the length of a list to be learned, that is, it imposed a greater memory burden. For HK individuals, on the other hand, increasing the number of context sentences resulted in a more integrated text representation, which was easier to recall. As Chiesi et al (1979, p. 271) note:

> *"HK and LK performance differences were due to the superior ability of the HK individuals to relate successive game actions and changes in game state to the game's goal structure ... The successive actions constituting a game 'episode' are interpreted in terms of the goal structure and related to each other as part of the episode."*

HK individuals are assumed, as a result of rapid access to well organised knowledge, to be able to recognise patterns, in the above case to organise sequences into recognisable baseball 'episodes' which are easily recalled.

A similar conclusion regarding the organisation of text representations formed by HK and LK individuals is reached by Spilich, Vesonder, Chiesi and Voss (1979), where subjects listened to and recalled a 5 minute long description of a baseball game. Not only did HK individuals recall more propositions than LK individuals, but as Spilich et al (1979, p. 284) point out, the nature and structure of information recalled was also different:

> *"The HK subjects recalled more information that was goal related than the LK subjects, they tended to integrate sequences of goal-related actions more than LK individuals, and they recalled information in the appropriate order more than LK individuals. Setting information of possible importance to the goal structure was generally recalled more frequently by the HK individuals, but information and actions not related to the goal structure were recalled more frequently by LK individuals. If one were to characterize the HK and LK differences in general terms, LK recall was like that of a free recall list, excluding a recency effect, in which primacy information was recalled and other*

specific items were also recalled. There was some organization in LK protocols, but the omissions often showed large gaps in the organization and the confusions suggested LK subjects had considerable trouble organizing information sequentially. For HK subjects recall was more coherent both with respect to the sequential nature of the account and to the completeness of description of important events."

Recht and Leslie (1988) modified the materials used by Spilich et al (1979) for children in grades 7 and 8. High-knowledge readers showed higher levels of and more accurate recall, with their recall organised more like that of adult experts than was the case with low-knowledge readers. When asked to produce a summary of the passage and to rate the importance of a list of statements from the text, high-knowledge children again performed more like experts than did low-knowledge children. The performance of children with high and low reading ability was affected to a similar degree by subject knowledge.

Similar findings are also reported by Schneider, Koerkel and Weinert (1989) from their study of children varying in soccer knowledge. HK individuals recalled more details, detected more errors embedded in the text, and answered more inference questions correctly. Another important finding was that HK children were more able to underline the ten most important sentences in the text, as determined by twenty adult soccer experts, than LK children.

These studies indicate that a high level of subject knowledge can influence many aspects of processing, including the ability to organise information into coherent episodes and to interrelate them at a global level. The studies consistently show that high-knowledge individuals are more able to determine which ideas are most important to the overall topic of the text. Fincher-Kiefer, Post, Greene

and Voss (1988, p. 426) also conclude that the effect of subject knowledge is not so much on the construction of individual propositions as on their organisation. *"... domain knowledge enables the individual to organize relatively large amounts of domain-related text, and via such organization retrieval structures are developed."* Clearly, then, an instructional programme needs to consider carefully the students' level of knowledge about the topics in the texts to be used to enable them to apply the strategies being taught.

1.4.2 Metacognitive knowledge

Flavell (1981, p. 37) refers to metacognition in broad terms as *"knowledge or cognition that takes as its object or regulates any aspect of any cognitive endeavour."* In this section it is not possible to discuss in detail the substantial literature on the subject. Rather, the aim is to focus on three main points related directly to the models of reading already discussed and the comprehension training programme to be developed further below.

1. Both Just and Carpenter (1980) and van Dijk and Kintsch (1983) propose that the nature and depth of processing is determined by the reader's goals. Thus, the question here is how this aspect of their respective theories can be related to wider metacognitive theorising.

2. The second point concerns the relationship between metacognitive and cognitive processes in the course of reading. It is widely acknowledged that people do not always use the cognitive skills available to them, in part because they do not have adequate metacognitive knowledge about the usefulness of such skills in different contexts. A question to be addressed here is whether the reverse also applies, that is, whether there are also cognitive reasons why readers are not able to apply the metacognitive knowledge available to them.

3. Finally, it will be useful to consider how metacognitive skills develop, and why they should be included in comprehension training programmes.

Metacognition and its relationship to reading

Brown (1980, p. 454-5) describes the influence of a reader's goals on processing in much the same way as Just and Carpenter (1980).

> *"The goal of reading is to achieve understanding. Yet understanding is not an all or none phenomenon; it must be judged against the criteria set by the reader as a goal of the activity. Readers' purposes vary and, as such, criteria of comprehension also change as a function of the particular reading task at hand. In some reading situations, the participant may be quite satisfied with gleaning a cursory overview of the gist, whereas in others the reader may set more demanding criteria of comprehension. The decision to process deeply and actively or merely to skim the surface will determine not only the strategies necessary for the task but also the reader's tolerance for intrusive feelings of failing to understand. In short, the reader's purpose determines how he or she sets about reading and how closely he or she monitors the purpose of reading, that is, understanding of the text."*

Flavell (1981) also acknowledges the influence of goals on processing. Cognitive goals are one of the four aspects of his discussion of metacognition, the others being cognitive actions, metacognitive knowledge and metacognitive experience. Of cognitive goals Flavell (1981, p. 40) says:

> *"Cognitive goals are the tacit or explicit objectives that integrate and maintain the cognitive enterprise. Goals naturally vary from one enterprise to another and can also change in the course of any given enterprise."*

Cognitive goals have a large influence on the selection of appropriate cognitive actions. Flavell points out that there is a dual aspect to the category of cognitive actions; some actions make progress towards cognitive goals, while others monitor that progress. In some cases the same action can fulfill both these functions. For example, readers preparing to recall the gist of a passage may decide after each paragraph to mentally review and summarise the facts presented. This one action has the effect both of encouraging integration of information and organising it for subsequent recall (a cognitive function), as well as allowing readers to check their understanding according to how well they can review and summarise (a metacognitive function). Thus, this cognitive action can give rise to metacognitive experiences, that is, an awareness of the extent to which the goal is being achieved.

Flavell (1981) includes in the category of cognitive experiences a wide range of thoughts and feelings associated with progress (or lack of it) being made in the task at hand. In relation to such a task *"we may currently experience the lack of some needed talent, puzzlement about some task-relevant information, an awareness of the difficulty of the task requirements, an inability to think of a means to a goal, or the opposite of these feelings."* (p. 48)

Metacognitive knowledge is similar to any other world knowledge, except that it focuses on the nature of and relations between human capabilities, task requirements, and strategies, aspects which Flavell refers to as person, task and strategy variables respectively. Person variables include beliefs about one's own processing efficiency, how one compares to other people, and how people in general function. Task variables include information about different materials and tasks on processing difficulty. The strategy variable includes information about the utility of various cognitive actions, or strategies, in a variety of contexts.

Clearly these three variables are interdependent; in order to determine appropriate strategies, readers have to evaluate their own capabilities in relation to particular task demands.

Jacobs and Paris (1987) divide metacognition into two broad categories, namely, self-management and self-regulation. Self-management includes the person, task and strategy knowledge discussed by Flavell (1981), including how, when and why to use particular strategies. Self-regulation involves planning, evaluation and regulation, that is, readers select cognitive actions to achieve a goal, they monitor progress towards the goal, and adjust strategies in relation to difficulties or changing task demands. This self-regulation category encompasses Flavell's cognitive goals and actions. Baker and Brown (1984a) adopt a similar categorisation to that of Jacobs and Paris.

Another model of metacognition is put forward by Borkowski and Kurtz (1987), one which, as will be shown later, is particularly useful for charting the development of metacognitive skills. The authors recognise four components in their model, these being Specific Strategy Knowledge, Relational Strategy Knowledge, General Strategy Knowledge and Metacognitive Acquisition Procedures.

In the context of memory tasks, Specific Strategy Knowledge involves knowledge of strategies such as rehearsal or organisation to enhance recall. From Specific Strategy Knowledge, Relational Strategy Knowledge and General Strategy Knowledge gradually develop. Relational Strategy Knowledge involves an evaluation of the relative merits and difficulties of particular strategies. General Strategy Knowledge is based on the understanding that a strategic approach to a task usually results in better performance, and that effort is required to apply strategies. This knowledge is abstracted over time from experience with using spe-

cific strategies. This category also incorporates a motivational component, that is, beliefs about self-efficacy and the attributional style of the learner. Learners who believe that they lack ability and that their achievement outcomes are not contingent upon effort will very likely not see any point in making the effort required to apply strategies. Finally, Metacognitive Acquisition Procedures involve knowledge of how and when to use known strategies, how to modify strategies, and how to adopt different strategies when problems arise. In other words, the procedures have a regulatory, executive function.

While these various approaches use different terminology and different categorisations of aspects of metacognition, they do have underlying elements in common. All recognise some form of declarative knowledge or belief about one's own ability and potential, about a range of strategies, including how and when they apply, and an understanding of task demands and goal setting. They also include a procedural component in which the above knowledge is applied to set a goal appropriate to the task demands, to select actions to achieve this goal, to monitor progress and take remedial action if necessary.

As already noted, goal setting is acknowledged as important by Just and Carpenter (1980, p. 350) when they comment: *"Goals can be represented in several aspects of the theory, but the main way is to require that each goal is satisfied or at least attempted before proceeding on to the next word, clause or sentence."* Careful readers attempting to learn from an expository text, for example, may set themselves the goal of understanding the text such that they can explain it to someone else or answer any questions about it. This goal, then, would influence many aspects of cognitive processing. For example, these careful readers would attempt to construct an adequate and textually consistent meaning for

each concept in the text. In other words, if a concept could not be related to other concepts already in memory, or if a meaning could not be integrated with previously processed text, these readers would be likely to take remedial action. In the case of relatively simple ambiguous words, Carpenter and Daneman (1981) show evidence of such monitoring and recovery from inconsistent interpretations through increased processing time and regressive eye movements. With more complex concepts other strategies may need to be called upon, such as rereading to construct a meaning from information given in the text, or consulting outside references. A less careful reader with different goals may well be satisfied with less complete understanding of each word concept.

The readers' goals also influence the criteria of coherence adopted, and hence the extensiveness and frequency of wrap-up operations. The careful readers assumed above would attempt to integrate each idea unit into their text representation, and hence would be likely to construct more bridging and even elaborative inferences in order to be able to do so. Having set stringent criteria of coherence, they would also be more likely to detect errors or omissions in the text and to be aware of lapses in their own concentration or understanding. Again, skilled and careful readers are likely to draw on knowledge of possible remedial actions in such circumstances. Readers with different goals, such as merely skim reading a text, would be satisfied with less frequent and less extensive integration.

Kintsch and van Dijk (1978) and van Dijk and Kintsch (1983) acknowledge the importance of the readers' goals in the formation of a macrostructure, the goals adopted influencing whether one will be formed at all, as well as its nature and degree of clarity. The careful readers referred to above may know that condensing a text into its gist is a way of aiding recall by providing a framework for

more detailed information, hence they adopt the strategy of devoting additional processing to macrorelevant information such as headings and topic sentences, and where necessary construct their own summary statements. Where the goals of the readers are less well defined, the macrostructure created is also likely to be vague, and such readers are likely to have more difficulty in assessing the thematic consistency of the text and their own level of understanding. As indicated previously, McKoon and Ratcliff (1992) also emphasise the importance of the reader's goals to the formation of a macrostructure.

In contrast to skilled and careful readers, children who are poor readers may set inadequate reading goals, such as merely pronouncing all words correctly and fluently (Garner and Kraus, 1981-2). As indicated above, such a goal would then influence the execution of all processes from lexical access to intersentence integration to macrostructure formation, as well as the monitoring of these processes.

The preceding discussion suggests that the self-regulation processes of planning, evaluation and regulation (Jacobs and Paris, 1987) can be accommodated within the models of reading proposed by Just and Carpenter and Kintsch. By implication, then, the declarative components of metacognition can also be accommodated, since self-regulation processes draw on what Flavell (1981) refers to as metacognitive knowledge and what Jacobs and Paris (1987) refer to as self-management skill, that is, knowing about various task demands and strategies appropriate for meeting these demands.

Limitations on the use of metacognitive knowledge

Much has been written on the importance of metacognitive knowledge and its use in the regulation of cognitive processing (see for example Paris, Wasik and

van der Westhuizen, 1987). There is a danger in the enthusiasm generated by this construct that poor reading achievement may be too hastily attributed to poor metacognitive skills or, as Jacobs and Paris (1987, p. 264) state: *"the construct can take on an undeserved ubiquity and executive power in explanations of cognitive processing."*

Markman (1981, p. 75) reflects a similar concern when she points out that much information about one's level of understanding arises directly from careful cognitive processing, rather than any conscious metacognitive monitoring.

> *"Note that it is not always necessary to have an explicit question 'Do I understand?' in order to obtain information about one's own understanding. In the process of attempting to understand, one can obtain information about how well one does understand. Metacognitive knowledge and abilities probably play an important role in all aspects of comprehension. Active attempts to comprehend, judgements about materials and task demands, judgements about potential explanations, etc all require some metacognitive knowledge. Without minimizing the role that such metacognition plays, I would like to emphasize that much information about one's comprehension is a by-product of active attempts to understand and not just of attempts to monitor."*

If readers are constructing meanings for all words and are able to carry out the processing needed to integrate each piece of information, they may not be aware of any kind of metacognitive experience. However, if the smooth operation of this processing is interrupted, the problem may reach consciousness as a metacognitive experience of puzzlement. As Baker (1985a, p. 168) says: *"readers are typically*

unaware of the processing operations they carry out to establish cohesiveness; it is only when they are unable to integrate the text that their efforts become conscious." Baker and Brown (1984b) and Flavell (1981) also agree that many metacognitive experiences arise from the application of cognitive actions aimed at producing comprehension, rather than from overt self-questioning.

Flavell (1981) suggests that the more complex forms of metacognitive experience, that is, knowing *why* as well as *that* you are confused, can arise only when there is sufficient working memory capacity available, since such complex experiences are very resource-demanding. Markman (1985) also draws attention to the importance of working memory capacity in metacognitive experiences. As already noted, she sees metacognitive experiences as arising from cognitive processing, especially the integration processes. However, if a reader's working memory capacity is exceeded during integration due to too much attention being required by the processing of smaller units, that reader will not be in a position to judge the consistency of the information being processed.

Markman (1985, p. 276) describes the capacity demands of various forms of comprehension monitoring as follows:

> *"Some types of comprehension failure can be discovered by processing small segments of material. An unfamiliar word can often be detected even in complete isolation from the remainder of the passage. One can also notice whether a single isolated sentence is confusing, ungrammatical or false. In contrast, understanding an argument, a story or expository prose requires more than understanding the individual words or sentences. To notice problems in comprehension thus depends on some difficulty discovered in integrating or relating one*

> *part of a passage to another. This requires inferential or constructive processing - the transforming, elaborating, expanding of incoming information. The difficulty of discovering an inconsistency should vary with the amount of processing required... Several studies with young children have demonstrated that, as information-processing demands decrease, children become better able to notice problems in their comprehension."*

Markman's comments here relate closely to the point made by Daneman and Carpenter (1983), namely, that the ability to monitor and repair comprehension depends to a large extent on how much working memory capacity is available as cognitive processing is taking place. Readers gave fewer correct responses to a probe question when the contextually primed meaning of an ambiguous word was inconsistent with subsequent information. The differences in percentage of correct responses between the consistent and inconsistent conditions were 52%, 40% and 22% for small, medium and large reading span subjects respectively, with reading span being used as a measure of functional working memory capacity. Thus, metacognitive and cognitive aspects of reading should not be considered separately, but need to be treated as part of one system, sharing limited resources.

According to Markman (1981), one factor which influences processing efficiency, and hence working memory capacity available for monitoring, is knowledge relevant to the topic of a text. She considers children as novices with regard to the subject of many texts, and points out some differences in knowledge organisation between experts and novices. Not only do experts know more facts than novices, who may hold vague or even erroneous beliefs, but the knowledge of experts is also more structured. They know more general principles to connect

discrete facts, and as a result are able to see more patterns in a body of information. Brown (1978, p. 142) points out how a deficient knowledge base relates to reading skill and working memory capacity.

> *"... the child's knowledge base is deficient in at least three ways: 1) the amount of information it contains; 2) the organization and internal coherence of that information; and 3) the number of available routes by which it can be reached. These differences impose severe limitations on the child's information processing abilities, even in such simple situations as reading information in STM. Such basic cognitive processes as ease of retrievability and speed of encoding, naming, and recognition are all influenced by restrictions imposed by an impoverished knowledge base."*

As a result of their processing efficiency, expert readers, in terms of subject knowledge, should not only be able to comprehend a text better, but should also be able to monitor comprehension more effectively for two reasons. First, because experts can readily detect patterns, they are more likely to note any inconsistent or incomplete information resulting either from a poorly written text or careless processing. Second, since knowledge contributes to processing efficiency, working memory capacity should be available to monitor and, if necessary, repair comprehension. An elaborate knowledge base should also facilitate the repair process.

Baker and Brown (1984b, p. 375) point out that focusing on the main points of a text, which is sometimes thought of as a metacognitive skill, depends on subject knowledge.

> *"If the text deals with topics unfamiliar to the reader, it will be*

> *difficult for her to understand the significance of the material, to select main points, and to disregard trivia. One has to understand the meaning of the material one is reading to be able to identify just what is important and what is trivial."*

However, Baker and Brown (1984b) also use phrases such as *"inherent structure"* and *"detect the logical structure"*. Brown (1980) refers to skilled readers having metacognitive knowledge which allows them *"to predict ahead of time what are the important elements of the texts"* (p. 465). It is important to keep in mind that skilled readers can 'detect' the main ideas in a text because their subject knowledge, rather than just metacognitive knowledge, allows them to see this pattern. A point made by Kuhn and Phelps (1982, p. 42) is relevant at this point; in a study observing the development of strategies to isolate the variable responsible for bringing about a particular chemical change, the authors note:

> *"The difficulty experienced by many of our subjects in mastering the problem we posed to them serves as a humble reminder of a fact occasionally forgotten by social scientists - 'evidence' does not exist in a body of data itself, it exists in the eye of the beholder."*

The same can be said of text structure and importance ratings of ideas. The work of Pichert and Anderson (1977) shows that how important an idea is judged to be is influenced by the goals of the readers. Van Dijk and Kintsch (1983) point out that macropropositions, representing main ideas, often have to be constructed by the reader, rather than being in the text waiting to be 'detected'. Clearly the text presents information, including cues as to structural importance, but the role of the readers' subject knowledge and processing activity in constructing

meaning, including logical relationships and judgements of importance, should not be underestimated.

From what has been said so far about the relationship between cognitive and metacognitive aspects of reading, it follows that training should not focus too heavily on metacognitive skills, since these must be seen in the context of the efficiency of the whole processing system. Baker and Brown (1984a, p.44) invite a similar conclusion when they note that poor comprehension monitoring may be the result rather than the cause of poor comprehension.

> *"The evidence is clear that less experienced and less successful readers tend not to engage in the cognitive monitoring activities characteristic of more proficient readers. Though it is tempting to conclude that inefficient monitoring of one's cognitive processes during reading is the cause of poor comprehension, we caution against such a precipitous conclusion. The majority of studies have shown that ineffective monitoring is associated with poor comprehension, but not that it is the cause. It may be that poor comprehension reduces the ability to monitor one's ongoing activities, or perhaps a third factor, such as impoverished background knowledge, is responsible for both problems."*

Thus, while metacognitive regulation undoubtedly plays an important role in developing skills such as reading, it must be considered within the constraints of the overall processing system. Metacognitive aspects of performance need to be examined in relation to knowledge of the subject matter in the text being read, as well as overall processing efficiency and working memory capacity.

Developmental aspects of metacognition

Studies to be discussed in chapter two show that metacognitive knowledge and self-regulation generally increase with age and skill. It is useful here to consider the course of this development with the intention of drawing out implications for training programmes.

In one study, Kuhn and Phelps (1982) observed children in grades 4 and 5 over a period of eleven weeks. In particular, they observed the strategies used by the children in their attempts to solve isolation of variables problems, where the children had to determine which of a range of solutions was responsible for bringing about a chemical reaction when mixed with one standard test solution. While Kuhn and Phelps observed the emergence of more sophisticated strategies over the eleven week period, that is, strategies emerged that were genuinely scientific and capable of solving the problem, they also noted a lack of consistency in the application of these strategies. A child would revert to less advanced strategies within the same experimental session, and would not necessarily carry over the advanced strategies from one session to the next. Kuhn and Phelps propose that the use of advanced strategies did not stabilise earlier because the metacognitive knowledge which would ensure consistent application had not yet developed. They describe the importance of this metacognitive knowledge as follows:

> *"Our study ... points to the importance of 'metastrategic' knowledge of what strategies are effective for a given problem, in contrast to strategic knowledge of how to execute effective (or ineffective) strategies. If, in the course of an encounter with the problem, subjects were doing no more than gaining practice in the application of a set of*

strategies to the problem, strategy use would remain relatively constant rather than change. They are also gaining knowledge about the problem and, in particular, knowledge about their own strategies as they apply to this problem. In short ... they are gaining understanding of how the capabilities they bring to the task bear on the demands posed by the task.

In a problem such as the one used in the present research, the metastrategic knowledge to which we are referring includes knowing that the most advanced strategy is the preferred strategy to apply to the problem, that is, knowing that this strategy works, exactly how and why it works, and why it is the best strategy to use." (p. 41)

Kuhn and Phelps suggest that it would have been relatively easy to teach the students the strategies necessary for isolating the relevant variable, but they note that using a strategy is not the same as understanding its significance. The absence of this understanding leads to what Kuhn and Phelps (1982, p. 41-2) describe as a 'generalization gradient', where: *"The less similar the transfer situation is to the original one, the less likely is the subject to apply newly-learned strategy, even though it is equally applicable and necessary in the new situation."*

Kuhn and Phelps make a distinction between competence and performance mastery, which is similar to the distinction made by Brown and DeLoache (1978) between behavioural and cognitive objectives. In both cases the former refers to the ability to apply particular strategies, while the latter refers to knowing when these strategies should be used. When only behavioural objectives are achieved, learners will apply strategies when told to do so, without having changed the way they independently approach similar problems in future.

Borkowski and Kurtz (1987, p. 127) adopt a position very similar to that of Kuhn and Phelps when they describe the development of the four components of their model as follows:

> *"The components in our model of metacognition develop at different times in a child's history, are dependent upon the integrity of other cognitive processes (such as primary or working memory), and are related to the type and quality of formal and informal educational experiences. In addition, an important internal dependency exists in the model - each component relies on the quality and form of earlier componential development. For instance, General Strategy Knowledge and its correlated beliefs about strategy efficacy emerge only after a number of specific strategies have been encountered by the child and considerable understanding about the roles and importance of these strategies is achieved. More importantly, the executive aspects of the system (represented by Metacognitive Acquisition Procedures) become operative only after detailed representations in Specific, Relational and General Knowledge components have been accumulated. Finally, the efficient operation of the entire system is linked to motivational/personality factors such as attributional style, self-esteem, and achievement motivation."*

In view of the above internal dependency, Borkowski, Carr and Pressley (1987) point out that it is particularly important to teach Specific Strategy Knowledge well, with sufficient and appropriate practice on a diverse range of tasks. In this way, students are in a position not only to master the application of the strategies,

but also to learn about the range of situations to which they apply. In addition, if strategies are learned well and succeed in enhancing performance, students can learn more about the value of strategic behaviour in general, including their own ability to control learning outcomes through the use of well directed effort. In other words, they are increasing their General Strategy Knowledge.

Motivational factors are an important component of this General Strategy Knowledge, and play an important role in determining whether students will apply available strategies independently (Borkowski et al, 1987). The student who, as a result of a history of successful strategy usage and performance, has a high academic self-esteem and attributes success to effort, has a considerable incentive to continue applying available strategies in order to achieve further success. The reverse applies to students with self-defeating beliefs. If strategies have been poorly applied in the past and this failure is attributed to low ability or other uncontrollable factors, the student is unlikely to believe that the effort involved in strategic behaviour will be rewarded in future. Performance is likely to remain poor, further lowering self-esteem and reinforcing inappropriate attributions for failure.

Borkowski and Kurtz (1987, p. 145) emphasise that the General Strategy Knowledge discussed above, as well as the executive functions of the Metacognitive Acquisition Procedures are important aspects of a training programme for children with learning difficulties.

> *"This is not to imply that deficits in Specific Strategy Knowledge are unimportant. To the contrary, specific knowledge about a variety of strategies is necessary for the emergence of higher-level metacognitive components. It would be foolhardy to train general or executive*

knowledge with a child who has little knowledge about specific strategies and their respective uses. We propose an integrated form of training that focuses simultaneously on the interaction of specific strategies (including when to use them and when not); general implementation and monitoring routines; and beliefs about the fundamental importance of deliberation, planfulness, and effort for successful learning."

Thus, Specific Strategy Knowledge should be taught in such a way as to also develop what Borkowski and Kurtz refer to as 'higher-level metacognitive components'. A similar view is also put forward by Baker and Brown (1984b, p. 381), albeit with less explicit emphasis on motivational factors such as self-esteem and attributions.

"... 1) training and practice in the use of task-specific strategies (skills training); 2) instruction in the orchestration, overseeing and monitoring of these skills (self-regulation training); and 3) information concerning the significance and outcome of these activities and their range of utility (awareness training). Students who receive only instruction in the skills often fail to use them intelligently and on their own volition because they do not appreciate the reasons why such activities are useful, nor do they grasp where and when to use them. Adding instruction in 'awareness', or knowledge about a skill's evaluation, rationale and utility greatly increases the positive outcomes of training studies. This is, in effect, training for lateral transfer, that is, explicitly instructing students about where, when and how to use a strategy in a variety of appropriate domains."

Kuhn and Phelps (1982) have provided some support for the notion that metacognitive, or what they call metastrategic knowledge is essential for the efficient and consistent application of newly acquired strategies. Empirical evidence to support Baker and Brown's view of the importance of self-regulation and awareness components in training studies will be presented in chapter three.

As a final point, Brown and DeLoache (1978, p. 13) suggest that novices need information not only about strategies, but also about the goals for which these strategies are useful. They state: *"The novice tends not to know much about either his capabilities on a new task or the techniques necessary to perform efficiently; he may even have difficulty determining what goals are desirable, let alone what steps are required to get there."*

The example was mentioned earlier of poor readers seeing accurate decoding as the goal of reading. Such children are unlikely to have a concept or model of a more appropriate goal, since they may never have had the experience of thoroughly understanding a text and as a result do not appreciate that this is the goal they should be aiming for. It is important, therefore, to provide means whereby they can develop reading goals which will encourage more thorough processing. Adequate reading goals are also important for the understanding of the significance of the cognitive actions being taught. The value of such actions is most fully understood in relation to progress towards a goal, and even very useful actions may not appear as such to the student who has very limited goals.

Problems with the construct of metacognition

While recognising its importance, Jacobs and Paris (1987, p. 257) draw attention to remaining problems with the construct of metacognition as follows:

> *"We believe that the enthusiasm and criticism generated by the construct of metacognition are both well deserved. Theories of reading have neglected cognitive aspects of strategies and self-regulated learning for too long, and metacognition offers some promise as a remedy. But the term may have been embraced prematurely. If metacognition is to be useful for researchers and teachers alike, it must be studied more extensively and in greater detail. In particular, more attention needs to be paid to the definition of metacognition and the creation of appropriate measures of metacognition."*

With regard to the measurement of metacognitive knowledge and self-regulation, Paris, Wasik and van der Westhuizen (1987) point out that all methods have certain inherent problems. Open-ended interview questions to assess metacognitive knowledge require subjects to introspect and articulate their cognitive processes. This task can be difficult, especially for children, and the resultant verbal reports may be distorted or inaccurate. It is not clear how such reports relate to actual performance, that is, if and how the thoughts articulated are applied in practice.

The problem of relating verbal reports to performance is illustrated in an article by Phifer and Glover (1982) entitled 'Don't take students' word for what they do while reading'. Students at college level read an expository passage in which they were to underline main ideas and to use any strategies they thought would assist recall. They were asked to report strategies they had used after free recall was taken. Good comprehenders underlined more main ideas while reading and recalled more main ideas than poor comprehenders, but contrary to expectation there was no significant difference in terms of reported strategy use. On closer

inspection it emerged that only three good comprehenders reported taking notes, whereas in fact eight had done so. In contrast, all the poor comprehenders who reported taking notes had done so.

Garner and Gillingham (1987) found a weak relationship between the knowledge articulated by children about paragraph construction and the knowledge demonstrated in the process of paragraph construction. Baker and Brown (1984b, p. 377) suggest one reason for poor relationships between verbal reports and performance, namely, that it is inherently difficult to examine and describe mental processes.

> *"Intuitively, declarative and procedural knowledge differ in the ease with which they can be subjected to deliberate examination. Consider, for example, the relative difficulty of reporting that tying your shoelace is a good way to avoid tripping as compared to describing how to tie your shoe."*

Paris, Wasik and van der Westhuizen (1987) also point to problems with other methods of articulating metacognitive knowledge. Questionnaires, where readers rate how often they use particular strategies or how useful they are, reduce the difficulty of introspection and verbalisation posed by open-ended questions, but also reduce self-expression and can encourage guessing. Subjects may be cued by the content of the questionnaire to make responses they would not otherwise have thought of. Think-aloud protocols, where readers report what they are thinking while reading, can interfere with normal processing and can also elicit fabrications. Subjects may feel under pressure to say something due to the demands of the situation, even if what they say is not an accurate reflection of their thinking.

While tutoring another person on how to perform a task is a more naturalistic expression of strategy knowledge, Paris et al point out that it depends on verbalisation and social interaction ability. Those subjects who are less able to articulate their knowledge or to relate easily to others may not be able to give an accurate demonstration of the extent of their knowledge.

Given the problems associated with verbal reports, some researchers have chosen to infer the level of metacognitive knowledge and skill from subjects' performance, for example, in error detection, self-correction of oral reading errors, or use of overt strategies such as underlining. However, Paris et al (1987, p. 8) point out that this method also poses problems:

> *"...it is not certain that poor performance indicates a lack of metacognition about reading. Poor readers may not use strategies because they lack specific knowledge, adequate practice, or motivation. Therefore, inadequate metacognition is only one plausible hypothesis for poor use of reading strategies. Poor performance does not necessarily imply poor metacognition anymore than proficient reading implies good metacognition."*

A study by Thorpe and Satterly (1990) illustrates the difficulties of making generalisations in this area. These authors used four tasks commonly thought to reflect metacognitive knowledge about memory processes, and they expected to find a common factor underlying performance. Children aged between 7 and 11 years completed the tasks of generating strategies, generating word lists, organising prose, and judging task difficulty. In the first task, the children generated as many strategies as possible to deal with simple problem situations, while in

the second they generated four words that would be easy to learn and remember in response to each in a series of cue words. In the third task, the children were presented with twelve sentences in random order about three aspects of one topic and they were asked to organise and explain their organisation of these sentences. Finally, the children were asked to judge which list in each of a series of word list pairs would be easier to learn and remember.

Factor analysis revealed four factors underlying the four measures, which were only weakly related and failed to show a common metacognitive component. Thorpe and Satterly (1990, p. 19) conclude that *"metacognition is specific to the task from which it is derived"*. They question the notion of a general metacognitive competence and refer instead to a *"disparate range of higher level cognitive skills."*

Perhaps this result is not surprising given the points made above by Paris et al to the effect that a performance depends on many factors other than metacognition, including specific knowledge and motivation. However, the study does illustrate the difficulty of defining metacognition precisely in the context of different task demands, and of delineating the contribution of cognitive and metacognitive processes to performance.

In spite of these difficulties, the construct of metacognition is one which cannot be ignored. As indicated above by Kuhn and Phelps (1982) and Baker and Brown (1984b), if cognitive strategies are not associated with metacognitive knowledge they are unlikely to be maintained and transferred to novel situations. Hence, a training programme should include a metacognitive component linked as closely as possible to the demands of reading comprehension to take into account the conclusions of Thorpe and Satterly.

One important aspect is to enable readers to develop goals appropriate to the purpose of learning from expository text, which is the focus of this study. In addition, training should include cognitive actions consistent with models of reading which, as Flavell (1981) indicates, can have a dual function. On the one hand, they foster comprehension, while on the other they provide information about the level of that comprehension. In other words, they can teach students how to make active attempts to understand, the process of which provides as a by-product evaluative information about that understanding (Markman, 1985). Finally, training should include what Jacobs and Paris (1987) refer to as self-management and what Flavell (1981) refers to as metacognitive knowledge, that is, information about how, when and why the particular cognitive actions or strategies are useful. Clearly, these various aspects are interactive, as implied by the model of Borkowski and Kurtz (1987). As a result of learning to apply certain cognitive actions, students' comprehension may improve and they may develop a clearer concept of what it means to understand a text. Subsequently they are able to set clearer reading goals and stricter criteria of coherence for themselves. Similarly, as a result of applying various cognitive actions they may refine their knowledge of when and why the actions are useful, and may see themselves as more competent learners, able to control learning outcomes.

1.5 Conclusion

In this chapter, the Just and Carpenter (1980;1987) model was used as a framework for discussing theoretical approaches to reading. It was shown, by reference to other models and empirical findings, that this view is well supported. However, some aspects of the model were expanded, in particular the processing of the global structure of the text, the organisation of knowledge necessary to support the processing described in the model, and the metacognitive knowledge and skills required to guide and monitor processing in a strategic manner.

In the introduction, criteria for the selection of skills to be taught in a training programme were mentioned (Dole, Duffy, Roehler and Pearson, 1991). The first of these is that skills should be consistent with a cognitive model of reading. To conclude this chapter, key aspects of the cognitive model of reading already developed will be summarised to provide a framework for the examination of differences between good and poor readers in the next chapter.

Working memory capacity is an important factor influencing comprehension performance (Just and Carpenter, 1980). The efficiency with which lower-level processes, such as word encoding and lexical access, are carried out is assumed to influence the capacity available for intersentence integration, macrostructure formation and comprehension monitoring. Other authors, such as Perfetti (1984) and Perfetti and Roth (1981) also emphasise the importance of accurate and efficient visual analysis of print. Just and Carpenter (1987) consider that speed of lexical access contributes to comprehension performance, although they do not quantify the magnitude of this contribution. The extent of lexical knowledge, in addition to speed of access, has been shown to influence comprehension and recall

(Stahl et al, 1991; Beck et al, 1982).

Readers use a small number of strategies to parse text into clauses (Frazier and Rayner, 1982; Clark and Clark, 1977). Readers also use a variety of converging cues in order to strategically assign case roles within each clause (van Dijk and Kintsch, 1983; Just and Carpenter, 1987). This intraclause integration results in the formation of individual propositions.

In order to establish coherence, these propositions must be integrated into a coherent representation of the text. Various forms of semantic entailment may be used to determine points of attachment between successive propositions (Just and Carpenter, 1987). Assigning referents to anaphoric expressions, such as pronouns, is another means by which propositions are integrated. As in the case of intraclause integration, readers use a variety of converging cues to be able to strategically make this assignment (van Dijk and Kintsch, 1983; Just and Carpenter, 1987). Where no explicit point of attachment exists between a new clause and the existing text representation, readers must construct a bridging inference (Haviland and Clark, 1974). In some circumstances inferences other than those required to establish local cohesion may also be made, but this is only likely to occur in highly constrained contexts (O'Brien et al, 1988; McKoon and Ratcliff, 1981; Goetz, 1979).

An overall understanding of a text involves more than establishing the local cohesion described above; it also involves abstracting the main ideas and their interrelationships to form a macrostructure. Readers operate within the constraints provided by macrorelevant information when establishing local coherence in a sequence of propositions (Kintsch and van Dijk, 1978). Macrolevel information, such as titles, has been shown to strongly influence the interpretation of a text

(Bransford and Johnson, 1972; Schallert, 1976). Directing readers' attention to macrolevel information has also been shown to increase recall of important information and problem solving ability (Loman and Mayer, 1983).

Readers draw on various sources of information to form a macrostructure while reading, including titles, thematic statements and topic sentences, especially in initial positions, knowledge about conventional text structures and knowledge about the topic of discourse (van Dijk and Kintsch, 1983). Some macrorelevant information is explicitly stated in a text, while some information often has to be inferred. The formation of a macrostructure is a strategic activity which depends on the reader having particular reading goals (McKoon and Ratcliff, 1992; Graesser et al, 1980), as well as adequate subject knowledge and processing capacity (van Dijk and Kintsch, 1983).

Reading goals are one aspect of metacognition, which includes knowledge about specific strategies and when and why to use them, beliefs about self-efficacy and the value of effort and strategic behaviour in general, as well as an executive component which plans, monitors and regulates processing (Jacobs and Paris, 1987; Borkowski and Kurtz, 1987). Metacognitive knowledge about when and why strategies can be effectively used is essential for the independent use and transfer of those strategies (Kuhn and Phelps, 1982; Baker and Brown, 1984b). Motivational factors, including self-esteem and attributions for success and failure, are important determinants of whether readers will invest the time and effort necessary to apply strategies (Borkowski and Kurtz, 1987). Metacognition involves setting appropriate reading goals, implementing strategies to achieve them, and monitoring progress towards these goals (Jacobs and Paris, 1987).

Finally, knowledge organisation has been reconceptualised as an associative

net, from which semantic structures are generated according to which subsets of the net are activated in a particular context (Kintsch, 1988). Access to a rich and interconnected knowledge base has been shown to increase explicit and inferential recall (Lipson, 1983; Pearson et al, 1979) and the ability to determine important ideas (Recht and Leslie, 1988; Schneider et al, 1989).

The second criterion for the selection of skills in a training programme is that those skills should distinguish good and poor readers (Dole et al, 1991). The second chapter will examine those aspects of the cognitive model summarised above which have been investigated from this perspective.

ldots

Chapter 2

Skill Differences Between More and Less Advanced Readers

2.1 Introduction

In the conclusion to the previous chapter, important components of a cognitive model of reading were summarised. The second chapter will examine skill differences between more and less advanced readers in some of these areas. The category of more advanced readers includes both readers who are older and more mature than the comparison group, and readers who are more skilled than their chronological peers. Thus, studies discussing differences between older and younger children, and between children of different reading ability levels will be examined. Areas to be considered include the following:

1. processing efficiency and working memory capacity;

2. interclause integration, in particular the construction of bridging inferences;

3. macrostructure formation, including the ability to generate and use information central to the gist of the passage;

4. metacognitive characteristics, including goal setting, strategy selection, comprehension monitoring, and motivational characteristics such as attributional style.

From the examination of differences in these areas, suggestions will be made regarding skills which could be targeted in an instructional programme. Two of the three criteria proposed by Dole et al (1991) for choosing skills to be taught will then have been met. Certain skills have been identified as important within a cognitive model of reading, and the extent to which these skills distinguish more and less advanced comprehenders will be the subject of this chapter. The third criterion, that the skills be amenable to instruction, will be addressed in the third chapter.

2.2 Processing efficiency and working memory capacity

Within the Just and Carpenter (1980) model, working memory capacity is an important determinant of comprehension performance. An indicated in the first chapter, Just and Carpenter assume that functional working memory capacity is determined by the efficiency of processing; the more efficiently processes such as word encoding, lexical access and intraclause integration are carried out, the more capacity there is remaining for integrative processes such as anaphoric resolution and inferencing to construct a coherent representation of the whole text.

In the first part of this section, differences between good and poor readers in the areas of word encoding and semantic memory access will be examined. In the second part, the relationship between these word level processes, working memory capacity, and comprehension performance will be discussed in order to arrive at a conclusion about the importance of word-level skills in a training programme.

2.2.1 Efficiency of word encoding and semantic memory access

Some researchers have contended that word-level processes are important in explaining reading ability differences among children (for example, Perfetti, 1984; Stanovich, 1980). The assumption made by such researchers is that inefficient word-level processing consumes cognitive resources which are then less available for higher-level integrative processing.

Perfetti and his colleagues have shown in several studies that good readers

have superior word recognition skills to poor readers. Perfetti and Hogaboam (1975) examined children in grades 3 and 5, with skill level determined by the reading subtest of the Metropolitan Achievement Test (MAT). Recognition speed for isolated high frequency words was not significantly different for skilled and less skilled readers. However, there was a large skill difference for low frequency words and especially pseudowords. The words in this study were also divided into those known and those not known by each individual child. When vocalisation latencies for these two groups of words, calculated separately for each child, were compared, the differences were larger for less skilled readers. In other words, the vocalisation speed of less skilled readers was more adversely affected by unfamiliar words. The less skilled readers also made more errors on unknown words.

Perfetti, Finger and Hogaboam (1978) tested children in grade 3, with reading skill level again determined by the MAT. Skilled readers were less affected by word length than less skilled readers, with the former showing only an 8% increase in vocalisation latencies as words increased from one to three syllables, compared to a 25% increase among the latter.

Such differences between skilled and less skilled readers in the speed of word recognition have also been found by Curtis (1980) in grades 3 and 5, by Samuels, Begy and Chen (1975/6) in grade 4, by Perfetti, Goldman and Hogaboam (1979) in grade 5, and by Perfetti and Roth (1981) in grade 4.

In addition to reading under normal conditions, Perfetti and Roth (1981) also examined the effect of print degradation on word processing. Print was degraded by removing between 0% and 35% of the dots in letters. While word recognition speed decreased for all readers with increasing levels of degradation, less skilled readers were affected to a significantly greater extent. This finding suggests that

the slower word recognition of less skilled readers is due to less efficient visual analysis skills; any condition which makes visual analysis more difficult then has a relatively larger effect on poor readers. The importance of visual analysis skills is also suggested by other studies cited above, which show that the word recognition of poor readers is more affected, relative to good readers, by word frequency and familiarity and pseudowords. In these cases, letter-level analysis is less likely to be facilitated by feedback from the word level (Rumelhart and McClelland, 1981), and poorer readers appear to rely more on such feedback.

There is also evidence to show that less skilled readers rely to a greater extent on contextual facilitation to recognise words, an observation which is incorporated into Stanovich's (1980) compensatory hypothesis. He proposes that when visual analysis is slow, the reader compensates by making greater use of the preceding context to recognise words.

Stanovich makes use of the Posner-Synder (1975) theory to explain how such compensation comes about. According to Posner and Snyder, there are two independent processing mechanisms which can operate simultaneously, but which have different characteristics. Automatic activation occurs very rapidly, without attention, without conscious awareness and without interference to other mental activity. Conscious processing, on the other hand, is slower acting and does require attention. While attention is being directed, for example, to a particular memory location or input channel, other conscious processing is inhibited because attentional capacity is strictly limited. However, automatic activation is not inhibited and can continue simultaneously.

Stanovich proposes that in fluent readers the skills involved in word recognition are so rapid, that any contextual facilitation effects can only occur through

automatic spreading activation. However, when the processing of visual information is slow, either because of developmental immaturity or stimulus degradation, the conscious processing of semantic and syntactic information has time to operate and influence word recognition by means of the expectancies generated.

Stanovich (1980, p. 57) points out that contextual facilitation due to automatic spreading activation carries no cost, but the use of attentional capacity to compensate for poor visual recognition skills reduces the capacity available for higher-level semantic integration.

> *"...a reader might use cognitive capacity in order to construct a new knowledge structure or to integrate new material into an old knowledge structure. On the other hand, a reader may also use the knowledge structures built up from previous context to aid in the processing of words that are currently undergoing recognition. Notice that the less cognitive capacity that is used to engage in the latter process, the more that is left over for the more important former processes. It is argued here that this is precisely what the fluent reader is doing. He recognises words rapidly and mostly on the basis of physical cues, so that expectancy processes that draw cognitive capacity are not necessary. Thus, his capacity is being used for comprehension, rather than conscious prediction processes that aid individual word recognition."*

In chapter one it was pointed out that context facilitates the word recognition of adult readers. This facilitative effect is also found among children of various ages and levels of reading skill (Biemiller, 1977/8; West and Stanovich, 1978; Perfetti, Goldman and Hogaboam, 1979; Perfetti and Roth, 1981).

West and Stanovich (1978) specifically investigated the extent and mechanism of contextual effects on word recognition among students in grades 4 and 6 and at college level. Target words were placed in congruous, incongruous or no-sentence contexts. Overall, speed increased with age and context facilitated word recognition. Within each age group, response time to words in congruous sentences was shorter than to words in the no-sentence context. However, comparing the incongruous context with the no-sentence condition, response times were longer for the children of both ages in the incongruous context, but not the adults.

West and Stanovich calculated a facilitation score (congruous - no context) and an interference score (incongruous - no context) for each age group. There was no significant age difference in contextual facilitation, but children showed significantly greater interference than adults due to an incongruous context. This finding supports the authors' hypothesis that among skilled adult readers word recognition is so rapid that contextual facilitation occurs only by spreading activation, which does not interfere with visual processing in the incongruous context. Less mature readers, with less rapid word recognition, are more likely to achieve contextual effects through conscious processing, which produces interference with visual processing in the incongruous context.

Finally, these facilitation and interference scores were related to the children's reading achievement, as measured by the reading sub-test of the Wide Range Achievement Test. Both scores were negatively correlated with achievement. Thus, the word recognition speed of better readers was less affected by both contextual facilitation and interference. In other words, better readers are less reliant on the conscious prediction described by Stanovich (1980), since they are able to efficiently analyse graphemic cues. As a result, word recognition

is less affected by an incongruous, and hence unpredictable context, although facilitation still occurs via relatively attention-free spreading activation. Poorer readers, relying more heavily on conscious expectancy, also gain the benefits of contextual facilitation, but at the expense of greater use of working memory capacity and inhibition of unexpected responses.

If reliance on context and conscious expectancy processes to facilitate word recognition are characteristic of poorer readers, then they should be more commonly observed among beginning readers. Stanovich, West and Feeman (1981) directly investigated the magnitude of contextual facilitation in relation to practice and the development of reading skill. Children in grade 2 were tested at the beginning and the end of the year. The task involved naming the final word in sentences such as *The pilot flew the* .... One condition provided a congruous word (*plane*), one provided an incongruous word (*bone*), and one provided no context at all. Half the target words were considered easy and half more difficult. Half the words in each group were practised on 15 successive days prior to the test.

Main effects on word recognition speed included context, practice, word difficulty and time of testing. Of particular importance here is that context had a larger effect on word recognition speed at the beginning of the year, on unpractised words, and on the more difficult words. Thus, Stanovich's (1980) contention that context will have a larger effect when word recognition based on visual analysis is slow was supported by these results.

This conclusion is supported by Biemiller (1970) who, in a study of readers in their first year of instruction, found definite trends as children improved in reading competence. He examined oral reading errors. Initially the words substituted

were contextually appropriate, but bore little visual resemblance to the text in terms of the shape of the printed word. In the next phase, children made a high percentage of non-response errors, that is, refused to say a word they found difficult, but at the same time more of their response errors were visually similar to the original. Finally in this first year, non-response errors declined and response errors tended to be both graphically similar to the original and contextually appropriate. Biemiller concludes that in the first phase children are minimising their use of graphemic information, but in the second phase they are beginning to attend to print and to slowly acquire decoding skill, which is being consolidated in the third phase. Biemiller concludes that to make progress in word recognition, children have to learn to rely on graphic rather than contextual information.

Consistent with Biemiller's conclusion, several studies show that less skilled readers continue to rely more heavily on contextual information beyond the early years of reading. Perfetti, Finger and Hogaboam (1978) show that skilled readers in grade 3 are less affected by the predictability of words than less skilled readers. The words to be read came from a small, closed set (the four seasons), a larger closed set (the twelve months), a small, open set (four animal names), and a large open set (twelve Christian names). The vocalisation latencies of skilled readers were unaffected by set membership, but less skilled readers took significantly longer to identify words from the least restricted category, namely, the large, open set.

In a study by Perfetti, Goldman and Hogaboam (1979) skilled and less skilled readers in grade 5 read a target word either in isolation or at the end of a context sentence. Overall, skilled readers recognised words more rapidly than less skilled readers. However, while skilled readers were unaffected by context, less skilled

readers responded more slowly when a target word was presented in isolation rather than in context.

This contextual effect was replicated in a further study by Perfetti et al (1979), in which subjects first attempted to predict the target word before reading it. Skilled readers made more accurate predictions than less skilled readers. Both groups showed increased vocalisation latencies when the words were not correctly predicted, but the increase tended to be larger for less skilled readers. Thus, skilled readers were both better at using context to make predictions and at the same time less dependent on context to facilitate word recognition.

Perfetti and Roth (1981) report an experiment with skilled and less skilled readers in grade 4, where a target word was read either in a highly predictable, an unpredictable but acceptable, or an anomalous story context. Context had a small effect on skilled readers, but a large effect on less skilled readers, with skill differences becoming increasingly large as contextual constraint decreased.

In this section so far it has been shown that primary school children who are poorer readers, as defined by standardised reading tests, have less efficient visual analysis skills and are more likely to make use of resource-demanding conscious expectancy processing to facilitate word recognition. Some studies have also shown that poor readers are less efficient at accessing information from semantic memory.

Howell and Manis (1986) found differences in the speed of naming objects among skilled and less skilled readers. The younger children were in grades 2 and 3, the older ones in grades 5 and 6. The less skilled readers were in remedial reading classes and performed poorly on the Woodcock Reading Mastery Word Identification Test. The children were shown a basic level statement (*This is*

a dog) or a superordinate statement (*This is an animal*), which they judged to be true or false in relation to a picture or word. The response to the basic level statement provided a measure of naming time, while response to the superordinate statement provided a measure of categorisation time.

Older children responded faster than younger ones, and normal readers responded faster than remedial readers, with the latter difference being more marked in the younger group. The differences occurred for both pictorial and verbal stimulus material, suggesting a general inefficiency in processing semantic information.

Chabot, Petros and McCord (1983) also found that less skilled readers took longer to make categorical decisions. Children in grades 2, 3 and 5 were categorised as good, average or poor readers according to their achievement level in the McGraw Hill reading series. They were shown two pictures, two words, or a picture and a word, and had to decide whether or not the two exemplars were from the same category.

Good readers responded faster than average ones, who in turn responded faster than poor readers. These differences were larger for verbal than for pictorial stimulus materials, in contrast to the findings of Howell and Manis (1986). It is not clear why this discrepancy exists, although it may be that the remedial readers studied by Howell and Manis showed more general and extensive processing deficits than the normal range of students studied by Chabot et al.

Chabot et al (1983, p. 139) conclude that the slower response time of poorer readers in their study was related to both word encoding deficits and slower access to semantic memory, which could lead to comprehension problems for the following reasons:

> *"... (1) large amounts of processing resources would have to be allocated to simple word encoding, lexical access and semantic memory access processes; (2) these slower word recognition processes would increase the difficulty of maintaining in short-term memory long enough to integrate semantic information across phrases, sentences and paragraphs; and (3) the utilization of cognitive resources for word recognition processes would limit the available capacity which could be allocated to higher order comprehension processes such as inference making."*

If rapid visual analysis and access to word meanings are essential components of skilled reading, then increasing the speed of word identification should improve comprehension. Fleisher, Jenkins and Pany (1979) examined the effect on comprehension of increasing the decoding speed of poor readers in grades 4 and 5, with children classified according to the reading subtest of the MAT. Poor readers during training practised reading a list of words from a target passage until they achieved the speed of good readers. These trained poor readers, as well as untrained poor readers and good readers, then read the list of practice words and the target passage, and completed a cloze test and tests involving inference and factual questions.

In terms of reading accuracy, the trained poor readers performed as well as the good readers on both the word list and the passage, while their reading speed was equivalent to that of good readers on the word list only. The good readers performed better in terms of word reading speed in context and the three measures of comprehension. Good readers differed from untrained poor readers on all measures, while the trained poor readers performed better than their untrained

peers on all oral reading measures, but on none of the comprehension measures. Thus, there was no indication of any comprehension improvement as a result of rapid word identification, and Fleisher et al (1979, p. 45) conclude:

> *"For instructional purposes it is important to note that short-term isolated word practice is possibly an inadequate strategy for producing the level of decoding automaticity that is required to enhance comprehension."*

In referring to the Fleisher et al study, Perfetti (1984) suggests that word identification speed reflects the accessibility of word representations, but can only be a rough approximation of automaticity, that is, the amount of processing resources consumed by the process. Perfetti prefers to conceptualise accessibility in terms of quality of representation rather than degree of automaticity, considering quality in terms of the precision, interconnectedness and redundancy of stored information. He gives the reason for this preference as follows (p. 55):

> *"It makes it easier to notice that, to improve reading efficiency, the reader's knowledge of words and orthographic patterns should be increased. Speed of identification should be thought of as a reflection of this kind of knowledge. Speed of access is not necessarily something to train independently of the representation system it works on."* (emphasis in original)

Perfetti is suggesting, therefore, that for students whose word recognition speed is slow enough to impede comprehension, training should be aimed at developing systematic and interconnected representations of orthographic and

semantic information, rather than just aiming for short-term improvements in recognition speed for lists of words, as was done in the Fleisher et al study. In that study, only naming speed was addressed, with no consideration of speed of access to word meanings and associations.

The kind of training advocated by Perfetti (1984) was incorporated into a study by Beck, Perfetti and McKeown (1982), referred to in chapter one in the section on lexical access. In this study, training aimed to build an interconnected memory representation of new words that could be quickly accessed. In achieving this aim, the study also succeeded in improving reading comprehension.

2.2.2 Working memory capacity and comprehension

As Just and Carpenter (1987) point out, reading involves the integration of sequentially presented verbal information. Since all of the integrative operations take place within a working memory of limited capacity, the reader whose processing is slow and inefficient is likely to lose more information from working memory through both decay and displacement, resulting in less efficient integration and reduced comprehension (Daneman and Carpenter, 1980). Daneman (1984, p. 368-9) explains this approach to the study of individual differences in comprehension as follows:

> *"The major tenet of the theory is that individual reader differences influence the integration processes of reading comprehension and that one crucial source of reader difference is processing efficiency in working memory. Poor readers may be less efficient at the processes that manipulate verbal material, processes such as retrieving letter or name*

codes from semantic memory. They may also be less efficient at some other semantic computations. These deficiencies have implications for the way in which readers allocate their working memory resources. Slower and less efficient processes may consume more of the available capacity and so lessen the amount of additional capacity for storing and maintaining information in working memory. If working memory has less information, subsequent processing will not be optimal. Without the relevant information accessible in working memory, poor readers will be less able to integrate successive elements of the text."

According to Daneman (1984), what is important in terms of reading performance is not a static storage capacity in working memory, as measured by digit or word span tests. Rather, it is the functional capacity which is important, and this depends on processing efficiency. The more efficient the processing, the greater the remaining capacity for storage. In other words, there is a processing-storage tradeoff.

Daneman and Carpenter (1980) devised a test involving both the processing and storage functions of working memory. In what the authors call a Reading Span test, subjects read aloud sets of unrelated sentences at normal pace (the processing component), and stored the last word of each sentence for recall at the end of the set (the storage component). Sets included from two to six sentences, and subjects' reading span was defined according to the longest set they were successfully able to recall.

Daneman (1984) notes that among the university students tested, reading span scores varied from 1.5 to 5, with less than 10% of subjects scoring 5. Thus, even a relatively homogeneous group of highly educated subjects showed a wide

range of scores, with only a small percentage operating with maximum efficiency. Reading span scores showed a significant correlation ($r = .59$) with verbal Scholastic Aptitude Test (SAT) scores (Daneman and Carpenter, 1980). Based on a number of studies, Daneman (1984) reports that reading span scores account for 25% to 35% of variance in verbal SAT scores.

Goldman, Hogaboam, Bell and Perfetti (1980) also demonstrated a relationship between reading skill and working memory capacity, in this case among children in grades 3 and 4. Skilled and less skilled readers were identified according to the reading subtest of the Metropolitan Achievement Test (MAT). Goldman et al hypothesised that brief memory of exact words and sentence structure was necessary in order to be able to construct meaning-based representations for long-term storage, and that working memory capacity would influence the duration of this verbatim memory. They used a probe procedure, whereby a word was repeated either two or five content words after it first occurred in a text. The children were to repeat verbatim all the words which followed the first occurrence of the word. Skilled readers recalled more words than less skilled readers, suggesting that they did, as hypothesised, have greater working memory capacity and should, as a result, be in a better position to construct meaning-based representations.

Perfetti and Goldman (1976) similarly showed superior short-term discourse memory among skilled readers in grades 3 and 5, with classification into good and poor readers again based on the reading subtest of the MAT. Children listened to two stories, in which some sentences were followed by a probe word, that is, a word which had occurred either one or two clauses previously. The task for the children was to state the word which had followed the first mention of the probe word.

Skilled readers showed superior recall of target words. In this study the need for decoding print was removed because the stories were heard. It is possible that skill differences in recall would have been larger had the stories been read without the opportunity to look back, because then the less efficient decoding skills of the poorer readers may have placed an even greater burden on their working memory capacity. This skill difference in recall did not emerge when the stimulus material was a string of digits instead of a text. Thus, the poorer readers did not suffer from a general memory deficit, but only showed reduced memory span relative to good readers when processing of verbal material was required. This finding supports Daneman and Carpenter's (1980) assumption of the importance of *functional* rather than static working memory capacity.

Thus, there is some evidence that reading span is related to overall reading skill. Several studies have investigated the relationship in more detail, in particular the relationship between reading span and the various subcomponents of reading.

Carpenter and Just (1989) examined processing efficiency among university students on the Reading Span test itself, using gaze durations as data. Readers with small spans (RS = 2) fixated each word in a sentence, excluding the last word, for an average of 77 msec longer than readers with large spans (RS = 4). Both groups of readers were affected to the same extent by increases in word length, that is, gaze durations increased to the same extent in both groups as word length increased. However, small-span readers were affected to a greater extent by decreases in word frequency, that is, their gaze durations on less frequent words increased to a greater extent than the gaze durations of large-span readers. Since word length was assumed to affect word encoding and word frequency was

assumed to affect lexical access, Carpenter and Just conclude from these findings that large- and small-span readers, at least at the university level, do not differ in the efficiency of word encoding, but that they do differ in efficiency of lexical access.

However, Herdman and LeFevre (1992) found a relationship between reading span and word naming latency among university students. These authors examined word naming latencies in single and dual tasks, the latter requiring subjects not only to name a word, but also to decide whether a tone was high or low. Both large- and small-span readers took longer to name words in the dual task. However, large-span readers showed a smaller increase in latencies in the dual task than small-span readers, suggesting that for them word recognition required fewer cognitive resources. In other words, their word recognition was more efficient, and may have contributed to their larger reading span, in contrast to the Carpenter and Just (1989) finding.

Carpenter and Just (1989) examined the relationship of reading span to syntactic analysis. They compared reading times for sentences of varying syntactic complexity, such as the following examples:

(1) *The paper that the senator attacked admitted the error.*

(2) *The paper that attacked the senator admitted the error.*

Sentence (1) is assumed to be more difficult to process because *the paper* is the subject of one clause (*The paper admitted the error*) and the object of the other (*The senator attacked the paper*). In sentence (2), however, *the paper* is the subject of both clauses, so syntactic analysis and case role assignment should be easier.

Readers with large spans showed no significant difference in reading times for these two types of sentences, while the reading times for small- and medium-span readers were significantly elevated for sentences of type (1), with gaze durations increasing especially on the first verb, where a decision has to be made as to which noun is its subject and agent. Thus, the syntactic analysis of large-span readers was more efficient.

Large-span readers, then, are possibly more efficient at word level processing, as well as at syntactic analysis. As a result of this greater efficiency, they should be at an advantage with regard to higher-level semantic processing. When integrating newly encoded information, concepts and relations from the immediately preceding text should still be active in working memory, both because processing is rapid and the previously activated information has not yet decayed, and because processing uses minimal capacity and so less rapidly displaces activated information from working memory. The reader with a large reading span is therefore in a position to integrate concepts into coherent chunks within the long-term text representation. Daneman and Carpenter (1980, p. 464) draw attention to further advantages accruing to the reader with a large reading span:

> *"...he should be able to detect more interrelations among these concepts, and to note their relative importance. The good reader's chunks should be richer, and more coherent, and contain different information. The presence of different interrelations could subsequently allow different inferences and generalizations to be drawn."*

These richer chunks should facilitate recall and integration processes, such as assigning referents to pronouns. Daneman and Carpenter (1980) examined

fact retrieval and pronoun assignment when the pronoun and its referent were separated by distances varying from two to seven sentences. Among university students with reading span scores ranging from 2 to 5, large-span readers correctly answered both more fact and more pronoun assignment questions than small-span readers. In addition, there was an interaction between reading span and distance; large-span readers were unaffected by the distance between the pronoun and its referent, whereas the ability of small-span readers to correctly answer the pronoun questions decreased monotonically as distance increased.

As Daneman and Carpenter (1980, p. 62) note, the differences between large- and small-span readers were not only quantitative, but also qualitative.

> *"Not only do good readers make fewer errors than poor readers, but their errors are more likely to reflect at least some understanding of the passages. On the other hand, when poor readers or listeners fail to retrieve a fact or a pronoun's referent, their errors frequently reflect a more fundamental misunderstanding of the passage. Their errors might contradict the whole theme of the passage."*

Large-span readers in the Daneman and Carpenter study were more able to supply an appropriate title for the passages they read. They had presumably constructed a more coherent representation of the gist or theme of these passages, so it is not surprising that their errors in response to questions were due to incompleteness or reasonable substitutes, rather than the confusion which was evident in the errors of small-span readers.

Finally, since large-span readers are hypothesised to construct more interrelationships among the concepts being processed, they should also be more aware of and more able to recover from inconsistent interpretations.

Daneman and Carpenter (1983) investigated the relationship between reading span and the ability to recover from errors. University students read paragraphs, displayed one word at a time, in which one meaning of an ambiguous word such as *bat* was primed by the context and then followed by disambiguating information which was either consistent or inconsistent with the primed meaning. The ambiguous word and the disambiguating phrase were either contained in one sentence or separated by a sentence boundary. Daneman and Carpenter used eye fixation data and answers to questions probing the meaning of the ambiguous word to infer the processing carried out by readers with large (RS = 4 - 5.5), medium (RS = 3 - 3.5) and small (RS = 2 - 2.5) reading spans.

All readers, regardless of reading span, increased gaze durations on the first disambiguating word, the whole disambiguating phrase, and the last word of the disambiguating phrase when an inconsistent meaning of the ambiguous word had been primed. Thus, all groups appeared to detect the inconsistency. However, reading span affected the ability to recover; small-span readers gave 52% fewer correct responses to probe questions when the passage was inconsistent than when it was consistent, compared to 40% fewer for medium-span readers and 22% fewer for large-span readers.

Small-span readers were also more adversely affected by sentence boundaries. In inconsistent passages, they gave 22% fewer correct responses to probe questions when a sentence boundary intervened between the ambiguous word and its disambiguating phrase than when this information was contained in one sentence. In contrast, medium-span readers gave 11% fewer correct responses, while large-span readers were unaffected by the sentence boundary. According to Daneman and Carpenter, the less efficient processes of small-span readers during sentence

wrap-up displace verbatim information from working memory, making it more difficult for them to retrieve and reinterpret the ambiguous word after a sentence boundary.

Daneman and Carpenter (1983) emphasise the importance of overall processing efficiency in the recovery from comprehension errors. In contrast to this largely cognitive view, the detection and repair of errors is often thought of as 'metacognitive'. Certainly there are metacognitive elements within the Just and Carpenter (1980) model. However, as Daneman (1984, p. 388) points out, in the case of error recovery it is not just a question of having appropriate metacognitive goals, but also of having the processing capacity required to achieve these goals.

> *"The development of strategies to monitor and repair comprehension errors has frequently been considered as evidence for the development of 'metacognitive' or 'metalinguistic' awareness. However, little attempt has been made to account for the mechanisms responsible for the development of such skills. The interesting implication that emerges from our study is that the type and sophistication of spontaneous attempts to correct reading errors may be best understood in terms of the interaction of these attempts with the information-processing characteristics of the reader."*

There is some evidence to support Daneman's (1984) contention that skilled readers have a greater functional working memory capacity than less skilled readers as a result of more efficient processing, and that this greater working memory capacity enables skilled readers to comprehend more effectively. However, Walczyk has questioned whether efficiency of subcomponent processing has such a

direct relationship to comprehension performance. Walczyk (1993) proposes a compensatory model where, beyond a certain level of skill, subcomponent inefficiency does not affect comprehension because readers can compensate for their inefficiencies. For example, they may read more slowly so that the rate of word encoding and lexical access is commensurate with their ability. In the case of inefficient pronoun resolution, readers can compensate by looking back in the text and rereading sections. However, according to Walczyk, subcomponent inefficiency will affect comprehension when pressure is placed on performance, as when reading time is limited or rereading is not possible.

Walczyk (1990) examined the relationship between subcomponent processes and comprehension among grade 4 children. Decoding speed and accuracy were assessed by having children read a list of unfamiliar words. To assess semantic memory access, children decided whether or not sets of two nouns belonged to the same superordinate category. In the working memory test, children heard pairs of sentences followed by a probe word which had occurred in the first sentence. Their task was to name the word which had followed the probe. Three measures of comprehension were included, the Iowa Test of Basic Skills reading subtest, a sentence verification test, and an error detection test. The sentence verification test involved judging whether or not test sentences had the same meaning as sentences in a test passage, and the error detection test involved finding explicit and implicit logical inconsistencies.

Decoding speed, decoding accuracy, and semantic memory access speed were intercorrelated; scores for semantic memory access accuracy were near ceiling level and were not included in the analyses. Working memory efficiency was correlated with only one lower-level measure (decoding accuracy) and only one

comprehension measure (sentence verification). With regard to comprehension measures, Iowa and sentence verification scores were correlated with each other, and both were also related to decoding speed and accuracy. However, error detection scores were not significantly correlated with any other measure. In a second experiment, error detection was related to scores on an inference test, so Walczyk (1990, p. 494) suggests that the strategic activity of going beyond the information given is crucial for error detection, rather than lower-level subcomponents of reading.

> *"This implies a threshold beyond which efficient encoding, semantic access and logical proposition integration afford no facilitating effects for strategic reading activities, at least as measured by error detection. Thus, a hierarchical relation between verbal efficiency and strategic activities may exist: For fourth graders, local text processing is sufficiently automated so that higher-level (strategic) text modelling is dependent on other factors. This interpretation does not deny but only constrains the importance of verbal efficiency."*

However, some of Walczyk's results are puzzling. Working memory capacity was not related to Iowa reading scores, although Goldman et al (1980) and Perfetti and Goldman (1976), using a similar probe procedure and children of a similar age, found that good readers as defined according to the MAT reading subtest had significantly greater working memory capacity than poor readers. Walczyk found no relationship between error detection and working memory capacity, although Daneman and Carpenter (1983) found that readers with a large reading span, that is, working memory capacity defined by a different measure, were

significantly better at error detection than readers with a small span. Walczyk also found no relationship between error detection and Iowa reading scores in his first experiment, although many studies to be discussed in the final section of this chapter show that good readers are significantly better at detecting logical inconsistencies than poor readers. Certainly Walczyk did not find the relationship between lower-level processes and working memory capacity on the one hand, and working memory capacity and comprehension on the other which writers such as Daneman (1984) have postulated. However, as indicated, this is not the only aspect of this study which is difficult to explain.

In a second study, Walczyk and Raska (1992) used children in grades 2, 4 and 6, and most of the same tests. The sentence verification test was replaced by inference tests. In one test, stories to be read contained several suggestive facts from which inferences could be drawn. The inferences were included as one of the options in a multiple-choice test. The other test involved free recall of a story in which certain facts were suggested but not explicitly stated.

In this study, working memory capacity was correlated with all measures of comprehension. However, working memory capacity was related to speed of decoding and semantic memory access only for the grade 6 children, and not for those in grades 2 and 4. Thus, working memory capacity was related to comprehension in the manner predicted by Daneman (1984), although the basis of working memory capacity differences in grades 2 and 4 is still not clear, since this measure was not significantly correlated with the efficiency of lower-level processes. However, it may be that lower-level processes, including all reading subcomponents necessary to encode information into propositional form, each make a small contribution to working memory capacity, but too small to individ-

ually appear as a significant correlation. Working memory capacity would then vary in relation to the combined efficiency of these subcomponents, rather than the efficiency of any single one. Such a view is not inconsistent with Daneman's (1984) position.

In summary, in this section it has been shown that children defined as good readers according to standardised reading tests recognise words and access information from semantic memory more rapidly than poor readers. According to the Daneman (1984) model of comprehension, the greater processing efficiency of good readers, including the above word-level processes, should leave more working memory capacity free for higher-level comprehension processes. Daneman and Carpenter (1980) devised the Reading Span test as a measure of functional working memory capacity, and showed that reading span was related to effectiveness of pronoun resolution and passage comprehension. Daneman and Carpenter (1983) also showed a relationship between reading span and recovery from inconsistent interpretations. Walczyk (1993) has questioned this relationship between word-level processes, working memory capacity and comprehension, arguing that beyond a certain threshold word-level processing is not a significant determinant of comprehension performance. However, taking into account the many studies referred to in this section, it was suggested that lower-level processing efficiency could influence working memory capacity and comprehension in a global way, without any one process individually showing a significant correlation to working memory capacity and comprehension. Indeed, the Reading Span test and the probe procedure used by Walczyk (Walczyk, 1990; Walczyk and Raska, 1992) measure the combined efficiency of all subcomponents contributing to proposition formation.

When teaching a normal, intact class of upper primary school children, as is the intention in this study, it is likely that there will be considerable variation in processing efficiency among them. However, in an intervention programme of limited duration, it is doubtful whether a focus on lower-level skills would be the most efficient way to improve comprehension of expository texts. In order to learn from such texts, it was shown in chapter one that strategic reading activities are required, such as generating and interrelating main ideas. Increased processing efficiency, while possibly providing a better foundation for strategic processing, does not necessarily enable children to carry out this type of processing which, as will be shown in a later section, they generally find difficult.

A few further points about processing efficiency can be made. The studies referred to in the first part of this section consistently show differences in the speed of word recognition between good and poor readers, and hence speed of word recognition appears to be a very important skill. However, these studies are done in a way which maximises observable differences. Good and poor readers are usually defined according to standardised reading tests, which are often timed. Under these circumstances, a strong relationship will appear between test performance and reading speed. It is therefore not surprising that children who read slowly do not perform well and may be categorised as poor readers. However, these same children may show themselves to be better comprehenders in normal classroom reading when they can take as much time as they need to finish a task. In this case, the relationship between reading speed and comprehension may be attenuated, as suggested by Walczyk's (1993) compensatory hypothesis. Indeed, compensatory strategies can be encouraged as part of a training programme, for example, by encouraging children to read at a pace where they can comfort-

ably process the information and pausing if necessary rather than trying to read quickly, or rereading if an idea is difficult to integrate. Finally, in a normal upper primary classroom it can be assumed, in the absence of information to the contrary, that the word-level processing of most children is efficient enough to enable them to benefit from strategy training.

Nevertheless, processing efficiency is important, and in an intact class it will be necessary to identify any children who do have decoding problems and/or impoverished semantic knowledge, since such children may not benefit from strategy instruction and ideally should receive remedial teaching in these areas. Also, in an on-going training programme, such as a teacher working with children over a whole year, vocabulary training such as that undertaken by Beck, Perfetti and McKeown (1982) could improve comprehension of expository texts. This approach aims to increase both speed of access to semantic information and the richness of the meaning representation, and hence should contribute to processing efficiency. Since expository texts from which children learn are likely to contain new concepts, children of all ability levels could benefit from such instruction, especially when used as an adjunct to strategy instruction.

2.3 Integration

In the first chapter it was pointed out that various processes are used to link successive segments of text into a coherent whole, including the establishment of lexical ties such as part-whole relationships, the assignment of referents to anaphoric expressions, and the construction of bridging inferences. Without such processes, the mental representation of the text would resemble a list of sentences, rather than a coherent discourse. Hence, these integrative processes are important, and could contribute to differences in comprehension ability between good and poor readers.

With regard to pronoun assignment, Frederiksen (1981) refers to one of his studies with high school students of unspecified age. These students were classified into four groups of reading ability according to the Nelson Denny Reading Test. Reading time was recorded as they read one sentence at a time. The sentences either contained a pronoun, or repeated a previous noun phrase. Reading times were longer overall for sentences containing pronouns, presumably because the pronoun entailed more decision making about an appropriate point of attachment. However, this difference was most marked amongst the least skilled group. Thus, increased skill level was associated with greater efficiency in the integrative process of pronoun assignment.

Yuille and Oakhill (1988) examined anaphor-antecedent integration among 8 year-old children formed into two skill groups. The groups were matched on sight vocabulary and oral reading ability, but differed in comprehension ability according to the Neale Analysis of Reading Ability. The narrative text to be read contained four types of anaphoric expressions, namely, pronouns, ellipsis and

substitution of verb phrases, and lexical ties such as part/whole relationships. The anaphor was either in the sentence immediately following the antecedent (immediate) or more than two sentences removed from the antecedent (remote). The children were asked what each of the target anaphors stood for with the text present, to eliminate memory problems.

Skilled readers correctly identified more antecedents than less skilled readers. The importance of anaphoric resolution to comprehension was particularly evident in the episode containing the 'punch line'. In this story, a boy went fishing in a boat, but became so seasick that he couldn't eat his lunch. As a result, he threw his sandwiches to the seagulls, one of which dropped a fish into the boat. The captain commented: *"That's in return for the sandwiches."* When asked what 'that's' referred to, 9 of the 16 less skilled readers could not answer this question, compared to only 2 of the 16 skilled readers. This failure to identify an appropriate antecedent means that a main point of the story was not understood by over half of the less skilled readers.

As another indication of the relationship between reading skill and anaphoric resolution, comprehension scores accounted for 48% of the variance in scores for remote anaphors, and 25% of the variance for immediate anaphors. Yuille and Oakhill (1988, p. 185) comment on their results as follows:

> *"... the two skill groups appeared to behave differently in resolving anaphors: less skilled comprehenders seemed to look for the most immediately preceding item of suitable syntactic form, or used world knowledge, inappropriately, whereas skilled readers presumably relied more on the model of the story they had built up from the text itself, and integrated this more appropriately with their world knowledge.*

... The results indicate that anaphor resolution is a significant problem for some children, and is affected by distance of the antecedent and by comprehension skill."

Bridge and Winograd (1982) investigated the comprehension of a wider range of cohesive ties. They included referential ties such as pronouns, lexical ties such as synonyms and superordinates, and conjunctive ties such as *then* or *and.* Five examples of each type of tie were deleted from a text to form a cloze test. The ninth grade subjects were divided into two skill groups according to the reading subtest of the Comprehensive Test of Basic Skills. One group included students in a remedial class who were reading at least two years below grade level, while the other group scored over the 68th percentile.

Skilled readers performed significantly better than less skilled readers, both in supplying a correct word to insert into the blank for each type of cohesive tie, and in giving adequate explanations as to how they had selected an appropriate word.

In summary, these studies indicate that skilled readers are more efficient than less skilled readers at finding appropriate points of attachment in their text representation for successive concepts as they are processed. This difference is an important one, as shown in the Yuille and Oakhill (1988) study for example, because without efficient integration, readers may miss the main point and fail to construct a coherent representation of the text.

Turning now to the subject of inferences, many studies have shown that with increasing age and skill, students are more likely to construct inferences from their reading. However, it is difficult to draw conclusions from these studies as to the contribution of observed differences in inference generation to overall

comprehension performance. First of all, the types of inferences investigated vary considerably. In the first chapter it was pointed out that while reading, skilled readers only routinely generate bridging inferences, as well as some elaborative inferences that are strongly suggested by the text. It is these inferences that are likely to affect comprehension performance. However, studies on inference generation among children do not always distinguish between inferences necessary to establish cohesion in a text and inferences which could be called 'optional extras'. Another reason why it is difficult to draw conclusions from these studies is that different testing procedures are used, including free recall, probed recall, cued recall and responses to short answer questions. These approaches may have different implications as to when the inferences were constructed. For example, when answering inference questions after reading, it is not certain that responses are retrieved from the text representation, rather than being generated in response to the question from explicit information in memory. In other words, it is not certain that the inference was made during reading rather than during testing. If the latter is the case, one would not expect the skill being demonstrated to directly influence overall comprehension ability.

In the section to follow, the studies of inference generation among children of various ages and reading skill levels will be briefly described before their implications are examined in more detail.

Paris and Lindauer (1976) presented single sentences, in which the instrument of the action was either explicit or implicit, to children in grades 1, 3 and 5. Recall was cued with the instrument. Children in grades 1 and 3 recalled fewer sentences with implicit instruments than sentences with explicit instruments. However, the grade 5 children recalled the two types of sentences equally well, suggesting that

the implied instrument was encoded as part of the memory representation and hence the instrument cue was effective in eliciting recall. However, it should be noted that even children in grade 1 made effective use of the cue to recall sentences with implicit instruments if they had first acted out the sentences. Thus, it was not lack of background knowledge which prevented them from generating inferences in the first experiment.

Paris, Lindauer and Cox (1977) conducted a similar study, but presented sentences which contained either the implicit or explicit consequences of an action to adults and children in grades 2 and 6. Subjects were given both an explicit noun and the consequence as recall cues for each sentence. The major age difference in performance occurred in sentences with an implicit consequence. For grade 2 children in this situation, consequence cues were significantly less effective than the explicit noun cues, while for adults and grade 6 children the effectiveness of the two different types of cues was not significantly different. Again, it appears that with increasing age subjects were more likely to spontaneously incorporate inferences in their memory representation of the sentences. However, grade 2 children could make effective use of consequence cues to recall sentences in the implicit condition if they first generated a story about each sentence, emphasising the earlier point that it is not lack of background knowledge preventing them from constructing inferences.

Paris and Upton (1976) used short narrative passages to investigate what they refer to as lexical and contextual inferences. The former included instrument inferences associated with particular verb phrases (chop wood-axe) and semantic entailments associated with particular nouns (the sister-the girl). The latter involved inferences associated with temporal and causal sequences and depended

on relations expressed in larger segments of text. Children from kindergarten to grade 5 heard the passages and then answered yes/no questions, half of them based on explicitly stated information and half on implied information. For both types of questions, the percentage answered correctly increased monotonically with age.

Paris and Upton (1976) in their second experiment asked children in grades 2 and 4 to answer explicit and inference questions about short narrative passages and then to recall the passages. With regard to questions, when scores for the explicit questions were used as a covariate in the analysis of inference question scores, there was no longer a significant age-related increase in lexical inferences (semantic entailment and instrument inferences), although there was still an increase with age in contextual inferences (temporal and causal sequences). Thus, recall of literal information did not totally account for the greater number of contextual inferences produced by older children, although the more localised lexical inferences were closely associated with recall of explicit information.

With regard to free recall, there was a significant improvement in performance with age. Although overall grade was the best predictor of explicit recall, contextual inference scores also explained a significant amount of variance, both across grades and within grades. Contextual inferences increased in importance as a predictor of explicit recall with increasing age. Paris and Upton (1976, p. 667) explain this relationship between inferences and explicit recall as follows:

> *"If a person constructs implied relationships and integrates them with original information, the derived memory representation may permit a temporarily ordered, logical, sequential unit that can be stored parsimoniously and accessed readily."*

Johnson and Smith (1981) used a long but simple narrative comprised of 16 paragraphs and containing 15 sets of premises which supported 15 inferences. Premises were either stated within one paragraph or in more than one paragraph. In one version of the story inferences were stated explicitly, while in the other they were not. As soon as all the necessary premise information had been presented, the subjects answered one question tapping inferential information and one tapping premise information. The subjects were children in grades 3 and 5.

Overall, children in grade 5 answered more inference questions correctly than children in grade 3. However, the relative performance of the two groups was modified by text version (implicit, explicit) and distance between premises (in one paragraph, in more than one paragraph).

Johnson and Smith calculated the percentage of inference questions answered correctly given that all necessary explicit premise information had been correctly recalled. Following the explicit version of the story, where 'inference' information was stated, there was no significant difference in the performance of grade 3 and grade 5 children. However, following the implicit version, where information had to be inferred, the grade 3 children performed significantly more poorly than those in grade 5. In other words, given that all relevant premise information was available in memory, the younger children were still less likely to construct their own inferences. As Johnson and Smith (1981, p. 1220) conclude: *"Although memory for the premises is an important factor, it is clearly not the sole factor. Younger children are less likely than older children to make an inference when reading a story, even when all the premise information is available in memory."*

Johnson and Smith also examined inference scores given premise recall in relation to distance between premises. Given that all premise information was

correctly recalled, grade 3 students were just as likely as grade 5 students to answer inference questions correctly when premise information was in one paragraph. However, the younger children were significantly less likely to answer inference questions correctly, again given that premise information was correctly recalled, when premise information was spread across paragraphs. In the latter case it was more difficult to integrate premise information, and here the younger children were less able to generate appropriate inferences.

These studies consistently show that older children in the primary grades are more likely than younger ones to construct inferences from explicit information being read. Similar differences have also been found between children of the same age but different levels of reading skill.

Oakhill (1984) adapted four of the passages used by Paris and Upton (1976), using short-answer rather than yes/no questions to assess recall of explicit and implied information. Two groups of 7-8 year-old children were matched on vocabulary knowledge and ability to read aloud, but were significantly different in terms of comprehension according to the Neale Analysis of Reading Ability.

Oakhill compared answers to literal and inferential questions when the text was still available for reference or no longer available. Skilled readers answered more literal questions correctly than less skilled readers, but only when the text was no longer available. However, skilled readers answered more inference questions correctly regardless of the condition. So again, access to explicit information was not sufficient to ensure inference generation among less skilled readers.

Bridge and Tierney (1981) examined free recall of good and poor readers in grade 3. All students could achieve an oral reading accuracy of at least 85%, but good and poor readers were distinguished according to the reading test of the

Stanford Achievement Test. The students read one narrative and one expository passage. Following free recall, general probes were instituted, referring to information already given in free recall and asking general questions such as: Can you tell me more about that?

In free recall, good readers produced more explicit information and more inferences than poor readers for both the expository and the narrative passage. In response to general probing, good and poor readers produced a similar number of additional inferences for the narrative passage, but good readers produced more than poor readers for the expository passage. Bridge and Tierney note that this passage was more difficult for all readers, and the probing encouraged them to elaborate on and make sense of this more difficult text, especially in the case of the good readers. However, the good readers comprehended the narrative well, and it was the poor readers who gained a relatively larger benefit from the stimulus of probing, raising them to the level of the good readers in terms of the number of additional inferences produced.

Holmes (1987) compared inferences generated from pictures, texts describing the pictures, or the pictures and print combined. The subjects were children in grades 5 and 6 who all had adequate word recognition skills but who differed significantly in comprehension ability according to the reading subtest of the SRA Achievement Series. After examining each stimulus in either picture, picture and print or print only form, students answered five inferential questions. Overall, skilled readers answered more questions correctly than less skilled readers, but in terms of the various conditions, the skill difference was significant only for print. In other words, less skilled readers were as capable as skilled readers of drawing inferences from pictures or print supported by pictures, suggesting that it is not

lack of background knowledge or a general reasoning problem which inhibits less advanced readers from spontaneously constructing inferences.

Bransford, Stein, Vye, Franks, Auble, Mezynski and Perfetto (1982) report a study in which successful and less successful students in grade 5 read a passage describing robots constructed for different window cleaning jobs. In one version of the passage the robotic structures described were explicitly linked to their function, while in the other version these links were not stated. Successful and less successful students were distinguished by scores on the language and reading subtests of the Comprehensive Test of Basic Skills.

Less successful students were able to answer more questions explaining the function of various structures after reading the explicit version than after reading the implicit version. For successful students, however, performance was similar after reading the two versions, suggesting that these students, unlike their less successful peers, had spontaneously elaborated structure-function relationships while reading the implicit version.

One reason why the generation of inferences is important is to increase the interconnectedness of information, thereby also increasing its retention. For example, Stein, Bransford, Franks, Owings, Vye and McGraw (1982) chose successful, average and less successful students in grade 5 on the same basis as Bransford et al (1982). These students read a set of sentences describing activities performed by particular kinds of men. The relationship between the type of men and the type of activities was deliberately made arbitrary to model the situation of a novice working in a new field. Subjects were asked to generate an elaboration for each sentence, which Stein et al classified as either precise or imprecise. They refer to a precise elaboration as one that clarified the relationship between the

type of man and the activity performed.

Successful students produced more precise elaborations than average students, who in turn produced more than less successful students. Recall of the sentences followed the same pattern. Analysis of variance showed a significant main effect of precision on recall, indicating that precisely elaborated sentences were recalled better.

There could be several reasons why less mature and less skilled students construct fewer inferences, including inadequate background knowledge and reasoning ability, incomplete access to premise information, difficulty with integrating premise information, and inadequate metacognitive knowledge, relating particularly to the goals of reading and strategies for achieving those goals.

None of the studies described gives any indication that lack of background knowledge or reasoning ability are responsible for the lower rate of inference generation. The most telling observation in this regard is that under favourable circumstances less mature or less skilled readers can perform at a level comparable to older or more skilled readers. These favourable circumstances include acting out stimulus sentences (Paris and Lindauer, 1976), elaborating stimulus sentences into a story (Paris et al, 1977), and using pictures rather than text as the stimulus material (Holmes, 1987).

The studies by Johnson and Smith (1981) and Oakhill (1984) suggest that incomplete access to premise information is also not the reason for lower rates of inference generation. Johnson and Smith show that when inference scores were calculated conditional upon all premise information being correctly recalled, grade 3 children still performed more poorly than grade 5 children. Similarly, Oakhill shows that skilled readers answered more inference questions correctly than less

skilled readers even when the text was available for inspection during testing.

On the other hand, there is some evidence that difficulty with integrative processing may be involved in lower rates of inference generation. In the Johnson and Smith (1981) study, inference scores conditional upon recall of premise information were not significantly different for grade 3 and grade 5 students when premise information was contained in one paragraph. However, the older children performed significantly better when premise information was contained in more than one paragraph. In the latter case, integration of the premises as a basis for generating an appropriate inference was made more difficult by the increased distance separating the two pieces of information.

Holmes (1987) in her study notes that the most common error in answers to inferential questions involved using only one of the relevant clues in the text. For example, a student would infer that the weather was warm because the sun was shining, but fail to integrate the fact that people were wearing coats. Stein et al (1982) make a similar observation when they note that in trying to generate precise elaborations, less successful students tended not to relate the particular type of man and the particular type of activity in their answers. As in the case of the Holmes study, the less successful learners were not using all the relevant information to generate an answer.

McCormick (1992) investigated errors in answers to inferential questions in some detail. In her study, reading disabled students with normal I.Q. in grade 5 read either an expository or a narrative passage every week for 20 weeks. They read along with an audiotape to minimise decoding difficulties, then reread and discussed the passage before answering literal and inferential questions. All questions tapped information central to the passage, that is, the macrostructure of

expository passages or the causal chain in narratives. Of the literal questions, 70% were answered correctly, compared to only 40% of the inferential questions, plus a further 36% which were partially correct.

McCormick classified errors in response to inferential questions into 26 categories. Of these, four accounted for over half the errors. These categories were: producing an under-developed written response which did not make its meaning clear, over-reliance on background knowledge to the exclusion of textual information, producing answers unrelated to the main points of the text, or answers that were too specific to reflect global concerns. In relation to the definition of these last two categories respectively, McCormick (1992, p. 74) states:

> *"Although there were attempts to infer in these responses, apparently the wrong text cues had captured the students' attention or they had not retrieved from memory the information of greatest significance. ... Errors of this type may indicate lack of experience in combining all pertinent information to form generalizations."*

That the less skilled readers have difficulty integrating all relevant pieces of information to generate inferences has already been noted in the Holmes and Stein et al studies. They may not combine relevant premises in the formation of an inferential response, or fail to integrate detailed information into a coherent macrostructure. It is then difficult to generate inferences consistent with the main ideas in the text.

Some of the studies cited link inferencing performance to metacognitive knowledge. For example, Bransford et al (1982) point out that when successful and less successful students in grade 5 were asked how difficult either the implicit or

the explicit version of the robot passage would be to learn, the successful students rated the implicit version with its unstated relationships as more difficult, and spent more time studying this passage. Judging by recall performance, these students probably spent some of the extra time elaborating relationships. Less successful learners, on the other hand, spent more time studying the explicit version. Bransford et al suggest that their main study strategy was rereading and, since the explicit version contained more words, it took longer to reread. It could be, then, that the successful learners had more metacognitive knowledge about task and strategy variables (Flavell, 1981), that is, what makes a task difficult and the range of strategies that could be adopted to deal with the difficulty.

Similar explanations are put forward by Paris and Upton (1976), Stein et al (1982), Paris and Lindauer (1976) and Paris, Lindauer and Cox (1977). Paris and Lindauer (1976, p. 225) comment as follows:

> *"Young children simply do not approach laboratory memory tasks with strategic plans. Their goal appears to be a monitoring task, whereby the words in the sentence are repeated and 'checked off' as a comprehensible sentence. Their task is accomplished at this point and the sentence is not rehearsed, monitored or recomprehended again. The goal of the child's behaviour does not appear to involve strategies directed towards storing the experience for later retrieval."*

If less mature and less skilled readers set limited reading goals, such as decoding fluently or checking that each sentence is comprehensible, then they are unlikely to generate the inferences which would be required if their goal were to construct a coherent text representation. However, while appropriate reading

goals are necessary, they are not sufficient; students must also be able to apply strategies which will achieve these goals. Stein et al (1982) point out that directing their less successful students to generate precise elaborations had no effect on their behaviour; more extensive step-by-step practice was required to master the strategy.

This last point is an important one, for although the level of metacognitive knowledge appears to contribute to children's propensity to generate inferences, this knowledge must also be considered in relation to processing difficulties. In the Johnson and Smith (1981) study, age differences were found when premise information was in more than one paragraph. In the Holmes (1987) study, skill differences emerged when the stimulus material was text, but not when it was in pictorial form. Thus, it was not case that younger or less skilled readers never generated inferences, but they did so less successfully as processing demands increased.

As a final comment on metacognition, some authors have suggested that motivational factors may reduce the inferential processing of less skilled readers. Motivational factors are included under the heading of General Strategy Knowledge within the model of metacognition set out by Borkowski and Kurtz (1987) described in chapter one. Holmes (1987) suggests that poor readers, due to their lack of confidence in handling print, may adopt the strategy of skipping anything that appears difficult and may not be willing to put in the time and effort necessary to construct a coherent text representation. McCormick (1992, p. 73) similarly suggests that less skilled readers have learned inappropriate strategies with inadequate integration of textual information:

"It may be that low achieving students have had less access to text

information because of word recognition difficulties and over time have developed a pervasive strategy of simply guessing at possible responses based on knowledge already stored in their schemata."

As indicated earlier, observed differences in inference generation between more and less advanced readers cannot be related directly to their overall comprehension ability, since a range of different inferences were included in these studies. The demonstration that skilled readers are more able to generate instrument inferences in a laboratory experiment, for example, does not necessarily mean that this inferencing ability is a reason for their high level of reading ability. Such a conclusion would only be valid if such inferences in the context of the particular stimulus material were necessary to construct a coherent text representation. If they are one of the many optional elaborations which could be generated, then their contribution to comprehension efficiency is less clear.

McCormick (1992) presents one of the few studies to indicate that inferences were related to information of central importance to the passages, that is, to the macrostructure of expository texts and the causal chain of events in narratives. The reading disabled students in this study answered only 40% of inference questions correctly, although there was no normal group of readers against which to compare this performance. In the study by Bransford et al (1982), the structure-function relationships of the window cleaning robot were also central to the passage, so it is of some importance that less successful students inferred fewer of these relationships than their more successful peers. In these cases at least, it appears that inference generation contributes to the more advanced readers' superior level of comprehension.

To summarise, it was shown in chapter one that some inferences are essential

to the construction of a coherent text representation. While the studies discussed in this section do not always make it clear whether the inferences under investigation are of this essential kind, there is some evidence that less advanced readers, both in age and skill level, are less likely to construct these important inferences. A number of reasons were suggested for this difference, relating to integrative processing, the setting of appropriate reading goals, the application of strategies to attain these goals, and sufficient self-confidence to persevere with such strategies. Potentially, then, the generation of important inferences is a skill which could be targeted in comprehension instruction programmes, and such instruction will be examined in chapter three.

2.4 Macrostructure formation

In chapter one it was pointed out that the macrostructure of a text consists of a statement of the overall topic at the highest level of generality, and the most important points in the development of this topic at the lower levels. The macrostructure is formed by deleting trivial information, selecting explicit macropropositions, and condensing relevant details into generalisations and constructions. Having generated a macrostructure while reading an expository text, a reader is in a position to recall something like: The text discusses the problems caused by industrial waste, and outlines the advantages and disadvantages of three possible methods of disposal. The first method is ..., and so on.

In chapter one several important functions of macrorelevant information were discussed. For example, Bransford and Johnson (1972) and Schallert (1976) show that using thematic information to guide processing influences the construction of conceptual meanings and relationships. Misinterpretations can occur when thematic information is not used in this way. For example, Daneman and Carpenter (1980) report that poor readers produced responses to fact and pronoun referent questions that contradicted the whole theme of the passage, and McCormick (1992) reports that disabled readers produced responses to inference questions that were unrelated to the main points of the text.

Another important function is to assist in organising the memory representation and recall. Lorch, Lorch and Inman (1993), for example, show that when macrostructure formation is made easier through the use of signalling, readers recall more of the topics and in a more organised form. Loman and Mayer (1983) show that when macrostructure formation is made easier through signalling read-

ers recall more of the important information in a text and are more able to use this information in creative problem solving. Therefore, being able to use macrorelevant information while reading is an important skill in learning from expository text.

Before examining differences between good and poor readers in relation to macrostructure formation, two important points made in chapter one will be briefly reiterated. These points concern interpretation of the ability to determine important ideas as a metacognitive skill, and the focus on knowledge and use of a small number of conventional text structures to determine the most important ideas.

The ability to determine the most important ideas of a text is sometimes considered largely as a metacognitive skill. For example, Brown and Smiley (1978) consider this ability under the heading of the task and strategy variables of metacognitive knowledge. Speaking of the ability to determine the most important ideas, Brown and Smiley (1978, p. 1087) state:

> *"... students need knowledge concerning texts, knowledge concerning strategies, and knowledge concerning the interface of these factors before they can study strategically. ... knowledge about texts must consist of general knowledge about consistent features of all texts and specific knowledge about the particular exemplar at hand."*

In the previous chapter, metacognitive knowledge was acknowledged as an important factor influencing macrostructure formation, particularly via the goals that are set for reading. Goals influence both the content and extent of macrolevel processing. McKoon and Ratcliff (1992) also emphasise that macrorelevant information will not be inferred or used to guide processing in the absence of specific

goals and strategic activity. Once the goal of constructing a framework of main ideas has been established, the reader needs strategies for achieving this goal. Looking for and making use of signalling information is a useful strategy, and knowledge about text structures should assist the reader in applying this strategy.

While metacognitive knowledge is undoubtedly important, the other factors mentioned in chapter one must also be kept in mind, these being level of background knowledge about the topic, and the processing demands of the text. Roller (1990), for example, points out that structural cues are not very helpful when the topic is unfamiliar; the reader has difficulty constructing meanings for basic concepts and is not in a good position to determine the global relations between them. With regard to processing demands, van Dijk and Kintsch (1983) refer to a study published in German by Schnotz, Ballstaedt and Mandl which showed that adults were not able to summarise the main points of a complex passage until after a second reading. Presumably the first reading made such large demands on cognitive capacity that these skilled readers were unable to form a macrostructure. Brown and Smiley (1978) also note that skilled readers may not be able to identify important ideas in a very complicated text. Thus, metacognitive knowledge is necessary but not sufficient for macrostructure formation.

Turning now to the second point, Meyer and her colleagues consider that knowledge and use of conventional text structures plays an important role in determining reading ability. For example, Taylor (1980, p. 400) states that: "*... differences in reading ability among adults may be due, in part, to differential use of expository text structure.*" Meyer identified five such text, or top-level structures, including cause-effect, problem-solution, comparison-contrast,

description and collection (Meyer, Brandt and Bluth, 1980).

Like Taylor (1980), Meyer, Brandt and Bluth (1980, p. 75) identify use of text structure as an important factor distinguishing good and poor readers.

> *"The top-level structures in these texts were expected to provide readers with a systematic, organized strategy for encoding information from text and retrieving it from memory. We expected good readers to approach text with knowledge about how texts are conventionally organized and a strategy to seek and use the top-level structure in a particular text. Use of the top-level structure was expected to relate to the amount of information remembered; organization has been shown to be a crucial variable in learning and memory. In contrast to good readers, poor readers were expected to approach text without knowledge of texts' organization and no effective strategy for utilizing the top-level structure of a passage."*

However, it was pointed out in chapter one that natural texts do not consistently conform to the text types described by Meyer. Just and Carpenter (1987) state that different types of text are sometimes divided by no more than a fine line, and Kintsch and Yarborough (1982) add that texts are usually composed of several different structural units. Thus, a description could be embedded in a problem-solution text, or a cause-effect sequence could be embedded in an overall comparison-contrast structure. Indeed, Hiebert, Englert and Brennan (1983), although working within the Meyer framework, acknowledge that pure text types are rare.

Nevertheless, attention to text structure cues during processing is of value, as set out by Englert and Hiebert (1984, p. 65).

> *"Text structure specifies the logical connectives in text, as well as the subordination of some ideas to others. These text structures provide organizational patterns that help readers identify and tie together the most essential related propositions."*

It was pointed out in the previous chapter that knowledge of conventional text structures was one source of information used by readers to generate a macrostructure. However, van Dijk and Kintsch (1983) emphasise that an expedient strategy operates on many kinds of information. The reader should be able to use a variety of cues, including headings, topic sentences, logical connectives, and general knowledge to infer implicit topic statements; in this way the reader will be sufficiently flexible to deal with texts that do not have a pure text structure. Thus, Meyer's contention that readers should approach texts with knowledge about how texts are conventionally organised should be seen as a subset of more general macrostructure formation strategies.

In the sections to follow, differences between more and less advanced readers, both in terms of age and skill, will be examined in relation to assessing the topical coherence of paragraphs, rating the importance of idea units in a text, including important information in summaries, and knowledge and use of conventional text structures to facilitate recall.

2.4.1 Assessing the topical coherence of paragraphs

A paragraph should contain information related to a specific aspect of a topic. In other words, the ideas should form a coherent unit of meaning. Sometimes this specific topic is explicitly stated in the form of a superordinate topic sentence, but sometimes it is left to the reader to infer the main idea being conveyed by the

paragraph. The ability to determine main ideas is an important one both from the point of view of focusing on and retaining information central to the message of the text, and of determining the contribution made by each paragraph to the overall structure of the exposition.

While there is general agreement that the ability to determine main ideas is an important one, Cunningham and Moore (1986) point out that 'main idea' is an umbrella concept that can be interpreted in different ways. They list nine possible interpretations for the concept, but when subjects in their study were asked to produce a main idea statement for paragraphs, four of these possibilities accounted for over 90% of the responses produced by teachers and would-be teachers and almost 80% of the responses from grade 6 children. The four types were topic sentence, gist, topic and selective summary. A topic sentence is the explicit sentence which most completely states what a paragraph is about. The topic sentence for the sample paragraph given was: 'Some horses are said to become 'loco' or insane from eating locoweed.' The gist is a generalisation formed from detailed information. For example, the gist of the sample paragraph could be: 'Locoweed is poison for horses'. The topic is a statement naming the subject of the paragraph but not how the subject is developed, as for example: 'What happens when horses eat locoweed'. A selective summary combines the most important points in the paragraph, for example: 'Horses go crazy, get sick, and die after eating locoweed'.

Although the four interpretations are subtly different, the term 'main idea' will be retained in the sections to follow because it is so commonly used. However, the more precise terminology will be used when studies make it clear that they are referring specifically to one of the categories identified by Cunningham and

Moore.

There is some evidence that older children know more about paragraph structure, in terms of statable knowledge, than younger children. Garner, Alexander, Slater, Hare, Smith and Reis (1986) asked children in grades 3, 5 and 7 to identify paragraphs and to explain the basis of identification. The children were deemed to be average readers on the basis of teachers' judgements and unnamed achievement test scores. Even the youngest children performed well on the identification task, while the older children achieved perfect scores. The majority of grade 3 and 5 children reported using physical cues such as indentation to recognise a paragraph, but the majority of grade 7 children referred to unity of meaning.

In a study by Garner and Gillingham (1987), children described what made a good paragraph as they attempted to construct one. The children were in grades 5 and 7, and were considered by teachers to be of at least average academic achievement. The older children were significantly more likely than the younger ones to refer to thematic unity and the importance of a topic sentence in their verbal reports. However, Garner and Gillingham also note that there was a low correlation between these verbal reports and the children's performance in constructing paragraphs from the sentences provided. In some cases children reported knowledge which was not demonstrated, and in other cases they demonstrated knowledge which was not reported.

Danner (1976) shows that with increasing age children are more able to demonstrate knowledge of paragraph construction. Children in grades 2, 4 and 6 were given the twelve sentences of each of two passages written on individual cards. Each set of twelve contained three topic sentences and three detail sentences related to each topic. There was a significant increase with age in the

number of children who scored maximum points for logically organising the twelve sentences, that is, forming three paragraphs, each beginning with an appropriate topic sentence, rising from 21% in grade 2, to 67% in grade 4 and 83% in grade 6.

Several studies have shown that older or more skilled readers are better at generating a title for paragraphs or writing a topic sentence to capture the main idea. In other words, they are better at determining the specific topic of the paragraph.

Williams, Taylor and Ganger (1981) investigated whether children could recognise the main idea of a paragraph. Children in grades 4 and 6 read paragraphs either with or without a topic sentence. They chose the best title for half the paragraphs from a set of four alternatives. Of the alternatives given, one related to the specific aspect of the topic being discussed, one to the general topic only, one to a single detail sentence, and one to an aspect of the topic not mentioned in the paragraph. For the other half of the paragraphs, the children generated their own titles.

In the selection task, the older children chose the specific title significantly more often that the younger ones. Similarly in the production task, the older children produced significantly more specific titles than the younger ones, although it should be noted that only just over half the titles produced by the older children were specific. In the case of the grade 4 children, about one third of the titles generated referred neither to the specific nor the general topic of the paragraphs, so these children were not particularly skilled at generating titles to summarise the topic of paragraphs. Contrary to expectation, the presence or absence of a topic sentence did not have a significant effect on overall performance.

In a similar study by Williams, Taylor and de Cani (1984), children in grades 3, 5 and 7, as well as graduate students, either selected the best title for each paragraph from the four options described by Williams et al (1981) or generated a main idea summary statement for each paragraph. Williams at al calculated the number of responses which accurately reflected the specific topic of the paragraphs. Only the normal paragraphs used by Williams et al will be considered here. These paragraphs were composed either of a topic sentence and three specifically related detail sentences or of four detail sentences.

In the overall analysis, older readers performed better than younger ones, the title selection task was easier than the summary sentence generation, and the presence of a topic sentence improved performance. Williams et al do not give further statistical analysis of age comparisons, but present average scores for the four age groups in graphic form. From this information, a steady improvement across the age groups is evident, in particular a sharp improvement in title selection between grades 3 and 5, and a substantial improvement in summary sentence generation between grades 3 and 5, and 5 and 7. Adults performed near ceiling level in both tasks.

Bridge, Belmore, Moskow, Cohen and Matthews (1984) also investigated the effect of topic sentences and title generation, but among a university undergraduate population. From a median split of scores on the Van Wagenen Rate of Comprehension test, half the students were designated as relatively good readers and half as relatively poor readers.

In the first experiment, subjects read a series of separate paragraphs, half with and half without a topic sentence. A time limit was set for reading each paragraph, and after each one subjects generated a title for the paragraph. The

recognition test which followed involved choosing the correct main idea statement from two options.

Good readers generated better quality titles than poor readers, and better quality titles were generated for paragraphs with topic sentences than paragraphs without them. However, there was no interaction between paragraph type and ability, so good readers were able to generate better titles than poor readers even for paragraphs with a topic sentence. There was no ability difference on the recognition test, possible due to a ceiling effect.

In the second experiment, subjects performed one of three tasks after reading each paragraph at their own pace. They either rated the comprehensibility of the paragraph, selected the correct main idea statement from two options, or generated their own main idea statement. After reading, they completed a recognition test, judging whether or not a list of main idea statements reflected the meaning of the paragraphs. They also completed an unexpected recall test of main ideas.

In relation to the comprehension tasks, both good and poor readers rated the paragraphs as equally comprehensible. However, good readers recognised more correct main idea statements, and generated better quality main idea statements than poor readers. The presence or absence of a topic sentence had no significant effect on any of these tasks, unlike in the first experiment where reading was not self-paced and readers may not have had sufficient time to infer their own topic sentences.

In free recall, good readers produced more complete and accurate main idea statements than poor readers. Overall, the generation task produced better recall than the rating task, with the recognition task intermediate between the two. Although there was no interaction between ability and task, Bridge et al note that

poor readers performed best relative to good readers in the generation task. The constructive processing demanded by this task appears to have been particularly beneficial for the poorer readers. There was no ability difference on the recognition test, and the presence or absence of a topic sentence had no significant effect on either the free recall or recognition test.

These studies consistently show that with increasing age children are able to both select and generate better paragraph titles and to generate better paragraph summary sentences. Similarly, good readers in an adult population generate better quality titles and summary sentences than poorer readers.

Another approach to assessing how well readers understand the topical coherence of paragraphs is to have them decide whether or not sentences belong in a paragraph, or to generate follow-on sentences which are consistent with the specific paragraph topic. In the study by Garner et al (1986), children in grades 3, 5 and 7 were asked to discriminate between sentences relating specifically to the aspect of a subject referred to in a topic sentence, relating only generally to the subject rather than to the specific aspect, or not relating to the subject at all. Given the topic sentences and a collection of specific, general and unrelated sentences, the children were asked to build what they thought were coherent paragraphs.

There was no significant age difference in children's ability to include specifically related sentences, but the grade 7 children were significantly more adept that the grade 5 and grade 3 children at excluding both unrelated and only generally related sentences. The grade 5 children performed at much the same level as those in grade 3 in this area, with both groups achieving only about 50% correct, compared to the near ceiling level performance of the grade 7 children.

Garner and Gillingham (1987) asked children in grades 5 and 7 to construct a good paragraph from a topic sentence, four sentences related to the specific topic identified in the topic sentence, and two unrelated sentences. Approximately half the students in both grades successfully excluded both unrelated sentences from their paragraphs. The percentages of students correctly placing the topic sentence in the initial position were 40% and 73% in grades 5 and 7 respectively, a difference which approached significance ($p < .07$). On the basis of these results, Garner and Gillingham (1987, p. 258) conclude that: *"... these students who face a number of reading and writing tasks that require them to attend to structural properties of text are not very well equipped to manage these tasks."*

In the study by Williams et al (1981), children in grades 4 and 6 were asked to determine whether all sentences in a paragraph belonged. Half the paragraphs began with a topic sentence stating the specific aspect of a subject to be discussed. Deviant sentences which did not belong to the paragraph could be either only generally related to the topic or totally unrelated.

Although in absolute terms the older children identified more deviant sentences than younger children, this difference was not significant. Both groups detected more unrelated sentences than only generally related sentences. The presence or absence of a topic sentence again did not significantly affect performance.

However, Williams (1984) did find a significant difference between children in grades 4 and 6 in their ability to assess the specific topic of discourse and sentences deviating from it. Paragraphs were either one or three sentences long, with only half containing a topic sentence. Thus, in the one sentence condition the stimulus was either the topic sentence or a detail sentence, while in the three

sentence condition the stimulus was either the topic sentence plus two detail sentences or three detail sentences. After each paragraph the children judged whether or not a further detail sentence belonged to the paragraph.

The grade 6 children performed significantly better than the grade 4 children, and the presence of a topic sentence improved overall performance. However, of particular interest is the fact that the grade 6 children performed as well when given three sentences without a topic sentence as they did for both paragraph types with a topic sentence. In other words, from the three detail sentences they appear to have been able to abstract the specific topic of discourse and to use it as a basis for judging whether test sentences belonged. In contrast, the grade 4 children performed quite poorly for both types of paragraphs without a topic sentence, achieving just over 50% correct.

Hiebert, Englert and Brennan (1983) investigated readers' ability to determine within-paragraph relatedness in terms of text structure. Paragraphs exemplifying four of the text structures discussed by Meyer et al (1980) and mentioned earlier were used, namely, description, comparison-contrast, sequence and enumeration, the last two being different aspects of Meyer's category of collection. The subjects were college students divided into two ability groups on the basis of the Degrees of Reading Power test.

Readers were given a paragraph containing a topic sentence, three appropriate subordinate sentences and two sentences on the same topic but not consistent with the text structure signalled by the topic sentence. Subjects rated the degree of fit of each subordinate sentence.

The higher ability group performed significantly better than the lower ability group, but the group difference was small for acceptance of appropriate sentences

and much larger for rejection of inappropriate sentences. Thus, the higher ability group was more able to recognise sentences which deviated from the text structure signalled by the topic sentence.

Englert and Hiebert (1984) examined the ability of children in grades 3 and 6 to carry out the rating task with the same four text structures. The children were divided into groups of high, medium and low reading achievement on the basis of the California Test of Basic Skills.

The ratings of the grade 6 children were more accurate than those of the grade 3 children, but only for rejecting sentences inconsistent with the text structure signalled by the topic sentence, not for accepting consistent sentences. The grade 3 children scored only 53% correct for rejecting deviant sentences, compared to 75% for the grade 6 children. However, there were ability differences in both acceptance and rejection, with the high achievement group performing better than the medium and low achievement groups, which did not differ significantly from each other. The high achievement group correctly rejected 74% of deviant sentences, compared to 56% and 55% among medium and low achievements groups respectively. Thus, the older children and the higher achievement group were more able to determine the theme and structure of a paragraph, to relate subordinate sentences to this structure and to determine whether or not the sentences were consistent.

Hiebert, Englert and Brennan (1983) presented their adult subjects with a topic sentence, or a topic sentence and one detail sentence, and asked them to generate two further sentences consistent with the content and text structure signalled by the topic sentence. The higher ability group generated a significantly higher percentage of acceptable follow-on sentences than the lower ability group.

Williams, Taylor and de Cani (1984) presented students in grades 3, 5 and 7 and at university level with paragraphs composed either of a topic sentence and three detail sentences or four detail sentences. The subjects were asked to generate one further sentence consistent with the specific topic of the paragraphs. There was an overall main effect due to age; the graphs presented by Williams et al indicate a large difference between adults and children. While the adults performed very well, only about half the responses from grade 7 students related to the specific topic of the paragraphs, compared to about a quarter from the grade 5 and grade 3 children. Thus, these children were not very skilled at generating specific follow-on sentences.

In summary, some studies have shown that older children are better at rejecting sentences which do not conform to the specific topic of a paragraph (Garner et al, 1986; Williams, 1984; Englert and Hiebert, 1984) and better at generating sentences conforming to the specific topic (Williams et al, 1984). In some studies no age differences with regard to rejecting deviant sentences were found (Williams et al, 1981; Garner and Gillingham, 1987), although in the Garner and Gillingham study children in both grades 5 and 7 performed rather poorly in this area and did not demonstrate a high level of awareness of topical coherence. With regard to ability differences, higher ability readers are better at excluding deviant sentences than lower ability readers (Hiebert et al, 1983; Englert and Hiebert, 1984) and better at generating appropriate follow-on sentences (Hiebert et al, 1983). Thus, there is a clear tendency for more advanced readers in terms of both age and level of reading skill to show greater awareness of the topical coherence of paragraphs.

As a final point, it would be intuitively sensible to assume that the presence of

a topic sentence would assist readers in determining the main idea of a paragraph. However, the effect of topic sentences has been rather mixed. In the Williams et al (1981) study the performance of children in grades 4 and 6 was not significantly improved by topic sentences, whereas the topic sentences did have a significant effect on the performance of children in grades 3, 5 and 7 in the Williams et al (1984) study, and children in grades 4 and 6 in the Williams (1984) study. Bridge et al (1984) found that topic sentences affected the performance of adults only when paragraphs were read under time pressure; when reading was self-paced these mature readers appear to have been able to infer the main idea of paragraphs for themselves.

Williams et al (1981) further investigated why the children in their study did not make better use of topic sentences. These authors initially examined the performance of adults in selecting an appropriate title from four options, generating their own title, and detecting deviant sentences in paragraphs with or without topic sentences. The adults chose the specific title on 87% of occasions and produced a specific title on 80% of occasions. Performance was significantly better with than without a topic sentence. Similarly, the presence of a topic sentence improved detection of deviant sentences within a paragraph.

Finally, the two experiments were replicated with grade 5 students to examine whether they could be induced to make greater use of topic sentences in the identification of specific paragraph topics if these topic sentences were graphically highlighted by having a box drawn around them. The topic sentences had a marginally significant effect on the selection and generation of specific titles. However, in the second task the presence of a topic sentence had a clearly significant effect on the ability to recognise deviant sentences in paragraphs. Of this

finding, Williams et al (1981, p. 863) note: *"The fact that graphic highlighting of topic sentences improved the ability of children to utilize topic sentence information corroborates the notion that readers have to learn what to pay attention to and that this has not been accomplished by the middle grades."*

As suggested by Williams et al, a topic sentence is unlikely to be of benefit to children if they do not devote special processing to it. As mentioned in chapter 1, adults spend longer reading topic sentences than non-topic sentences (Lorch, Lorch and Matthews, 1985). According to Williams et al, children in the primary grades still need to learn to do the same. Williams et al (1984) add a further suggestion, namely, that the younger children in this age group may still need to improve their categorisation skills in order to more precisely understand the relationship between superordinate (topic sentences) and subordinate (detail sentences) information.

While teaching children to attend to topic sentences would appear to be a useful strategy, the point made by Bridge et al (1984) and Winograd and Bridge (1986) should also be kept in mind; a minority of expository texts written for both adults and children have been found to have topic sentences, and even fewer have topic sentences in the initial position of paragraphs. Therefore, if children are to understand the main idea of paragraphs, they need to be able to deal with less than considerate texts.

A study by Hare, Rabinowitz and Shieble (1989) emphasises the point that young readers in particular have difficulty processing more natural texts, as opposed to the well-structured texts often used in experiments, containing explicit topic sentences and a simple enumeration structure. In a first experiment, grade 4 students reading grade-appropriate expository paragraphs were as capable of

underlining explicit topic sentences as students in grades 6 and 11 if the topic sentences occurred in the initial position. However, the grade 4 students were less likely than the other two groups to identify topic sentences as the main idea when these sentences occurred in medial or final positions. As Bridge et al (1984) and Winograd and Bridge (1986) indicate, topic sentences in natural texts are not always placed in the initial position, thus possibly creating a problem for younger readers.

In a second experiment, children from the same grade levels read grade-appropriate expository paragraphs exemplifying four of Meyer et al's (1980) text structures, namely, listing (enumeration), sequence, compare-contrast and cause-effect. Students were directed to underline the main idea if explicitly stated, and to generate a main idea statement if none was present. Although grade 4 students underlined fewer explicit topic sentences than grade 11 students, they nevertheless performed moderately well. However, all groups performed quite poorly when required to generate a topic sentence; even the grade 11 students generated only about one-third correct responses for the easiest two text structures, and less for the other two. It should be noted, though, that the paragraphs used were quite conceptually dense, and the integration of information in them in order to construct an appropriate topic sentence was likely to be difficult.

Nevertheless, this study makes two points regarding the processing of more natural texts. Firstly, even students beyond the primary grades may have difficulty inventing topic sentences and, secondly, determining main ideas is more difficult in some text structures than others. Natural texts include a variety of structures, while experimental texts often have a simple description or enumeration structure. In the Hare et al study, all groups found the compare-contrast and

especially the cause-effect texts more difficult than listing and sequence texts.

Ideally, then, children should be able to determine the main idea of paragraphs with topic sentences that are not in the initial position, with no explicit topic sentences, and with different text structures.

2.4.2 Determining the importance of ideas in texts

In order to form a macrostructure, readers need to be able to use textual cues to identify important information, as well as being able to construct macropropositions not explicitly stated, for example, by subsuming the details of a sequence of events into one global statement of the event. As indicated in the introduction to this section, knowledge of both the topic of discourse and text structure are important; knowledge of conventional structures and methods of signalling global relationships can help readers to identify the high-level structure of a particular text, but without background knowledge about the topic of discourse the reader will have difficulty understanding the relationships being signalled, and will be less able to determine the importance of different ideas in the absence of explicit cues.

Brown and Smiley have conducted several studies to assess children's awareness of structural importance in folktales. As indicated above, fluent readers would be expected to use both structural and contextual knowledge to determine the contribution made by each idea unit to the overall theme, and hence whether it was an important idea unit or not.

Brown and Smiley (1977) established four levels of importance in two Japanese folktales through a norming study with college students. These subjects eliminated the quarter of idea units considered least important, then the next least

important quarter, and so on. The average of these judgements was used to determine whether a particular idea unit should be assigned to level 1 (least important), 2, 3 or 4 (most important).

In their study of developmental differences, Brown and Smiley chose college students and children in grades 3, 5 and 7 to listen to each of the two folktales. All subjects were readers of at least average ability. For one story free recall protocols were collected, while for the other the least important ideas were progressively eliminated, as in the norming study.

While grade 3 students recalled less than all other groups, the recall of all age groups was affected by importance level. In other words, all groups recalled more level 4 than level 3 units, more from level 3 than from level 2, and more from level 2 than level 1.

However, age differences were more marked in the ability to rate importance level. The grade 3 students did not distinguish consistently between levels of importance, while the grade 5 students were able to determine level 4 units only. However, by grade 7 ratings approached the adult pattern. Grade 3 students rated in an idiosyncratic way that was not correlated with adult ratings. The ratings of grade 5 students showed a modest correlation ($r = .33$) with those of adults for both stories, while by grade 7 these correlations were .87 and .76 respectively for the two stories. The ratings of the college students in this study were almost identical to those of the norming group.

With regard to the levels effect apparent in the free recall of all age groups, the effect may have resulted from the integrative processes carried out during reading. In this situation integration was made easier by removing the need for decoding and by using stories for which background knowledge was readily accessible, even

for grade 3 children. It is quite possible, then, that even the younger children were able to coherently integrate all the ideas. If important ideas (level 4) participate in repeated integrative episodes because they are related to several other ideas, and unimportant ideas (level 1) do not, as suggested by Meyer, Brandt and Bluth (1980), then high-level idea units should also be easier to recall, producing the observed levels effect in recall. It would appear, though, that the ability to deliberately examine the products of this processing in order to make explicit judgements of importance is rather late developing.

If integration is not carried out efficiently, then there is unlikely to be a clear levels effect in recall, as may have happened in the study by Smiley, Oakley, Worthen, Campione and Brown (1977). These authors compared the sensitivity of above and below average grade 7 students to levels of importance. One group of students was reading at or above grade level, while the other was at least two years below grade level. These students heard one of the folktales and read the other, each followed by immediate recall. In a subsequent experiment, children in grade 1 heard and recalled one of the stories.

Overall, good readers recalled more idea units than poor readers. In addition, good readers recalled more units from level 4 than level 3, and more from level 3 than levels 2 and 1. However, poor readers distinguished only between level 4 and the rest. For poor readers, after both reading and hearing the stories, their overall amount of recall and its distribution across levels of importance was similar to that shown by grade 1 students.

Brown and Smiley (1978) examined the effect of additional study time on recall of ideas from different levels of importance. In two separate experiments, these authors tested students from college level and grades 5, 7/8 and 11/12.

Subjects recalled a folktale after hearing it once. They then had 5 minutes to study the written text, during which time they could use any strategies they wished to help them remember the text. A second free recall was then taken. For the second story, subjects were given the mild prompt that underlining or note taking helped some students to learn.

As in the previous study, idea units were more likely to appear in the adults' protocols the higher their level of importance. While recall increased after the study period, the increase was limited to level 3 and 4 units. Thus, adults focussed their attention on the most important ideas during their study.

For grade 5 students, on the other hand, there was no difference between immediate and delayed recall. Grade 7/8 showed some improvement in recall of level 3 and 4 units, while this improvement was as marked in the grade 11/12 students as in the adults. As in the previous study by Brown and Smiley, it was not until grade 7 that students as a group began to show an adult-like ability to deliberately focus on important ideas.

However, there were within-group differences. Among the students who spontaneously underlined or took notes during the study period, even students in grade 5, recall of important idea units improved after study. The spontaneous use of these strategies increased with age. On the other hand, prompting students to use a strategy when they did not do so spontaneously was not effective, and non-strategic subjects of all ages did not improve their recall of important ideas as a result of study.

These results suggest that the ability to study strategically depends on being able to determine the most important ideas within a text. Some students in grade 5 were beginning to develop this ability, but in general it was not until grade 7

that students began to study text in a manner similar to adults. The results also suggest that merely knowing about a strategy and its usefulness is not sufficient to improve performance; the reader must also be able to carry out the processing necessary to apply the strategy, in this case to be able to determine the relative importance of ideas.

Grabe (1980) investigated the ability to determine the importance of ideas in a different way, namely by being able to apply a particular perspective to a text. Grabe used the story by Pichert and Anderson (1977), which described boys truanting from school, but also contained information which could be relevant to a burglar wishing to rob one of the boys' homes. Thus, which ideas were of greatest importance would depend on the reader's goal, whether it was to follow the narrative as written or to glean information relevant to a robbery.

The subjects were good and poor readers in grades 4 and 6, selected according to the vocabulary subtest of the Iowa Silent Reading Test. They were told to remember as much as possible from the text, but the perspective group was told in addition to focus particularly on information that might be relevant to a burglar. Students were told to highlight information they thought was important as they were reading.

Older children recalled more than younger children, and good readers recalled more than poor readers. The latter difference, however, was limited to perspective-relevant information, not irrelevant ideas. In other words, good readers were more able to apply the burglar perspective to their reading, to determine the importance of ideas accordingly, and to retain important ideas. In terms of highlighting, good readers highlighted more ideas than poor readers, with this difference again limited to perspective-relevant ideas. Good readers recalled a

significantly higher percentage of highlighted ideas than poor readers. It appears, then, that good readers were not only more able to determine important information, but were also more likely to devote additional processing to it such that it was likely to be retained and produced in recall.

This study is an interesting one because it involves the ability to determine importance without the benefit of an overall signalling structure. Since the burglar perspective was not the central theme of the passage, importance was unlikely to be determined through the use of textual cues and knowledge of text structures. Rather, readers had to keep in mind the burglar theme and judge the importance of idea units to this theme on the basis of their background knowledge. In this study, good readers were more able to carry out this processing than poor readers.

Being able to determine important ideas is an essential skill for the construction of text summaries. Several studies have investigated the ability of good and poor readers to condense important information into summary form.

Brown, Day and Jones (1983) examined the ability of subjects to summarise the two folktales used in the Brown and Smiley (1977; 1978) studies after comprehension had been achieved and the stories had been learned to criterion. Subjects from college and grades 5, 7 and 11 wrote a free summary and then a 40- and a 20-word summary. The grade 7 students performed much like the older subjects in including mainly level 3 and 4 idea units in their summaries. The grade 5 students showed some sensitivity to levels of importance, although not to the same extent as the older subjects. These youngest children did include more level 4 units than others in their free summaries, and did drop level 1 units to a greater extent than others in their restricted summaries, but that was the limit of their discrimination.

Brown and Day (1983) investigated the rules applied to summarising by different age groups after reading expository texts. Students in grades 5, 7 and 10 as well as college freshman were all capable of deleting trivial and redundant information. However, grade 5 and 7 students selected explicit topic sentences for inclusion in their summaries on only about one-third of the occasions when it was appropriate to do so, compared to two-thirds of occasions for college freshmen. The invention of topic sentences when they were not explicitly stated was the most difficult rule to apply for all age groups; grade 5 and 7 children used this rule on only 14% and 28% of appropriate occasions respectively, compared to 52% for college freshmen and 84% for rhetoric experts. Thus, while the grade 5 and 7 children were quite capable of identifying and deleting trivial and redundant details, they performed quite poorly when identifying important statements and especially when attempting to generate their own summary statements.

K. Taylor (1986) examined summarising skills among children in grades 4 and 5 who had been identified by teachers as average or above average readers. The children were instructed to summarise one expository and one narrative passage in no more than six sentences. A panel of four judges scored each summary for accuracy/clarity, inclusion of main ideas, brevity and use of own words. For each criterion a maximum of 3 points and a minimum of 1 point was awarded by each judge, so a child could score between 4 and 12 for each criterion.

Contrary to expectation, scores for the narrative passage were not significantly different to those for the expository passage. Scores were high for brevity and use of own words. Most children also performed reasonably well on the accuracy/clarity criterion. The weakest area was the inclusion of important ideas; about one third of the children scored less than 8, 8 being the score which would

reflect a judgement of moderate achievement from each of the four judges.

On the basis of think-aloud protocols collected while children were summarising, K. Taylor (1986, p. 198-9) puts forward one reason why some children did not include important information in their summaries:

> *"... rather than selecting the most important ideas from the text for a general audience, several of the children had the confused notion that the most unusual or unfamiliar ideas should be chosen. For instance they stressed that a summary should contain what the audience would like to know and not just the important ideas in the article."*

Another characteristic of poor summarisers was that they were too specific and tended to concentrate on concrete details, which perhaps was also related to their focus on interesting information.

Winograd (1984) conducted a more detailed study of summarising among good and poor readers in grade 8 and adults at university level. Reading ability was assessed by means of the comprehension subtest of the Standford Achievement Test (SAT). In addition, decoding speed and accuracy were measured for each student, and non-verbal I.Q. was assessed by means of the IPAT Culture Fair I.Q. test. The children were also asked a number of questions to assess awareness about the task of summarising, including what it means to summarise.

All subjects read six expository passages. After each one they answered from memory a series of multiple-choice comprehension questions based on information explicitly stated in the text. They rated the importance of each sentence on a scale of 1 to 5 and constructed a 60-word summary with the text present.

Contrary to expectation, task awareness did not distinguish good and poor readers; over two-thirds of both groups stated that a summary should include

the most important ideas. However, good and poor readers did differ in terms of what they considered to be the most important ideas. The importance ratings of sentences given by good readers were considerably more like those of adults ($r = .71$) than were those given by poor readers ($r = .46$).

However, poor readers were not idiosyncratic in their judgements, but appeared to have different criteria for determining importance. According to Winograd, sentences considered to be important by poor readers, but not others, tended to be rich in visual detail and perhaps appealed to personal interest. This tendency was also noted by K. Taylor (1986) among less skilled summarisers in grades 4 and 5.

Among poor readers there was a weak relationship between information they rated as important and information they included in summaries. On the other hand, this relationship was moderately strong for good readers and stronger still for adults.

Winograd examined the types of transformations applied to textual information to reduce it to summary form, with the transformations including combinations amalgamating the main elements of several sentences, inventions representing inferred topic sentences to summarise the gist of paragraphs or larger segments of text, run-on combinations incorporating several sentences in great detail, as well as direct reproductions from the text. Good readers used more combinations and inventions, and fewer run-on combinations and reproductions than poor readers, but only the difference in combinations reached significance. Combinations were the type of transformation used most often by adults.

Finally, Winograd examined the contribution of various skills to comprehension and summarisation performance. Summarisation scores were calculated for

children according to the extent to which they included and excluded information in the same way as adults. Factors which explained a significant percentage of variance in these scores included non-verbal I.Q., sensitivity to the importance of ideas, and effective rule use, that is, the percentage of statements in summaries which were combinations or inventions.

Significant predictors of passage comprehension were non-verbal I.Q., decoding accuracy, sensitivity to importance and, marginally, effective rule use. Significant predictors of SAT comprehension scores were non-verbal I.Q., decoding speed and sensitivity to importance. As Winograd (1984, p. 420) notes, the ability to determine the importance of ideas emerged as an important predictor of both summarising and comprehension performance.

> *"...the most striking result is that the ability to identify important elements in a text accounted for a significant proportion of the variance in all three dependent measures, even after I.Q. and decoding ability had been taken into account. There is strong evidence that higher-order comprehension difficulties may be linked to strategic deficits. This also indicates that the ability to identify important elements in a text is a strategic skill that underlies both comprehension and summarization."*

Although Winograd draws attention to this similarity between comprehension and summarisation, namely, that sensitivity to structural importance contributes to the performance of both, he also points out that summarisation requires specialised skills in addition to those used for comprehension. The same conclusion can also be drawn from the studies by K. Taylor (1986) and Brown, Day and

Jones (1983). In the latter study, age differences in summarisation skill appeared even after comprehension to criterion had been achieved by all students. Thus, if the aim of an instructional programme is to improve children's comprehension, rather than their summarisation skill, these results suggest that students should be taught to evaluate the importance of ideas to the overall theme, since this skill contributes to both comprehension and summarisation, rather than summarisation techniques, which have a more specialised function.

With regard to evaluating structural importance, Winograd (1984, p. 422) emphasises the importance of being able to use cues provided by the text to assist in determining importance and interrelationships.

> *"...poor readers seem to have difficulty in using textual cues. Without such guidance it must be very difficult to construct an accurate, organized representation of what the author intended to communicate."*

The extent to which readers use textual cues and their knowledge of text structures to organise their mental representation of a text has been investigated in several studies. As already mentioned at the beginning of this section, Meyer, Brandt and Bluth (1980) hypothesise that readers use their knowledge of text structure to encode and retrieve information from a text in a systematic way, increasing their recall as a result. Poor readers, on the other hand, are hypothesised to have no organised way of using a text's top-level organisation.

In the Meyer et al study, students in grade 9 were classified as good, average or poor readers according to the comprehension subtest of the Stanford Achievement Test. They read two passages, one with a compare-contrast top-level structure,

the other with a problem-solution structure. They recalled the passages both immediately, and again after one week. Recall was scored both for total number of idea units and for the importance of idea units, that is, whether they were part of the main message, major supporting details or minor supporting details.

In recall of both passages there was a significant main effect of ability on both total recall and recall of the main message, with good readers recalling more than average readers, who in turn recalled more than poor readers. Protocols were also categorised according to whether or not they followed the texts' top-level structure, in one passage, for example, organising ideas under the headings of 'problem' and 'solution'. Readers who used the texts' top-level structure recalled more major and minor details at both times of testing, and especially more main message units. Most of the good readers used what Meyer et al refer to as the 'structure strategy', that is, looking for conventional patterns of organisation in the text and using them to encode and retrieve information. Most of the poor readers, on the other hand, used the 'default strategy', just trying to remember something. As Meyer et al (1980, p. 99) state:

> *"As suggested by the default strategy of the model and supported by the data, the ninth grade students who did not employ the structure strategy simply tried to list collections of descriptions from the passage without interrelating them. In contrast, those employing the structure strategy compared viewpoints or related solutions to components of the problem, and consequently, developed a rich retrieval network."*

McGee (1982) obtained similar results with younger children. Good readers in grades 3 and 5, and poor readers in grade 5 with the same range of scores on

the SRA test as the children in grade 3, read two passages with a description structure. Recall protocols were scored for total number of idea units and the importance of idea units, that is, whether they were superordinate (levels 1 and 2 in the Meyer system) or subordinate (level 3).

Overall, good readers in grade 5 recalled more than poor readers, who in turn recalled more than the good readers in grade 3. However, on closer analysis this difference was significant only for superordinate, not subordinate idea units.

Recall protocols were also examined for the extent to which they followed the texts' organisation of ideas. The majority of grade 3 students showed no structure and produced a list-like collection of ideas, the majority of poor readers in grade 5 showed a partial structure in recall, while most of the grade 5 good readers showed a full structure, that is, following the same organisation as the texts. However, this group of readers still recalled less than half the level 1 and 2 idea units, and less than 40% of the information overall, even though the texts were of grade 3 readability level.

Taylor (1980) investigated the use of text structure amongst good and poor readers in grade 6 and good readers in grade 4, with skill determined by the Iowa Test of Basic Skills comprehension subtest. One text with a description top-level structure was written at grade 6 readability level for the grade 6 good readers and at grade 4 level for the younger and poorer readers. The text was recalled both immediately and after a two day delay.

In immediate recall there were no significant group differences in the number of students using the same top-level structure as the text in their protocols. However, both good and poor readers in grade 6 recalled more than grade 4 students. In delayed recall the grade 6 good readers were significantly more likely

than the other two groups to use the top-level structure of the text in their protocols. These good readers also recalled significantly more idea units than the grade 6 poor readers and the grade 4 good readers, who did not differ significantly from each other.

The amount recalled was not significantly affected by the use of text structure in immediate recall. However, in delayed recall both good *and* poor readers who used the top-level structure recalled more than good and poor readers who did not use this structure.

These results are somewhat ambiguous as to the importance of using the text's top-level structure to encode information, since good readers were not more likely to use this structure than poor readers in immediate recall, and use of this structure was not related to amount recalled. Only in delayed recall did the predicted group differences emerge. Here use of top-level structure was related to amount recalled for good and poor readers, so for the latter lack of awareness and use of text structure would appear not to have been an important factor in their lower reading ability. Taylor (1980) does note that only one text was used, so results may have been influenced by passage-specific characteristics, for example, the familiarity of the content to different children.

Ohlhausen and Roller (1988) examined how readers decide which information in a text is important, that is, the respective roles of content and structure schemata. Adults and children in grades 5, 7 and 9, with no information given as to reading ability, read a passage in one of three versions. The normal text had a description top-level structure which was made more salient by tight signalling. In an other version, in order to prevent readers from using content schemata to determine important ideas, significant content words which could trigger the

schemata were changed into nonsense words (content deformed version). The structure remained unchanged. In the third version, in order to prevent readers using structure schemata, all signalling was removed and sentences were presented in random order, but content was unchanged (structure deformed version).

After reading one version of the text, readers were asked to underline the seven most important ideas and to explain the basis on which they had selected these ideas. Explanations were scored for degree of elaboration and categorised according to whether they were based on content or structural features.

Adults performed better on the underlining task than children in grades 9 and 7, who in turn performed better than children in grade 5. Not surprisingly, the ability to underline the most important ideas was greater with a normal text than with the two deformed versions for all readers except those in grade 5. In terms of scores for structural explanations, adults scores significantly higher than all children, and grade 9 students scored higher than those in grade 5. Scores were higher for the normal and content deformed texts than for the structure deformed text. With regard to scores for content explanations, adults scored higher than children in grades 7 and 5, and scores were higher for the normal and structure deformed texts than for the content deformed text.

In short, there was a developmental trend in the ability to underline important ideas and to explain the basis of this underlining. As expected, readers used content and structure schemata as appropriate to determine important ideas. However, Ohlhausen and Roller note that when readers have the opportunity to use content schemata they do so, and make less use of structure schemata.

In the first chapter, it was pointed out that knowing 'what a text is about', that is, being able to condense it into its most important ideas, is an important

comprehension skill. This section has shown that more mature and more skilled readers perform better at tasks which require an understanding of main ideas than younger and less skilled readers.

The comparison of various studies can be difficult because experimental materials cover different topics, subjects are required to perform a variety of tasks, and texts could be of various structures and levels of complexity, all of which could influence how well the subjects perform. However, keeping in mind this difficulty, certain trends have been noted from the data.

Firstly, several studies indicate that adults and older students often do not perform perfectly, that is, even relatively mature readers still have room for improvement. For example, in Williams et al (1981), when adults selected and produced titles for simple paragraphs, they scored 87% and 80% for the two tasks respectively. In Hiebert et al (1983), good and poor readers at college level scored 81% and 71% respectively for rejecting sentences not consistent with the topic and structure of paragraphs. In Brown and Day (1983), college freshmen produced only about half the possible number of invented topic sentences in their summaries, and year 11 students in the study by Hare et al (1989) still had difficulty generating their own topic sentences for paragraphs.

Secondly, from a number of studies it appears that children by the end of the primary years, which in South Australia is grade 7, are approaching an adult level of performance for some, but not all tasks involving main ideas. This trend is particularly evident in the studies by Brown and Smiley (1977; 1978), where grade 7 children were beginning to rate the importance of idea units in a manner similar to adults and, like adults, were able to devote additional study time to improving recall of important ideas. In Brown, Day and Jones (1983), grade 7

students included important idea units in their summaries similar to the adults. In all these studies grade 5 students, by contrast, showed only a rudimentary understanding of the importance of different ideas.

In Garner et al (1986), children in grade 7 when constructing paragraphs achieved almost perfect scores for excluding sentences not related to the specific topic, compared to average scores of only about 50% for grade 5. When constructing paragraphs in Garner and Gillingham (1987), children in grade 7 were more able to place the topic sentence in the initial position (73%) than children in grade 5 (40%). In Danner (1976), more children in grade 6 (83%) were able to place sentences into logical paragraphs with topic sentences than children in grade 4 (67%). Children in grade 6 in Williams et al (1981) were approaching the adult level for correctly selecting a specific title for paragraphs (78%), compared to children in grade 4 (58%). Grade 7 children in their verbal reports were more able to identify topical coherence as the basis of good paragraphs than children in grade 5 (Garner et al, 1986; Garner and Gillingham, 1987).

Finally, there is evidence that skilled readers are more able than less skilled readers to identify important ideas and generate summary statements. Among adults, good readers were more able to generate titles and summary statements in Bridge et al (1984), and more able to reject deviant sentences from paragraphs in Hiebert et al (1983). Among children, good readers were more able to reject deviant sentences (Englert and Hiebert, 1984), and more able to identify important information and include it in a summary (Winograd, 1984). Good readers were also more able to organise their recall of information according to the same logical top-level structure as the text (Meyer et al, 1980; McGee, 1982).

In conclusion, the various aspects of macrostructure formation would appear

to be useful skills to include in a training programme for children in the upper primary grades in general, and for poor readers in particular. From the evidence presented, all children could potentially gain from such instruction, since even good readers are not performing at ceiling level. Children may benefit from being taught not only to make use of explicit structural cues, such as headings, topic sentences and various other methods of signalling importance and structural relationships, but also to invent topic sentences or summary statements when these are not provided by the text.

2.5 Metacognition

Paris, Wasik and van der Westhuizen (1987) in their review draw attention to the large number of articles published in recent years on the subject of metacognition and its role in skilled reading. Paris and Jacobs (1984, p. 2083) describe the importance of metacognitive knowledge and skill applied to reading, what they call 'reading awareness', as follows:

> *"The development of reading awareness is an important cognitive attainment because it distinguishes beginning and advanced readers. Skilled readers often engage in deliberate activities that require planful thinking, flexible strategies, and periodic self-monitoring. They think about the topic, look forward and backward in the passage, and check their own understanding as they read. Beginning readers or poor readers do not recruit and use these skills. Indeed, novice readers often seem oblivious to the strategies and the need to use them."*

In the first chapter, several theoretical views of metacognition were presented. For example, Paris and Jacobs (1987) divide the construct into self-management and self-regulation aspects, the former including declarative, procedural and conditional knowledge, the latter involving the planning, evaluation and regulation of cognitive activity. Borkowski and Kurtz (1987) divide the construct into four aspects, these being Specific Strategy Knowledge, Relational Strategy Knowledge, General Strategy Knowledge and Metacognitive Acquisition Procedures. From the use of specific strategies readers build up knowledge regarding their relative usefulness and difficulty, conditions of applicability, and the value of strategic behaviour in general. General Strategy Knowledge also involves a motivational

component, including attributional style, while the Metacognitive Acquisition Procedures fulfill an executive, regulatory function.

In the sections to follow, differences between more and less advanced readers in several of these areas will be examined. The first section will discuss differences in metacognitive knowledge, including awareness of appropriate goals and strategies. The next section will examine some of the motivational differences included by Borkowski and Kurtz in their General Strategy Knowledge category. The final section will focus on differences in comprehension monitoring.

Before proceeding, two points made in the first chapter will be briefly reiterated. Firstly, when analysing a performance, the point made by Paris and Jacobs (1987) should be kept in mind; metacognition should not be accorded an 'undeserved ubiquity' in explanations of cognitive processing. It was pointed out in chapter one that much information relevant to cognitive monitoring and regulation is a by-product of processes such as integration, and does not always depend on metacognitive self-questioning. Thus, if a reader fails to detect an inconsistency, for example, the problem may be due to inadequate integration in a very resource-demanding text or, a related problem, inadequate background knowledge of a topic of discourse. In both cases the reader may be unaware of the inconsistency, but the basis of the problem may be largely cognitive rather than metacognitive.

The second point is an extension of the previous one, namely, that each method of assessing metacognition has inherent problems. In the section to follow a variety of approaches will be examined in the hope of finding converging evidence. If evidence obtained by different approaches tends to suggest the same conclusion, then the conclusion can be accepted with greater confidence than if

the evidence is obtained through a narrow range of approaches.

2.5.1 Metacognitive Knowledge and its Application

In chapter one, Flavell's (1981) conception of metacognitive knowledge was described as including knowledge about person, task and strategy variables. Myers and Paris (1978, p. 680) point out the importance of such knowledge to performance as follows:

> *"First, children need to know about their own enduring characteristics and transient conditions that influence performance. Learners or memorisers need to appraise realistically in order to engage in skills commensurate with their ability. Second, children need to know about the purposes, scope, and requirements of the task before the problem can be efficiently attacked. Third, one needs to be aware of the existence of relevant strategies and to recognize the need to apply them. Further, one must form plans, generate hypotheses, check own's progress, evaluate results, and generalize behaviour."*

This section will examine differences in metacognitive knowledge between more and less advanced readers, both in terms of age and reading skill. Metacognitive knowledge in the studies to be discussed was assessed using a variety of approaches, including interviews, think-aloud procedures, and analysis of observable strategies used during performance.

Myers and Paris (1978) compared the metacognitive knowledge about reading possessed by more and less experienced readers, in this case children in grades 2 and 6. More children in grade 6 were able to state the basis on which paragraphs

were organised, and the function of the first and last sentences in paragraphs. The older children reported focusing on meaning when retelling a story while about half the younger ones reported that verbatim reproduction was more important. More of the older children realised that gist recall was easier than verbatim reproduction, and the older children were much more likely to report using different strategies to prepare for gist as opposed to verbatim recall. In a situation where only some of the words in a story could be read, the most popular reported strategy amongst the younger children was to focus on easy words, while older children were more likely to report focusing on informative words. The older children were more likely to report useful strategies for dealing with problems, such as looking up unknown words in a dictionary, asking someone if a whole sentence was not understood, or going back to the beginning of a paragraph or story. On the other hand, more grade 2 children reported sounding out unknown words and not knowing what to do if a whole sentence was not understood, and they were less likely to be able to explain the reason for rereading sections of a story.

In general, then, the more experienced readers showed awareness of a wider range of strategies and were more able to match strategies with task goals. They showed more emphasis on meaning construction, as opposed to the greater orthographic and word level emphasis of the younger students.

Gambrell and Heathington (1981) used the same set of questions as Myers and Paris (1978) to assess metacognitive knowledge. Learning disabled adults reading at or below grade 5 level according to an informal reading inventory were compared to adults drawn from a business administration class, who were presumably good readers. These good readers were more able to state the basis

on which sentences were organised into paragraphs. Good readers to a far greater extent reported that reading for gist recall was easier than reading for verbatim recall. Poor readers, on the other hand, were more likely to report that oral reading was faster than silent reading. When asked what makes a good reader, the good readers were more likely to state that it was someone who could grasp the meaning, while poor readers were more likely to think it was someone who could read quickly. When confronted with an unknown word, all good readers reported using an independent strategy, such as sounding it out or using context to deduce meaning, whereas one third of the poor readers did not report any independent strategies and either did not know what to do or would ask someone. Gambrell and Heathington conclude that the responses of the adult disabled readers resembled more closely those of grade 2 children than those of grade 6 children in the Myers and Paris (1978) study. These disabled readers showed less knowledge about appropriate strategies and the effect on reading of various task demands, and placed less emphasis on construction of meaning as the goal of reading.

Garner and Kraus (1981/2) conducted an interview with grade 7 good and poor comprehenders, who were selected on the basis of teacher judgements. The children were asked what was involved in being a good reader, how to judge efficiency of reading, and what made something difficult to read. The responses of poor comprehenders focused almost exclusively on words, that is, knowing and being able to fluently pronounce the words in a text. On the other hand, good comprehenders focused on larger units of meaning, noting, for example, that good readers were the ones who understood the main ideas, and that being able to get the main ideas was a good test of reading well. Good comprehenders suggested

using context to work out something they didn't understand, and they were more likely to suggest rereading a difficult part, whereas poor comprehenders were more likely to suggest skipping the difficult section.

In contrast to the above studies using general interviews about reading, Olshavsky (1976/7) used think-aloud protocols to infer strategies used by students in grade 10. Subjects were divided into good and poor readers according to the comprehension section of the Iowa Silent Reading Test. In the experimental text, a dot appeared after each clause in the short story to remind readers to stop and report what they were thinking and doing at that point. Olshavsky grouped responses into nine types of strategies.

Good readers did not report significantly more types of strategies than poor readers, although there was a tendency for good readers to report using strategies more frequently overall. In spite of the lack of significant difference in reported strategy usage, good readers gave a more accurate paraphrase and were more able to retell the story. If knowledge and application of appropriate strategies is as important to skilled reading as has been claimed, then this result is surprising. However, there are problems with think-aloud protocols to assess strategy usage, as indicated in the first chapter. The type of task used by Olshavsky, namely forcing readers to stop and report thoughts at specified intervals, could be potentially disruptive to normal reading, and could encourage the fabrication of reports because subjects feel compelled to report something. There is also the question of how effectively the strategies reported were actually applied.

Kavale and Schreiner (1979) also collected think-aloud protocols, but while subjects were answering questions rather than while reading. Children in grade 6 were classified as average or above average readers according to the compre-

hension subtest of the Gates-MacGinitie Reading Tests. They read a series of short paragraphs, each one followed by a multiple choice question, and thought aloud as they considered reasons for accepting or rejecting options. The specific reasoning strategies identified were classified as either textual or non-textual depending on whether or not they referred to information which had been read. A textual response could be either discourse or choice dependent, the former referring to information in the paragraph, the latter to information in the question alternatives.

Most responses were textual, with above average readers producing more of these responses than average readers. About half the responses were discourse dependent, with again the above average readers supplying more of these responses. The most commonly used discourse dependent strategies were Generalisation, that is, a synthesis of pieces of information from the paragraph, and Key Lexical Item, where a reasoning strategy revolved around a particular word deemed to be important. Above average readers used these two strategies more frequently than poor readers.

Although there were group differences in the overall frequency of strategy usage, average and above average students used similar types of strategies. However, above average students used strategies more successfully. They were able to achieve an 85% success rate in answering questions with each of six strategies, while poor readers successfully used only two strategies. In other words, above average readers could effectively use six strategies compared to only two among poor readers. In contrast, above average readers used three strategies with limited success, compared to six strategies used unsuccessfully by average readers. In summary, above average and average readers did not use different strategies,

but the former used strategies more frequently and were able to use a wider range of strategies successfully. Overall, above average readers also answered more questions correctly than average readers.

Kaufman, Randlett and Price (1985) included in their study both retrospective reports on an actual task and reporting of reading behaviour in general. Two groups of college students were selected according to scores on the Nelson-Denny Reading Test; the top and bottom quarters were chosen to represent good and poor readers. Subjects read several short passages and answered questions after each one. They were then interviewed about how they thought they had performed, whether there were any confusions, how they responded to confusions, and how they read in general. Finally, subjects responded to a forced-choice questionnaire in which a variety of possible reading strategies were listed.

In response to the question about how they lessened confusion, good readers gave more responses than poor readers. In response to general questions about reading behaviour, good readers reported more "inside the head" (category 3) strategies, such as making connections between pieces of information or using context to deduce meanings. There was no difference between the groups in the reported use of observable strategies such as rereading (category 2), or environmental strategies such as finding a quiet place to read (category 1). Specific items that good readers were more likely to see as applicable to themselves than poor readers were: believing in personal control over performance on a reading test, looking for main ideas before and during reading, and finding it easy to separate important from unimportant information. Kaufman et al speculate that poor readers are less able to apply higher-order thinking skills because of less efficient decoding skills, which leave less working memory capacity for construction of

meaning. The authors point out that the teaching of category 2 strategies will be of little value unless readers are able to engage in these higher-order skills:

> *"It would be possible for a reader to engage in all the category 2 strategies and still not understand the meaning of what is being read. Unless readers actively, mentally manipulate the words and phrases, unless readers make mental connections and inferences, they may understand little of the meaning in the text. Reading the words in the subtitles, skimming, pointing, rereading, etc does not guarantee comprehension. To understand what is being read, the reader must become actively involved in the reading task."* (p. 9)

Paris and Myers (1981) classified children in grade 4 as good or poor readers according to the California Achievement Test. They read a story at grade 4 level containing four unknown words. The children were told they could do anything they thought would help them to understand the text, and any strategies used were recorded. In addition, the children rated the degree to which various factors helped or hindered reading, and recalled the text both immediately and after one week.

Good and poor readers did not differ in the amount of time taken to study the text, although significantly more good readers demonstrated strategies for determining unknown word meanings. More good than poor readers made notes, used a dictionary, asked the experimenter for meanings and learned the definition of at least one word in the course of the experiment. Good readers recalled marginally more than poor readers on the immediate recall test, but significantly more on the delayed recall test due to greater forgetting by poor readers, who

also produced less organised recall protocols. Good and poor readers were equally aware of the facilitative effects of positive strategies, but poor readers were less aware of the negative effects certain actions could have, for example, watching television while reading, or skipping difficult parts. In summary, strategies for determining meaning and enhancing recall were used more frequently by good readers and the use of these strategies was associated with higher recall.

Garner, Wagoner and Smith (1983) attempted to externalise the strategies used by good and poor readers in grade 6 by setting them the task of helping average readers in grade 4 answer questions about an unfamiliar passage. Three of the questions required information from the text for the formulation of an answer, while two began 'Do you think ...' and asked for the reader's opinion. Tutors were instructed to give advice to their tutees on what to do to answer the questions.

Good readers to a greater extent than poor readers encouraged their tutees to look back to the text. Good readers also made a greater distinction between types of questions, referring tutees to the text for text-based but not on reader-based questions. Good readers were more directive in their text sampling, referring tutees to specifically relevant parts of the text. Thus, good readers not only used lookbacks more frequently, but were more sophisticated in their use, knowing, when, where and why lookbacks could be productively used.

In a study by Scardamalia and Bereiter (1984), students in grade 6 and grade 10 attempted to reorder two sets of scrambled sentences to form coherent texts. Older students were more likely than younger ones to reorder a sentence after initial placement and to group sentences into clusters prior to placement. The authors considered three types of reading strategy associated with the task, namely,

survey content by reading ahead of the sentence being placed, reread and check by rereading more than the sentence being placed and indicating that the placement of the sentence was potentially provisional, and summarise content by stating the gist of information contained in more than two sentences. Older and younger children did not differ in their use of the 'survey content' strategy, although grade 10 students used 'reread and check' and 'summarise content' more frequently than grade 6 students. Scardamalia and Bereiter (1984, p. 392) conclude:

> *"Findings considered up to this point indicate a developmental trend toward handling sentence arrangement problems by means of tentative groupings and placement of sentences guided by rereading and checking and by summarizing portions of text already assembled. The approach taken by the more mature students appears, therefore, to parallel the process of gist-construction posited by the Kintsch and van Dijk (1978) model. This parallel is reinforced by a significant correlation between use of summarization in sentence arrangement and the use of interpretations oriented toward macro-structure in ordinary reading. Younger students, on the other hand, tend to approach sentence arrangement in the same element-by-element fashion that has been hypothesized they use in normal reading, making seldom reconsidered judgements about the placement of sentences, much as they appear to make once-and-for-all judgements about the significance and meaning of statements as they encounter them serially in reading."*

When asked to justify the placement of sentences, grade 6 students either referred to argument overlap between adjacent sentences or some reason which

was not explanatory, such as liking the general effect or sentences sounding good together. On the other hand, grade 10 students were more likely to refer to the structural role of the sentence in that position or to construct a superordinate proposition for the sentences under consideration to show their coherence, leading Scardamalia and Bereiter (1984, p. 397) to state:

> *"What the sentence arrangement task seems to do in a case such as this is to produce an overt display of the macro-operations that are presumed to go on covertly in normal text comprehension (cf Kintsch and van Dijk, 1978). For sophisticated students, then, sentence arrangement appears to consist of constructing a mental representation of text meaning and is accomplished by the processes normally employed in comprehension."*

Wong and Wong (1986) examined knowledge about the effect of study time and passage characteristics on recall, as well as children's actual study behaviour and recall. Subjects were above average and average readers in grades 5 and 6, and below average readers in grades 5, 6 and 7, with reading ability defined according to the comprehension section of the Gates-MacGinitie Reading Test. Two pairs of passages were used for the experiment. In the first pair, one passage contained more difficult vocabulary than the other, and in the second pair one passage was thematically organised while the other was scrambled.

In an interview, children were presented with the hypothetical situation of students studying these passages for varying amounts of time, and were asked which students would remember more and why. Children were scored for both correctness and justification of response. Three weeks later the children actually studied the passages for a free recall and comprehension test.

Above average readers achieved higher scores for the interview about study behaviour than both average and below average readers, who did not differ significantly from each other. On both the free recall and comprehension tests, above average readers performed significantly better than below average readers. Average readers were intermediate between the two but did not differ significantly from either. More comprehension questions were answered correctly following the passage with easy vocabulary than the one with difficult vocabulary, and more information was recalled following the well-organised passage than the poorly organised one.

With regard to study time, more time was spent on the disorganised than the organised passage, and more time was spent on the passage with difficult vocabulary compared to the one with easy vocabulary. However, there was an interaction between passage type and reader ability. Above average readers were most sensitive to the organisation variable and showed a significantly greater increase in study time for the disorganised passage than the other two groups, who were not very sensitive to this variable. In contrast, below average readers were particularly sensitive to vocabulary difficulty, showing a larger increase in study time for the passage with difficult vocabulary than the other two groups, with above average readers showing little increase in study time due to this variable. Thus, although above average readers showed awareness of the effect of difficult vocabulary in the interview, they did not adjust their study behaviour to take account of this factor, and although below average readers showed little awareness in the interview, they did adjust their study behaviour. As Wong and Wong explain, metacognitive knowledge and strategy usage are not always directly related, and *"...the lack of clarity in the data reflect basically substantive problems*

in metacognitive assessment." (p. 109)

This result highlights one of the many problems in studying differences in strategy usage between good and poor readers, namely, that when the two groups are given the same texts to read, the task is considerably more difficult for the poor readers than for the good readers, and this in itself may affect strategy usage. It could be that the difficult vocabulary in the Wong and Wong study was not in fact difficult for the above average readers, so examination of study time says more about ease of lexical access than application of metacognitive knowledge. This point is an important one with regard to the choice of content in an instructional programme. If poor readers have adequate strategy knowledge, but are not able to apply it to grade-level texts because of the cognitive demands made by these texts, then instruction needs to focus on improving processing, not increasing strategy knowledge.

The relationship between text difficulty and strategy usage was investigated by Kletzien (1991), who selected good and poor readers in grades 10 and 11 on the Basis of the California Test of Basic Skills. All students read three social studies passages, which were written at the seventh, eleventh and fourteenth grade level respectively for good readers. In order to ensure an equal level of difficulty for poor readers, Kletzien wrote the passages for them at a level two grades lower, although she does not indicate whether the poor readers were in fact reading two years below grade level. For both groups, then, one passage was easy (independent level), one was moderately difficult (instructional level), and one was very difficult (frustration level). Twelve context dependent words were deleted from each passage to produce a cloze test. Students filled in the blanks as they read and then explained how they had determined the chosen word. Kletzien

categorised these responses in terms of fifteen possible strategies.

With regard to frequency of strategy usage, there was a tendency for good readers to use specific strategies more often than poor readers at all levels of difficulty, but the difference was significant only at the frustration level. Frequency of specific strategy use declined for both groups with increasing text difficulty, but more so for poor readers. Regarding the type of strategies used, the two groups did not differ significantly at the independent level, although good readers used a wider variety of strategies at the more difficult instructional and frustration levels. Kletzien (1991, p. 79) describes the group differences as follows:

> *"At the independent level, the strategies used by the two groups were indistinguishable; however, differences became apparent at the instructional level. Good comprehenders appeared to have a greater ability to control their strategy use by changing types of strategies for the passage. ... Differences between the two groups became more pronounced on the frustration level passage. Good comprehenders continued to try different strategies to construct the meaning of the text, whereas poor comprehenders experienced a precipitous decline in strategy usage (both in variety of strategies and in number of times strategies were used). Because the passages were of the same relative difficulty for both groups, and because both groups of readers appeared to be familiar with the same strategies (judging by strategy use on other passages), one can conclude that the difference between the groups was in regulation, rather than knowledge, of comprehension strategies."*

Kletzien suggests that the difference may be a motivational one, in terms of effort attributions and sense of personal efficacy. In other words, poor readers may

be less likely to believe that increased effort will result in improved performance when the text is perceived as very difficult. However, the strategy use of all readers declined at the frustration level. At this level, some strategies were used in a narrow way, for example, rereading involved only a few words next to the blank. Of this observation Kletzien (1991, p. 82) says:

> *"Because vocabulary and sentence structure were much more complex at this level, subjects had to work much harder at the lower-level tasks of word recognition and understanding individual sentences, leaving less cognitive capacity available to understand the meaning of the passages as a whole and to integrate the content with what they already know."*

In many studies, poor readers are reading texts which are relatively more difficult for them than they are for good readers, so, based on Kletzien's results, it is not surprising if they report or demonstrate fewer strategies. The reason may well be processing difficulty leading to capacity limitations, not just inadequate metacognitive knowledge.

One puzzling aspect of the Kletzien study is that there was no significant difference in cloze performance between good and poor readers at any level of difficulty, in spite of the fact that good readers used a wider range of strategies more frequently at the higher levels of difficulty. There is no obvious explanation for this finding, which may again indicate that the relationship between performance and verbalised strategy use is not a simple one. It was pointed out in the section on macrostructure formation that Garner and Gillingham (1987) found a weak relationship between verbal reports of what makes a good paragraph and the performance of constructing a good paragraph.

However, Paris and Jacobs (1984) did find a relationship between reported knowledge and comprehension ability. Their metacognitive measure of reading awareness involved 15 open-ended questions tapping knowledge about evaluation, planning and regulation strategies in reading. They found a significant correlation among children in grades 3 and 5 between reading awareness and three different measures of comprehension, including a cloze test, error detection, and the comprehension subtest of the Gates-MacGinitie Reading Tests.

To conclude, several studies have shown that older and more skilled readers are more likely to report that good reading involves construction of meaning, in contrast to the focus on decoding and reading quickly reported by younger and less skilled readers (Myers and Paris, 1978; Gambrell and Heathington, 1981; Garner and Kraus, 1981-2). In these studies, the more advanced readers also reported more strategies for dealing with problems, such as rereading or using context to determine the meaning of an unfamiliar word, while less advanced readers reported few independent strategies as well as some damaging strategies, such as skipping difficult parts.

Good readers reported using strategies which involve more active constructive processing, including focus on important ideas (Kaufman, Randlett and Price, 1985; Scardamalia and Bereiter, 1984). Although good readers do not always report or demonstrate a larger number of different strategies, several studies show that they use available strategies more successfully to answer questions, recall a text, or tutor younger children (Kavale and Schreiner, 1979; Paris and Myers, 1981; Garner, Wagoner and Smith, 1983).

However, certain problems in the data have also been noted, in particular, that reported knowledge and strategy use was not always reflected in observed

behaviour and performance (Olshavsky, 1976-7; Wong and Wong, 1986; Kletzien, 1991). Such findings are not altogether surprising given the difficulties in measuring metacognitive knowledge pointed out in the first chapter.

Nevertheless, some general implications can be drawn from these studies. In some cases at least, less advanced readers have limited and inappropriate concepts of the purposes of reading. They may also know fewer strategies for achieving such goals and for overcoming difficulties. In other cases, though, the problem may not be lack of strategy knowledge, but difficulty in carrying out the processing necessary to apply the strategies and insufficient confidence to persevere when faced with a difficulty. Thus, while instruction in this area would appear to be important for less advanced readers, it needs to focus not only on increasing knowledge, but also on enabling students to master the application of known strategies.

2.5.2 Motivational influences on performance

In the first chapter it was pointed out that motivational factors are an important component of Borkowski and Kurtz's (1987) category of General Strategy Knowledge. The reader who does not attribute successful outcomes to effort or, conversely, unsuccessful outcomes to insufficient effort, is unlikely to see the point of strategic behaviour and persistence in the face of difficulty. Some of the studies referred to in the previous section indicate such lack of persistence among poor readers. For example, in Kletzien (1991) good and poor readers used a similar number of strategies with comparable frequency when reading a relatively easy text, but the strategy usage of poor readers declined much more than that of good readers when reading a difficult text. In Garner and Kraus (1981-2), poor

readers reported skipping difficult sections rather than adopting strategies to deal with the difficulty. Several researchers have investigated the relationship between motivational factors, strategy usage and reading performance in more detail.

Butkowsky and Willows (1980) studied three groups of boys in grade 5, with groups formed on the basis of discrepancy between their non-verbal I.Q. as measured by the IPAT Culture Fair I.Q. Test, and their reading ability as measured by the comprehension subtest of the Gates-MacGinitie Reading Tests. The three groups had a comparable average I.Q., but significantly different average reading scores, and so were labelled good, average and poor readers. The poor readers were those students reading at a level below that which their I.Q. suggested they were capable of, that is, they were underachievers.

All children completed two tasks, one involving anagrams described as a reading test, the other tracing line drawings without lifting the pencil. Each child experienced either success or failure on the two tasks, depending on whether or not they were given solvable problems. Prior to each task students were asked to estimate how many problems they would solve correctly. After each task students were again asked how many of a new set they would solve. After each task they were also asked to explain why thought they had or had not done well.

There was no significant difference between the groups in time taken to complete solvable problems. However, good and average readers persisted 40% longer than poor readers on unsolvable problems. In other words, the poor readers gave up much more easily in a difficult situation. With regard to attributions for failure, poor readers referred to internal and to stable factors significantly more often than average and good readers, in particular to lack of ability. Poor readers referred to external factors more often as the reason for success, that is, luck and

the ease of the task.

More specifically, poor readers attributed success to internal factors on only 59% of occasions, compared to 77% and 84% for average and good readers respectively. In other words, poor readers were less likely to claim credit for success. On the other hand, poor readers attributed failure to stable, and hence insurmountable factors on 93% of occasions, compared to 60% for both average and good readers. Good and average readers were more likely than poor readers to attribute success to ability and failure to lack of effort.

Prior to each task, poor readers gave a lower estimation of their expected success rate than average and good readers. Poor readers showed a significantly greater decline in their expected success rating after each task than average and good readers when they had experienced failure. In other words, the poor readers were not only less confident to start with, but also showed a greater loss of confidence following failure.

As Butkowsky and Willows (1980, p. 419) point out, the lack of persistence which results from the poor readers' beliefs and related attribution patterns is likely to have a detrimental effect on their learning and chances of future success.

> *"Poor readers' markedly lower persistence in the face of difficulty may be seen as a self-defeating behaviour, one that is integrally linked to the maintenance of a low self-concept of ability. Children who give up easily in the face of difficulty may never persist long enough at a task to discover that they do possess the capacity to achieve outcomes that exceed their expectations. When such children do succeed, it is likely to be at tasks that have lower difficulty levels. Such successes are far less self-reinforcing (and far less likely to be reinforced) than*

those associated with greater effort expenditure. These data provide convincing evidence in support of the notion that children with reading difficulties may display an eroding motivation in achievement situations that increases the probability of future failure."

Johnston and Winograd (1985) also draw attention to the importance of motivational factors in explaining poor reading performance. They note that poor readers may not be deficient in skills, but rather may be passive learners who do not apply their skills in situations where such skills are required to perform well. Johnston and Winograd refer to the work of Dweck in relation to this point; Dweck and Repucci (1973), for example, showed that children who Johnston and Winograd would consider as passive learners performed much less efficiently in a situation where they had experienced failure than in a situation where they had experienced success, even though the problems to be solved in the two situations were almost identical. The cognitive correlates of passive failure, then, include reduced and less effective strategy usage, and poor monitoring and assessment of performance. The motivational correlates include attributing failure to low ability and success to external factors, resulting in low expectations of success and low persistence. The affective correlates include low self-esteem, resignation and apathy.

Johnston and Winograd consider that teachers can encourage passive failure if they set common reading tasks in a competitive atmosphere, making it likely that some students will experience frequent failure, and if they interact with lower achievers in ways which further suggest that their status is fixed and unalterable. On the other hand, Johnston and Winograd suggest that passive failure can be minimised if teachers focus on and reward appropriate strategy use rather than

a final product, and if they encourage and help students to assess their own work and the outcome of their strategy use. In this way, students may become more aware of the importance of effort in determining outcomes, although in an atmosphere which removes the fear of teacher evaluation and failure.

Like Butkowsky and Willows (1980), Carr, Borkowski and Maxwell (1991) compared normal and underachieving students, in this case in grades 3, 4 and 5. Underachievers were those children reading at a level below that which would be predicted from their I.Q. Students completed an attribution questionnaire in which they expressed their beliefs as to why particular success and failure outcomes had occurred. They completed the Reading Awareness Questionnaire developed by Paris and Jacobs (1987) to assess knowledge about evaluation, planning and regulation of reading, as well as conditions under which particular strategies were useful. To assess strategy use, children were given writing materials and told to do anything that would help them to remember the paragraphs which were read to them twice. Recall of paragraphs was also scored. Finally, the children completed a self-esteem questionnaire.

Three factors distinguished the normal and underachieving children in a discriminant analysis, these being reading awareness, self-esteem and attributions for success. Underachieving readers showed less reading awareness, lower self-esteem, and were less likely to attribute success to effort. Of the relationship between self-esteem, attributions, strategy use and performance, Carr et al (1991, p. 117) conclude:

> *"... children's successes are dependent, in part, on their belief that effort counts and that they are in control of academic progress. There is otherwise little reason to apply strategies to problem-solving tasks*

and to acquire metacognitive knowledge. If the impetus for achievement is external to children (e.g., a belief in luck or in the necessity of help from others), it is unlikely that they will develop feelings of self-esteem and a repertoire of high-level metacognitive skills (especially executive skills) necessary for good performance."

To conclude, a training programme needs to take into account these motivational factors. Whether or not poor readers are deficient in strategy knowledge, they are unlikely to apply either familiar or newly acquired strategies in an optimum way if they do not believe that the effort involved in such strategy use will lead to a successful performance.

2.5.3 Comprehension Monitoring

The skill of comprehension monitoring is one which is assumed to be important to proficient comprehension, and hence may be another source of difference between skilled and less skilled readers. Baker (1984b, p. 289-90) describes the importance of such a difference as follows:

"An important component of effective reading and of effective listening is the ability to evaluate the current state of one's own comprehension ... Children who fail to decide for themselves when they understand and when they do not are at a serious disadvantage in a wide range of situations which demand adequate comprehension ... The disadvantage arises because children who do not evaluate their comprehension also do not regulate it; that is, they do not attempt to remedy comprehension problems. The logical relationship between these

two components of comprehension monitoring is that evaluation must occur prior to regulation; consequently, any difficulty in the former activity will necessarily impair the latter."

Comprehension monitoring is often assessed by means of the error detection paradigm, that is, various types of errors are deliberately inserted into texts and the ability of readers to detect these errors is assessed. It is assumed that the careful and efficient reader who is monitoring comprehension will find the errors. However, some problems with the error detection paradigm will be discussed further below.

Markman (1979) has pointed out that the reader must engage in constructive processing in order to be in a position to notice certain types of errors, such as internal inconsistencies. If successive statements are not related to each other by the reader, the inconsistency will not become apparent. Since age and reading skill differences in constructive processing generally have been shown in previous sections, and since such processing is a necessary condition for at least some types of error detection, one would predict that age and reading skill differences would also be found in comprehension monitoring.

To begin with an examination of age differences. Harris, Kruithof, Terwogt and Visser (1981) observed that children in grade 6 were more able to report textual anomalies than children in grade 3. In this case the texts were script-based narratives, such as going to the dentist. A target line in each text was either consistent or anomalous given the events normally associated with a particular script. These lines were chosen in such a way that they were consistent with one title, such as Going to the Dentist, but anomalous given another title, such as Going Shopping. Although the younger children read more slowly overall than

the older children, both age groups were equally affected by the anomaly, that is, their reading time increased to a similar extent. Harris et al (1981, p. 227) assumed that *"... variation in reading time reflects the ease with which a set of new concepts can be either retrieved or integrated with the subject's representation of the story up to that point."* Based on this assumption, Harris et al conclude that both grade 3 and grade 6 children were engaging in constructive processing, but the older children used more effective checking procedures to enable them to detect anomalies.

Zabrucky and Ratner (1986) similarly found no age difference in the effect on reading time due to anomalies, but did find differences in other aspects of comprehension monitoring and regulation. Children in grades 3 and 6 read short narratives, half of which contained an internal inconsistency, with the contradictory pieces of information either in adjacent sentences (close) or separated by four sentences (far). Children were not warned of errors, but were told to read for meaning and were videotaped while reading. They recalled what they could of each story and were asked whether everything made sense (first probe). They were presented with context and target sentences, the latter containing the contradiction in the inconsistent condition, plus two filler sentences and asked whether or not these sentences had occurred in the story. Finally, the pieces of contradictory information were presented in one sentence and children were asked whether the sentence made sense (second probe). Thus, Zabrucky and Ratner used several measures of monitoring and regulation, including the on-line measures of reading time and lookbacks scored from the video recording, plus the off-line measures of free recall, recognition memory, and verbal reports of error detection in the first and second probes.

Grade 3 students read target sentences more slowly than grade 6 students, but both groups were equally affected by an incongruous target sentence. In other words, the increased processing difficulty slowed their reading to a similar extent, thus replicating the results of Harris et al (1981). However, on other measures of error detection and regulation of comprehension there were age differences. Grade 6 children reported more errors after the first general probe than grade 3 children, and similarly reported more errors after the second probe, when contradictory information was presented in the same sentence. In this situation, most of the grade 6 children but only about half of the grade 3 children reported the error. Grade 6 children were more likely than grade 3 children to look back to previous sentences, particularly to the context sentence when it was four sentences prior to the target sentence in the far condition. Zabrucky and Ratner take this finding to show that the older children more actively regulate their comprehension, that is, attempt to repair errors they have detected.

The poorer error detection performance of the younger children could not be attributed to poor memory for the contradictory pieces of information, since there were no age differences in the recall of context or target sentences, nor were there any age differences in the recognition of incongruent context and target sentences. Grade 3 children did not make inferences to repair errors; there were only seven fix-up inferences of a possible total of 160 instances of inconsistency that could have been repaired. The poor performance of the younger children also could not be explained by their inability to make an adequate verbal report; the error detection results were much the same with lenient scoring where children were credited with error detection for only saying the story didn't make sense, without any further explanation. In trying to explain why an age difference was not found

on-line (effect of inconsistency on reading time), but was found in verbal reports (error detection), Zabrucky and Ratner (1986, p. 1414) suggest that:

> *"...beginning readers process information in a constructive fashion and show rudimentary skills at applying both internal and external standards of evaluation, but that their processing may not result in conscious awareness of non-comprehension or that they may not know how to interpret such awareness."*

There were very few systematic relationships between different measures. There was little or no relationship between reading times and error detection, reading times per sentence and frequency of lookbacks, reading times and memory, frequency of lookbacks and memory, or lookbacks and error detection. However, for grade 6 but not grade 3 children there was a strong relationship between memory, both free recall and recognition, and error detection. Thus, the picture of possible relationships between what are postulated to be measures of monitoring and regulation is confused, although Zabrucky and Ratner state (1986, p. 1416) regarding the link between memory and error detection in the older children: *"This finding certainly fits our expectation that evaluation of understanding is an important skill for reading comprehension and that comprehension of material should be reflected in higher retention of material."*

Tunmer, Nesdale and Pratt (1983) found improvement with age in the ability to detect both implicit and explicit internal inconsistencies. Children in preschool, grade 1 and grade 2 heard passages composed of three sentences from a puppet, and had to decide whether the story was sensible or silly. For example, in one explicit text the first sentence stated that one cannot ride a bicycle with a broken

wheel, while in the implicit version the first sentence stated that a boy had received a bicycle as a present. The second sentence for both versions stated that the boy's bicycle was run over and had a broken wheel. The final sentence either stated that the bicycle was taken to be fixed (consistent) or that the boy rode away on the bicycle (inconsistent).

Explicitly stated contradictions were easier to detect, and there was an improvement with age in error detection. Even the youngest children showed clear evidence of ability to monitor comprehension, although they had some difficulty in the implicit condition. Even in the implicit consistent condition they sometimes rejected stories as silly on empirical grounds. Tunmer et al (1983, p. 107) conclude:

> *"These children seemed less able to engage in the inferential and constructive processing necessary to obtain information about the relationships that can occur among sentences in implicit stories. Instead, they tended to question the truth of individual sentences, rather than integrate the story as a whole and examine its overall logical structure."*

Baker (1984a) compared children in preschool, grade 1 and grade 3. All children were informed that there might be errors in the narratives to be heard, and were given examples of the type of nonsense words, falsehoods and internal inconsistencies that could occur. Children in grade 3 detected more errors that children in grade 1, who in turn detected more errors than preschoolers. However, all children showed some ability to use the three standards of evaluation under the 'ideal' conditions in this experiment, where simple stories were heard twice,

children were informed of the type of errors involved, and, if children failed to find the errors, they were pointed out by the experimenter.

Baker (1984b) incorporated the same three types of errors into expository texts. Children in grades 4 and 6 were to underline any problems as they read. Half the children were told only that problems might exist (general instructions), while the other half were given examples of each type of problem (specific instructions). If a problem was not underlined, increasingly specific probe questions were asked to direct attention to the problem.

Overall, the grade 6 children underlined significantly more problems than the grade 4 children (53% and 34% respectively). However, this difference occurred only amongst good readers, as defined by the California Achievement Test; amongst poor readers the older children did not perform better than the younger ones. If probing was required, older children found problems sooner than younger ones. Not all children used all of the three standards they were instructed to use in the specific condition. Baker (1984b, p. 307-8) suggests that even when particular standards of evaluation are within the repertoire of children, they may not be able to apply them simultaneously for a variety of reasons:

> *"The data suggest that one reason younger and less proficient readers identify fewer problems in text than older and more proficient readers is that they adopt fewer different criteria for evaluating their understanding. However, this is not the only explanation, for even when the necessary criteria are available in the child's repertoire, the younger and poorer reader still identifies fewer problems. This implicates differences in the consistency and effectiveness of standard use. Children may not use the standards consistently when there are many*

competing demands on resource allocation ... Thus, a less experienced reader may have more difficulty simultaneously checking external consistency, internal consistency and word understanding, and so only checks one or two dimensions. The effectiveness of standard use is limited by logical reasoning abilities and general world knowledge, as children's responses to the follow-up questions revealed. The younger and poorer readers had difficulty realizing that certain parts of the texts were in fact problematic. Just as basic cognitive skills and capacities play a crucial role in comprehension so too do they contribute to comprehension monitoring."

Beal (1990) also draws attention to the role of constructive processing, not only in error detection, but also in error repair. Children in grades 4 and 6 read four types of stories, including normal versions, versions with key sentences missing so that no resolution was given to the problem described in the story, versions containing anomalies violating prior knowledge expectations given the context, and versions containing logical contradictions. The children had the stories read to them with the text present. They were asked for their help in locating parts that were not clear and in making any such parts clearer.

Grade 6 children detected significantly more errors than the grade 4 children, who performed quite poorly. All children found the missing sentences easier to detect than the anomalies and inconsistencies. Given that an error was detected, there was no age difference in terms of the adequacy of suggested repairs.

In the first experiment, the problem of the younger children appeared to be in detection rather than repair of errors. However, since without detection there could be no repair, and detection rates among the grade 4 children were low, the

first experiment gave an incomplete picture of the younger children's ability to repair errors. In a second experiment, if children did not detect an error the story was read again, and if the error was still not detected the experimenter pointed it out and asked the children how it might be repaired. Grade 4 students required such prompting in an average of half the stories, compared to less than one-third in the case of the grade 6 children. The older children were able to suggest more adequate repairs for the errors exposed during prompting.

The final experiment investigated whether children were able to evaluate the quality of repairs they did not have to generate themselves. A sentence was deleted from stories, with the deletion marked by a blank line. Students were asked to explain why the deletion was problematic, and if they did not know they were told. They were then asked to evaluate a sentence which was placed on the blank line. The versions included the original sentence, a detail sentence which restated the problem but offered no solution, an implicit solution which suggested enabling conditions but no explanation as to how they contributed to the solution, and a solution which was inconsistent with the theme or setting.

More children in grade 6 than in grade 4 were able to explain the nature of the problem, to judge the adequacy of the proposed sentence, and to suggest a better one if the proposed sentence was inadequate. Thus, in this study overall, grade 6 children were more able to detect errors, to repair errors, and to judge the quality of proposed repairs.

Beal suggests that constructive processing and inferential comprehension are required to detect errors of the type included in this study. The memory representation which results from such processing then also provides a basis for generating and evaluating repairs. In the absence of such processing, repairs are difficult to

carry out, even when errors are highlighted by the experimenter.

In a study by Englert, Hiebert and Stewart (1988), students in grade 6 detected more errors while reading expository texts than students in grade 3. All children were informed that there might be problems in the texts and were shown how to respond to errors. In spite of this quite specific caution, only 38% and 18% of the errors were detected and repaired by grade 6 and grade 3 children respectively. Among grade 6 students, 41% of errors were neither detected nor repaired, compared to 57% among grade 3 students. The errors in this case were falsehoods, internal inconsistencies, and sentences not fitting into the overall text structure, for example, a descriptive sentence in a compare and contrast text.

Vosniadou, Pearson and Rogers (1988) used narrative texts which contained either a falsehood or an internal inconsistency, that is, a statement which contradicted a principle stated earlier. In one example, a magnet picked up sugar but not iron filings, although it had been stated that magnets picked up iron and not sugar. Children in grade 3 detected more errors than children in grade 1 after listening to the stories, and children in grade 5 detected more errors than children in grade 3 after reading the stories. When children failed to detect an error, most commonly they could not recall the critical propositions. Vosniadou et al deduce from this observation that children's problems in detecting errors lie primarily in establishing an adequate representation of the text, which in turn is related to the familiarity of the information being processed. In support of this claim, Vosniadou et al show that even children in grade 1 can detect internal inconsistencies when the subject matter of a text is familiar and easy to represent. Vosniadou et al (1988, p. 36) conclude:

> *"Children often lack the background knowledge necessary to form an adequate mental representation, or they may have beliefs or theories that are contrary to the information included in the text. Inadequate and/or incomplete prior knowledge may result in distortions of the text information, failures to detect inconsistencies, repair of inconsistencies, or a disposition toward evaluating the text for truth rather than inconsistency."*

Walczyk and Hall (1991) asked children in grades 2 and 4 to listen to narratives with a particular perspective in mind. For example, in one story children were asked to focus on information relevant to a burglar, in an other on information relevant to a safety inspector. Some of the stories contained falsehoods. The authors reasoned that if readers were asked to adopt a perspective, relevant prior knowledge would be activated and the probability of detecting a falsehood (prior knowledge violation) would be increased.

Recall protocols of the stories indicated that children in grade 4 could apply a perspective on-line while listening, but grade 2 children could not. In other words, the older children recalled more perspective-relevant than perspective-irrelevant information, while the younger children did not make a distinction. The older children also detected more errors, and again, unlike the younger children, were influenced by perspective in that they detected more perspective-relevant than perspective-irrelevant errors. The age difference in error detection occurred in spite of the fact that a quiz revealed all children had the necessary general knowledge to detect the falsehoods. In this respect the Walczyk and Hall study does not support the conclusion of the Vosniadou et al study regarding inadequate general knowledge as a reason for failure to detect errors. However, like Vosniadou

et al, Walczyk and Hall found that in most cases where the grade 4 children did not detect an error they could not recall the critical sentence; only 13% of these children recalled all the necessary information, but failed to detect the error.

Although the studies referred to so far show significant age differences in error detection, not all results are so clear cut. For example, Markman (1979) wrote short expository texts containing internal inconsistencies which were either explicit or required an inference to become apparent (implicit). Children in grades 3, 5 and 6 heard the texts twice. They were not warned of any errors, but were asked to make suggestions that would make the texts easier to understand. After each passage the children were asked some general questions, then recalled the passage and, if the inconsistencies had not been reported, specific probing was initiated.

There was no grade difference in error detection in either the explicit or the implicit condition. After hearing three passages, each containing an error, 96% of all children missed all or all but one of the implicit contradictions, while 43% missed all or all but one of the explicit contradictions prior to the initiation of specific probing. This relatively poor performance occurred in spite of the fact that 86% of recall protocols contained no significant error or omission, and that children could answer specific probe questions, indicating that they had the necessary general knowledge and inferential capacity to detect the errors.

In a subsequent experiment, children in grades 3 and 6 repeated either each entire text, or two sentences at a time to see if having the contradictory information in working memory contiguously would increase error detection rates. In the implicit condition, repetition had no effect on performance, and grade 6 children did not detect more errors than grade 3 children. However, in the explicit

condition, the older children detected significantly more errors than the younger ones. Repetition of two contradictory sentences did not guarantee error detection among grade 3 students; over two-thirds of these children failed to detect any of the explicit contradictions, compared to less than one-fifth of the older children.

In the final experiment, Markman warned half the children that the texts may contain problems. Without this warning, grade 6 students performed as poorly as grade 3 students. With the warning, however, the older children performed significantly better than the younger ones and improved their detection of both explicit and implicit contradictions. Thus, the older children had the requisite knowledge and skills, but did not apply them spontaneously to comprehension monitoring.

A similar result was obtained by Markman and Gorin (1981), who used expository texts containing either falsehoods or internal inconsistencies. One group of children was given neutral instructions, the task being to decide if everything was easy to understand (control group). Two other groups were told of the possible existence of errors and given examples of one of the error types (specific groups). Ten year-old children detected more errors than eight year-olds only in the specific conditions; in the control condition the older children performed no better than the younger ones. So age differences depended on being informed of the existence of errors. As in the Markman (1979) study, older children had the skills to detect errors if they were told to do so, but they did not apply these skills spontaneously any more than younger children.

Scardamalia and Bereiter (1984) had students in grade 6 and grade 10 read two of the passages containing internal inconsistencies devised by Markman (1979). These students thought aloud as they interpreted the passages. The grade 6

students made marginally more detail interpretations than the older students, that is, they referred to only one sentence in their interpretation. On the other hand, grade 10 students made significantly more macro interpretations, which related statements to each other or the overall gist of the passage.

Most students recognised the anomaly in the passage about ice-cream. However, one third of readers made no comment at all about anomalous information in the fish passage, and only about one third in some way recognised the anomaly. The authors suggest that the ice-cream passage was easier because the anomaly could be found by questioning the plausibility of individual statements, while in the fish passage individual statements had to be linked. The importance of what Scardamalia and Bereiter refer to as a macro interpretation is highlighted by a significant correlation between successful error detection and a more sophisticated reading strategy involving macroproposition formation.

Zabrucky and Moore (1989) point out that whether or not age differences are found can depend on the standard of evaluation being examined, for example, whether that standard is word understanding or internal consistency. In this study, children in grades 4, 5 and 6 read expository texts which contained either a nonsense word, falsehood or internal inconsistency. All the children were told that there could be errors in the texts, although half of them were also given examples of the three types.

Children in grade 6 detected more of each kind of error than children in grade 4, but the position of the grade 5 children relative to the other two groups varied depending on the standard of evaluation being examined. This middle group performed at a similar level to grade 6 children in the detection of nonsense words, but at a similar level to grade 4 in the detection of internal inconsistencies. Thus,

the application of a lexical standard of evaluation, asking whether each word makes sense, appears to develop early, while the application of an internal consistency standard, asking whether each idea fits in with previous ideas, develops later. The detection of falsehoods increased linearly from one grade to the next. Zabrucky and Moore point out, therefore, that 'comprehension evaluation ability' is multidimensional, that the use of the various standards involved develops at different rates, and therefore results regarding age differences may vary depending on which standards are being studied.

Grabe, Antes, Kahn and Kristjanson (1991) compared adults and grade 7 students in their ability to detect falsehoods and internal inconsistencies in short expository passages. A pilot test was conducted to ensure that the problematic information was familiar to grade 7 children. Half the subjects were informed that errors might be present and were given specific examples of both types, while the other half were not informed.

Error detection performance was assessed only among informed subjects. Here adults were no more able than children to detect internal inconsistencies. However, adults did find more of the falsehoods, 53% compared to the 28% found by children. Overall less than half the errors were detected.

In summary, the studies reviewed in this section show improvement with age in the ability to detect errors, which is assumed to reflect increasingly efficient comprehension monitoring. However, there are also certain exceptions to this general conclusion. The spontaneous performance of older and younger children did not differ in Markman (1979) and Markman and Gorin (1981); only when given specific instructions did older children find more errors. Older children performed better than younger ones only if they were proficient readers in Baker

(1984b); there were no age differences among poor readers. Whether or not age differences were found depended on the standard of evaluation in Zabrucky and Moore (1989) and Grabe et al (1991).

Although older children in general are better at detecting errors than younger children, it must also be kept in mind that the performance of adults is often far short of perfect. As already mentioned, in the Grabe et al (1991) study, adults detected only 53% of falsehoods after being warned that problems might exist. In Baker (1979), where subjects were not informed of errors, only 38% of errors were detected when passages were re-examined; subjects claimed to have noticed 23% of errors during the initial reading. These results reflect somewhat less poorly on the comprehension monitoring ability of the subjects when examined in more detail. These adults found 62% of errors embedded in superordinate, topic sentences but only 14% embedded in subordinate, detail sentences, which is consistent with the goal expressed by many of them that they were reading for main idea comprehension. Also, internal inconsistencies were detected much more efficiently than inappropriate connectives; in the case of the latter, some subjects were not credited with correct identification because they classified the problem as an internal inconsistency rather than an inappropriate connective. The point is, though, that they did detect a problem.

In Baker and Anderson (1982) adults found 66% of internal inconsistencies in topic sentences and 62% in detail sentences. Being informed of possible errors did not affect performance. Subjects spent longer reading an inconsistent sentence than its consistent counterpart in an intact passage. They also reread inconsistent sentences more often. However, Baker and Anderson note that they found *"little individual consistency and few systematic relationships"*. Only 25% of subjects

correctly identified errors in all three passages, and subjects who spent more time reading inconsistent sentences were no more likely to identify the error than those who did not spend the extra time.

In Baker (1985b) 34% of errors, including nonsense words, falsehoods and internal inconsistencies were detected. All subjects were informed that errors might exist, but error detection was higher when examples of the three types of errors were also given. Detection rates were high for nonsense words, but relatively low for internal inconsistencies. Baker reports that many subjects exclaimed: *I can't believe I missed that* when errors became apparent during probing.

Yussen and Smith (1990, expt 3) also used internal inconsistencies with informed subjects, and found that these adults averaged 5.30 of a possible score of 8 for the detection of inconsistencies in topic sentences and 3.65 for detection in detail sentences.

Glenberg, Wilkinson and Epstein (1982) examined not only the error detection performance of adults, but also confidence in their performance. Subjects read passages of either one or three paragraphs, in which the last two sentences were contradictory. In the last sentence, the contradictory sentence was syntactically marked as either given or new. The authors hypothesised that with three paragraphs there were more potential points of attachment for incoming information which needed to be checked, thereby making error detection more difficult. If information was marked as new, the authors predicted that readers would be less likely to search for a point of attachment and therefore would be less likely to detect an inconsistency. Subjects were told to read carefully and were forewarned of the presence of errors, which they were asked to find. After reading, they rated their confidence in their comprehension on a scale of 1 to 4.

The probability of detecting an internal inconsistency ranged from a high of .89 in the given-1 paragraph condition to a low of .38 in the new-3 paragraph condition. Of particular interest were confidence ratings in relation to error detection. Glenberg et al determined students with an illusion of knowing (IK) on the basis of a failure to detect an inconsistency, but a high confidence rating. The probability of IK ranged from .06 in the given-1 paragraph condition to .51 in the new-3 paragraph condition. Thus, the latter condition produced not only a high rate of failure to detect an inconsistency, but at the same time a still high level of confidence in comprehension success. Glenberg et al (1982, p. 602) explain this finding as follows:

> *"The frequent occurrence of the illusion of knowing is inconsistent with an assumption of active, accurate, on-line comprehension monitoring as the source of this confidence. Instead, we propose that confidence arises from a default assumption made by readers, that is, in dealing with expository text, a reader assumes by default that comprehension is progressing smoothly, and this assumption is maintained until the reader is alerted otherwise by an error signal arising from a failure in some comprehension process... Readers may favour the comprehension monitoring procedure afforded by the default assumption on the grounds of economy of processing. No special processes need be developed or activated to assess comprehension. Importantly, the process of waiting for an error signal can operate in parallel with comprehension processes, while requiring little, if any, capacity. Readers need not interrupt comprehension processes to assess their efficacy. Also, except for texts in unfamiliar domains, the assumption of com-*

prehension is often reasonable. Unfortunately, for students and educators, it may be that when reading in new domains our ability to detect ignorance is most impoverished. On such occasions, the illusion of knowing becomes an obstacle to effective learning and instruction."

Less than perfect assessment of comprehension has been demonstrated without using error detection in a study by Pressley, Ghatala, Woloshyn and Pirie (1990). In this study adults read expository passages and then answered either a multiple choice or short answer question on general thematic or detailed information from the passage. Depending on how they evaluated the correctness of their response, readers had the option of moving on to the next passage or rereading the current passage. Around 75% of the thematic questions and 50% of the detail questions were answered correctly. Following detail questions subjects had a moderate probability of making a correct judgement, that is, correct response - move forward, or incorrect response - reread. However, for thematic questions subjects tended to move forward regardless of whether or not their response was correct; the probability of rereading after an incorrect response was only .27 for short answer questions and .20 for multiple choice questions. Pressley et al (1990, p. 242) conclude that:

"Given that participants claimed after the study that they decided to reread the passage only when they were not sure that their answers were correct, we conclude that subjects were confident about their answers to thematic questions even when those answers were in error ... errant responses were in fact related to the passages in question, although the answers were either incomplete or there was

some distortion of the passage content. A reasonable hypothesis is that readers have great confidence in whatever general interpretation they have constructed while reading a passage. If so, this is a very serious deficiency in comprehension monitoring, because awareness of miscomprehension is crucial to regulation of reading and especially to regulation of text reprocessing."

Like Pressley et al, Schommer and Surber (1986) also investigated the comprehension monitoring of adults without using an error detection task. They used two different texts, one more difficult than the other because of longer sentences and more embedded clauses. Subjects were given either shallow or deep processing instructions by being set different goals for the reading activity. In the shallow condition, readers were to judge whether the passage could be understood by the average college freshman, while in the deep condition, they were to read in order to be able to teach another student enough information to pass a test. In the latter condition the subjects were told to produce a summary of what they intended to teach their 'student'. The authors hypothesised that the shallow condition would encourage processing only at the level of individual micro-propositions, leaving open the possibility that the reader would be unaware of not having understood the main ideas. In other words, these readers would be susceptible to illusions of knowing. The deep condition, however, should encourage readers to construct a macrostructure, the success of which could be used as a measure of gist comprehension. It is more likely, then, that the reader would be alerted to a failure to understand.

Subjects read one passage, rated confidence in their comprehension, completed a summary of the passage, rated their confidence in the quality of the summary,

answered three multiple choice questions on main points, and rated their confidence in these responses. A subject was classified as having an illusion of knowing (IK) if two or three of their multiple choice responses were wrong, but they nevertheless expressed a high level of confidence in their comprehension. Readers in the shallow condition showed more IK, especially when reading the more difficult text. One quarter of these adults were classified in the IK category, that is, they failed to understand the main ideas of a passage in the manner intended by the author, and yet had a high level of confidence in their comprehension. These results support the findings of Pressley et al.

Schommer and Surber examined the accuracy of concepts and relations included in summaries. Those students who both scored well on the multiple choice questions and rated their confidence highly were very likely to produce model summaries, while students performing poorly and rating their confidence as low produced impoverished summaries, missing several concepts and relations. The IK group, on the other hand, produced not so much impoverished as distorted summaries. The authors suggest that they had enough information to construct a macrostructure, albeit an inaccurate one, and were not alerted to the fact that they had not understood the main ideas.

The studies discussed thus far show a trend for older subjects to be more able to detect errors and evaluate their comprehension, although even the performance of adults is not perfect. In the next section, the error detection performance of subjects varying in reading skill will be examined. There is considerable evidence that skilled readers are more able to detect errors than less skilled readers.

Yussen and Smith (1990) investigated detection of internal inconsistencies among adults. There was a strong correlation ($r = .61$) between reading pro-

ficiency, as measured by the Stanford Diagnostic Reading Test, and ability to detect errors while reading. However, the correlation between reading proficiency and error detection while listening was not significant.

Zabrucky and Commander (1993) examined on-line detection of errors and rereading as a strategy to overcome comprehension problems. University students classified as either good or poor readers according to the Nelson Denny Reading Test read short expository texts one sentence at a time from a computer screen. Half the texts were consistent and half contained an explicit logical inconsistency. Information presented in a context sentence was later either reaffirmed or contradicted in a target sentence. Students were not warned of the presence of errors but were told only to read carefully.

Incongruous target sentences were read more slowly than congruous target sentences by both groups of readers, suggesting that both encountered a problem when processing the incongruous sentences. While good readers spent more time rereading the target sentence than poor readers, poor readers spent more time than good readers rereading sentences other than the context and target sentences, showing that they were less selective in their application of the rereading strategy.

Zabrucky and Commander discuss and eliminate several possible explanations for this finding. For example, while it is possible that poor readers reread more because everything made less sense to them, this explanation is unlikely given that the texts were short and simple, written at year 10 readability level. Zabrucky and Commander (1993, p. 452) favour an explanation in terms of working memory limitations, since selective rereading involves holding problematic information in memory. As they state: *"In order to reread effectively it is necessary for an*

initial signal of comprehension failure to be generated and for the reader to hold onto the signal while simultaneously attempting to find earlier information that is inconsistent."

As a final point, there was no ability difference in terms of verbally reporting a problem; the mean rate of reporting incongruities after reading was .64, supporting the earlier point that the performance of adults is still short of perfect.

Englert, Hiebert and Stewart (1988) showed that children with high reading ability detected more errors in expository text than children of medium ability, who in turn detected more than children of low reading ability. The children were in grades 3 and 6, the errors involved falsehoods and internal and text structure inconsistencies, and ability was measured by the reading subtest of the California Test of Basic Skills.

Garner (1980) asked good and poor comprehenders, as determined by teacher judgement, to rate the comprehensibility of expository texts. The children were in grades 7 and 8. Some segments of text had been altered to include an internal inconsistency. Overall the comprehensibility of altered segments was rated as lower, although this trend was more evident among good comprehenders. Poor comprehenders did not distinguish to the same extent between altered and unaltered segments in their ratings, and commented on long words as a reason for lower ratings. Good comprehenders, on the other hand, could point to the internal inconsistency as the reason why a segment was harder to understand.

Garner and Kraus (1981/2) asked good and poor readers in grade 7, determined by teacher judgement, to read short narrative texts which contained either inter- or intra-sentence internal inconsistencies. The detection of inter-sentence contradictions was low among both good and poor comprehenders, the latter

finding none at all. The intra-sentence inconsistencies were detected by 80% of the good comprehenders, but again none of the poor comprehenders.

In interviews with these children Garner and Kraus found that the poor comprehenders focused almost entirely on knowing and fluently decoding words in their judgement of what made a good reader and what made something difficult to read. Good comprehenders, on the other hand, focused on larger units of meaning and considered a good reader to be one who understood main points. Linking together their two findings, Garner and Kraus suggest that poor comprehenders focus on individual word meanings, do not integrate these meanings into larger units, and hence do not find internal inconsistencies. The focus on individual words was not due in this case to decoding problems, since only adequate decoders were chosen for the study.

Garner (1981) further examined the hypothesised individual word emphasis of poor comprehenders in grades 5 and 6, classified according to the comprehension subtest of the Maryland screening instrument. Children read short narrative passages in which some segments were altered to contain either long and difficult words or an internal inconsistency. They were asked to rate the comprehensibility of segments, hence were not alerted to the existence of errors. Comprehensibility ratings were lower for segments containing long words than for both consistent and inconsistent segments, which did not differ significantly from each other. Readers pointed to the long words as the reason for their lower ratings, but were vague about why consistent and inconsistent segments had been rated in a particular way. Garner interprets this finding as an example of 'piecemeal processing', that is, a focus on individual word meanings without integration into larger units of meaning.

Paris and Myers (1981) had children in grade 4 read aloud two texts written at grade 3 and grade 5 level respectively, then underline anything that didn't make sense in two further texts of similar difficulty. Two nonsense words or phrases had been inserted into each text. The children were divided into good and poor readers according to the SRA Assessment Survey, with the two groups matched according to age, sex and maths achievement.

In oral reading, hesitations, repetitions and self-corrections were taken as evidence of monitoring. There was no overall group difference in frequency of hesitations and corrections, although monitoring was more accurate among good readers for the grade 5 but not the grade 3 text. In other words, the good readers on the more difficult passage showed greater evidence of having noticed the nonsense words and phrases, while the poor readers were more likely to show hesitations or self-corrections with acceptable words.

With regard to the second task, good readers underlined more nonsense information, although this difference was significant only for nonsense phrases, not words, with both groups performing poorly in the latter case. Again there was no difference in the overall frequency of underlining, but the good readers were more likely to underline problematic information while the poor readers underlined more legitimate information. Good readers also produced more accurate recall.

Winograd and Johnston (1982) used skilled and unskilled readers in grade 6, classified according to the Stanford Diagnostic Reading Test. Subjects read four short passages, into each of which an anomalous sentence had been inserted. Skilled readers detected significantly more of the problems than less skilled readers (56% compared to 18%). However, the authors note that the overall result was

lowered due to a particularly poor performance on one passage. In this case, many children constructed an inference which made the supposedly anomalous sentence meaningful, hence no error was reported. Error detection rates when this passage was excluded were 88% and 45% for skilled and unskilled readers respectively.

Garner and Reis (1981) examined responses to comprehension obstacles rather than error detection per se. Children read a narrative passage divided into three sections and answered questions after each section. Some of the questions asked after the second and third sections referred back to the first section. Considering the detailed information being tapped, these questions were difficult to answer without looking back. The experimenters recorded both the incidence of look-backs and what they considered to be signs of ongoing monitoring, that is, both verbal and non-verbal signs of puzzlement or dissatisfaction.

Good comprehenders were more likely than poor comprehenders to show signs of puzzlement, to look back to the first section, and to answer the look-back questions correctly. These differences were larger among the grade 8 than the grade 6 and 7 children. With increasing age, good comprehenders showed more look-back and monitoring behaviour, but poor comprehenders showed very little such behaviour even in grade 8. Thus, good readers became increasingly able to both monitor and repair comprehension where necessary, that is, to become aware that they did not have the knowledge to answer look-back questions and to take appropriate steps to find the answer.

A number of studies have found not only reading ability differences in error detection, but also differential effectiveness of specific instructions.

Baker (1985b) compared the error detection performance of adults with either higher or lower Verbal-SAT scores. The subjects read expository passages containing either nonsense words, falsehoods or internal inconsistencies. All subjects were informed that there could be errors in the texts (general instructions), but half of them were also told the exact nature of the errors and given examples (specific instructions). Error detection performance was superior among higher Verbal-SAT subjects, but only in the specific condition. Both the higher and lower ability groups benefited from more specific instructions, but the higher group more so. When a reader failed to detect errors, a probing procedure was initiated, with increasingly specific questions drawing attention to the problem. Although Baker (1985b, p. 310) reports that during probing many readers expressed surprise at their failure to detect a seemingly obvious error on initial reading, she also notes that probing sometimes revealed a lack of relevant knowledge underlying the failure to report a problem:

> *"...the responses to the follow-up questions make it clear that the poorer problem identification by the lower ability students is not simply a product of less frequent or less effective use of the targeted standards of evaluation. More fundamental limitations in background knowledge and logical reasoning also play a role."*

There was no difference between the higher and lower Verbal-SAT students in the frequency of using the lexical standard of evaluation, but the higher group used the internal consistency standard more frequently than the lower group. Half of the latter did not use the internal consistency standard even when instructed to do so. Baker (1985b, p. 311) comments:

> *"A number of component skills are involved in the evaluation of internal inconsistency, including the ability to retrieve and integrate nonadjacent propositions and draw logical inferences. These prerequisite skills may well be limiting the lower verbal students' use of the standard."*

Baker (1984b) also found, this time among children rather than adults, that good readers were more able to make use of guidance to increase their error detection rate. Children in grades 4 and 6 were divided into skilled and less skilled readers according to the reading subtest of the California Achievement Test. The children read expository texts which contained either nonsense words, falsehoods or internal inconsistencies. Subjects were again given either general or specific instructions. Probing was initiated after all texts had been read.

Overall, skilled readers detected more errors than the less skilled readers (54% and 29% respectively), but there was also an interaction with instructional condition. Although both groups benefited from specific instructions relative to general instructions, skilled readers gained greater benefit. Overall there was no difference between the two groups in their frequency of using standards of evaluation, that is, in the number of times the clarity of passages was questioned. From the results above, though, it is clear that skilled readers used the standards more effectively to find errors. Also, although the two groups did not differ in the frequency of standard use overall, they did differ in terms of which standards were used most frequently. Less skilled readers made more use of the lexical standard than skilled readers, but used the internal consistency standard less frequently. In response to probe questions, skilled readers found previously undetected problems more quickly than less skilled readers, again indicating, as in the case of the

specific instructions, that they are more able to make use of guidance.

In a study of children in grades 4, 6 and 8, Garner and Taylor (1982) also found that skilled readers are more able to make use of guidance. In this case, the children were identified by their teachers as either good or poor comprehenders, with all being adequate decoders. The children read narrative passages containing an internal inconsistency. They were given a score depending on whether or not problems were identified spontaneously and, if not, how quickly problems were detected during probing.

Good comprehenders detected more errors, and did so more quickly than poor comprehenders, but even the more skilled readers needed assistance. Only 7 of the 90 students identified problems without probing, all of these being good readers. On the other hand, almost all of the poor comprehenders in grade 4 and over half in grade 8 were unable to find any errors, even when the two contradictory sentences were pointed out by the experimenter, with the warning that some people thought the sentences did not fit together. Many good comprehenders who did not detect errors spontaneously nevertheless were able to make use of this assistance to improve their performance.

Zabrucky and Moore (1989) studied children in grades 4, 5 and 6, who were divided into good, average and poor readers according to comprehension scores on the Stanford Diagnostic Test. They read expository texts into which either nonsense words, falsehoods or internal inconsistencies had been inserted. The children were given either general or specific instructions. Good readers detected more errors overall than average readers, who in turn detected more errors than poor readers. This pattern applied to both nonsense words and falsehoods, but for internal inconsistencies the average readers were no more successful than poor

readers. Specific instructions resulted in a higher rate of error detection, but only for good readers. As Zabrucky and Moore point out, it is therefore not just a lack of awareness on the part of average and poor readers that results in their lower level of error detection. Rather, they may not have the skills to apply the standards of evaluation they have been instructed to use.

A final point to emerge from the Zabrucky and Moore study is that ability to apply one standard does not imply ability to use another, so it is difficult to measure a 'global' comprehension monitoring ability. The correlation between ability to detect nonsense words and internal inconsistencies was a low .22, increasing to .37 and .38 between falsehoods and nonsense words on the one hand and falsehoods and internal inconsistencies on the other. Markman and Gorin (1985, p. 325) reach a similar conclusion when they state: *"... it has become clear that there is no single capacity that underlies an ability to monitor one's comprehension, and no single age at which children suddenly acquire this ability."*

August, Flavell and Clift (1984) examined children in grade 5, who were divided into skilled and less skilled readers according to the comprehension subtest of the Comprehensive Test of Basic Skills. The groups were matched on nonverbal intelligence, and all had adequate decoding skills. They read narratives and were informed that some might have a page missing.

Skilled readers were more able than less skilled readers to determine which stories had a page missing, where it was missing, and what it should say. Skilled readers to a greater extent than less skilled readers showed increases in reading time for the sections of text following an omission, thus indicating that they had detected (not necessarily consciously) a problem. Given signs of detection, skilled readers were still more likely than less skilled readers to report the existence of a

problem. For both skilled and less skilled readers, the failure to report an error was most often due to inferences being made to repair the problem.

Once all the stories were read, a review procedure was carried out, with the children being asked to reread the final sections of the stories (where an omission could occur) and to answer an increasingly specific set of probe questions. Here too the skilled readers performed better because they, unlike the less skilled readers, were helped by the review procedure.

Garner and Anderson (1982) failed to resolve the question of whether poor readers could be assisted by more specific instructions. Three groups were formed from intermediate grade students at a summer reading clinic. All performed at or above grade level on word recognition but at least one year below grade level on the comprehension subtest of the Maryland screening instrument. All subjects read a narrative about a holiday in a cabin without electricity. The internally inconsistent piece of information involved children plugging in a television set. The first group was told to read the passage in order to be able to discuss it later, the second group was told to ensure that everything made sense, while the third group was told that there was something wrong and that they should find it. Four of the eight children in each if the first two groups reported the error, compared to six of the eight children in the third group. Garner and Anderson, therefore, could not reach a firm conclusion about the effect of instructions on error detection.

Chan (1988), in contrast to the previous five studies referred to, found that specific instructions benefited poor readers. In a first experiment she matched, according to reading age, poor readers aged 11 in a remedial class with 8 year-old average readers. In the general instruction condition, an anomalous line in

practice passages was pointed out, without any explanation as to why it didn't fit in with the story. In the specific condition, an explanation was given. The children then read stories containing similar anomalies. The remedial readers benefited from specific instructions, while the average readers did not. The remedial readers performed better than the average group only in the specific condition. Chan (1988, p. 29) comments as follows:

> *"In the present study, subjects in the specific instruction condition received explicit explanation on how to compare one proposition with others in a given text. In effect, this explicit explanation served two functions. Firstly, it provided instructions in the use of a cross-referencing evaluative strategy. Secondly, it provided students with clear knowledge of the criterion task, a metacognitive skill allowing them to engage in task-appropriate learning activities."*

In a second experiment, remedial readers in grades 5 and 6 were matched to average readers in grade 5 on chronological age and to average readers in grade 3 on reading age. In this case, the average readers in grade 5 detected more anomalies than the other two groups. Specific instructions improved performance only for the the remedial readers and the average readers in grade 3, so presumably average readers in grade 5 were able to work out for themselves, without specific instructions, why anomalous lines did not fit into a text. Grade 3 students in this case were assisted by specific instructions, as the author notes, perhaps because they were slightly older and more proficient readers than the group in the first experiment. Chan concludes that explicit instructions of the form used here assist students with reading difficulties to evaluate and regulate comprehension in a

more active way. However, on the basis of the five studies discussed previously, it would seem that the situation is not as simple as Chan indicates, and that less skilled readers generally need more than specific instructions in order to be able to improve their error detection performance. It may be that the anomalies used by Chan were easier to detect than the errors in other studies, for example, involving more familiar content or embedded in texts making lower processing demands. In this case, specific instructions could be sufficient to prompt the less skilled readers to use skills available in their repertoire.

Reis and Spekman (1983) similarly attempted to improve the error detection performance of poor readers through specific instructions. The subjects were in grades 6 and 7, and were two to three years below grade level on the Stanford Diagnostic Reading Test. They read two narratives, each of which contained three errors, in one case falsehoods and in the other internal inconsistencies. The children were not warned of errors but were only told to judge if everything made sense. Of the possible score of 3, the average score for detection of falsehoods was 1.3, for detection of internal inconsistencies 0.5.

In a second experiment, Reis and Spekman selected all students who had detected none of the three internal inconsistencies. Two groups were formed, both with an average score of 1.08 for detection of falsehoods. One group received 10 minutes of instruction in two sessions one week apart. During these sessions, the experimenters explained the role of an editor, practised finding errors in single sentences with individual students, practised finding inconsistencies across several sentences, and set out three rules or things for the children to do to find both internal inconsistencies and falsehoods. These children, and the control group who had received no instruction, then read two narratives and decided whether

everything made sense. The same procedure was followed again a week later.

Following the first session there was no significant difference between the experimental and control groups for detection of either falsehoods or internal inconsistencies. After the second session, however, the experimental group detected significantly more falsehoods, while there was still no significant difference between the groups for detection of internal inconsistencies. Reis and Spekman link the lack of training effect in the latter case to an incompatibility between reading goals and the task at hand. They suggest that the instruction to decide whether a story made sense encouraged reading for gist understanding, whereas the internal inconsistencies involved minor details which did not interfere with gist understanding and so were not detected.

While this explanation is certainly plausible, it is also possible that detection of internal inconsistencies, with contradictory pieces of information separated by five sentences, made greater processing demands on readers than detection of falsehoods. For example, Baker (1985b) has pointed out that poor readers may be limited in their ability to retrieve and integrate non-adjacent propositions. In this case it would not be sufficient to tell poor readers what to do in order to establish appropriate goals, but it would be necessary to engage in much more extensive practise to improve the efficiency of integrative processing. As suggested by Markman (1979), specific instructions are only of value when students have the required skills in their repertoire, but fail to use them spontaneously.

Prior to reaching a conclusion about comprehension monitoring as an important area of difference between skilled and less skilled readers, it is important to ask whether the error detection paradigm provides an accurate measure of comprehension monitoring. Winograd and Johnston (1982) have expressed reser-

vations in this regard. They summarise, from previously published studies, twelve possible reasons *other than* failure to monitor comprehension why readers may not report errors. To this list they add some of their own reasons. The following are some of the main criticisms, which will be discussed in turn below:

1. insufficient background knowledge to detect errors
2. failure to recall premises necessary to detect errors
3. use of inferences to repair errors
4. insufficient logical capacity
5. reading goals inappropriate to the task of detecting errors

1. As already indicated, Vosniadou et al (1988) present a strong case that failure to detect errors may be due to insufficient general knowledge, making the information to be read difficult to represent in memory. These authors showed that even grade 1 children can detect internal inconsistencies when the content is very familiar. Markman (1979, p. 654) supports the importance of adequate knowledge when she says:

> *"There are several ways in which the familiarity of material could affect comprehension monitoring. To take one example, familiar material can be more efficiently organized or 'chunked', thus resulting in a saving of capacity to be used for the drawing of inferences, comparison of information, etc."*

Tunmer et al (1983), like Vosniadou et al (1988), found that children in preschool and grades 1 and 2 show evidence of comprehension monitoring

provided that the content is familiar. Walczyk and Hall (1989) reach a similar conclusion as a result of their experiments, in which grade 3 children searched for errors in narratives with more or less familiar themes. Children found more numerical contradictions in texts involving familiar rather than unfamiliar themes, although this difference was not significant for logical inconsistencies. Walczyk and Hall (1989, p. 151-2) conclude:

> *"We are hypothesizing that comprehension monitoring tasks, involving unfamiliar material containing logical inconsistencies are novel. Thus, they require large amounts of control processing for decoding and comprehending at a literal level. This leaves insufficient attentional resources for both retaining material in memory and checking for consistency."*

Baker reports that among both adults (1985b) and children (1984b) limited background knowledge can be a factor which impairs the ability to detect errors. At the same time, however, having sufficient background knowledge does not guarantee error detection. Markman (1979) notes that the children in her study had sufficient knowledge to answer probe questions. Grabe et al (1991) pilot tested their material to ensure that the content was familiar to grade 7 students, and Walcyzk and Hall (1991) conducted a quiz to ensure grade 2 and grade 4 children had the knowledge necessary to find errors. In each of these cases, error detection rates were far short of perfect, even though the required knowledge was available. Thus, relevant general knowledge is a necessary, but not sufficient prerequisite for error detection.

2. Turning now to recall failure, Vosniadou et al (1988) found that this factor was the main reason for not reporting errors; children did not remember the problematic premises after reading the passages, and so did not have the necessary information in memory to report an error. Vosniadou et al relate this recall failure to insufficient content knowledge, which made the incoming information difficult to represent in memory. In Walczyk and Hall (1991), in 65% of failures to report an error there was also failure to recall problematic information. In contrast, in Markman (1979) children could recall passages without significant errors or omissions and yet still failed to report errors. Zabrucky and Ratner (1986) found no consistent relationship between memory for relevant passage information and error detection among grade 3 children, although there was a strong relationship among grade 6 children. Garner and Anderson (1982) found that 7 of the 10 readers who failed to detect an internal inconsistency could recall both pieces of contradictory information. Most recall protocols included all the main points, showing no evidence of glossing over details or distortions. So, as suggested in Winograd and Johnston (1982), recall failure can be a reason for not reporting errors, but again does not apply in all circumstances. Nevertheless, on theoretical grounds one would expect a relationship between recall and error detection, since the integrative processing which should enhance retrievability should also provide more opportunity to detect inconsistency.

3. Winograd and Johnston (1982) in their study of grade 6 children found one particularly marked example of the use of repair strategies to make an apparently anomalous sentence meaningful. In a passage with a church

theme, the sentence: *The teacher said it was time for us to put the toys away* was intended to be anomalous. Children performed poorly in reporting this anomaly, apparently because some re-interpreted the context as being Sunday School.

Baker (1979) also notes that some of her adult subjects made inferences to resolve confusions. August et al (1984) report that half the children who failed to report a missing page made inferences to resolve the problem. On the other hand, repair strategies did not account for failure to report problems in Vosniadou et al (1988) or Markman (1979). In the latter case, children could accurately recall contradictory prepositions without alteration or addition and still fail to notice a contradiction. Zabrucky and Ratner (1986) also report very few repair inferences.

It appears, then, that readers do at times make inferences to resolve problems, and therefore the failure to report the problems does not necessarily indicate a failure to monitor comprehension. In other studies, however, such inferences cannot account for error detection failure.

4. Readers may not have the logical capacity to compare two statements and detect an inconsistency. Baker (1984b, 1985b) in relation to children and adults respectively, points to deficits in logical reasoning ability as a factor which limits the application of standards of consistency.

   These deficits may reflect inability to construct logical relationships, possibly also related to insufficient content knowledge. For example, in Markman (1979) children would need to reason: one needs light to see colour, there is very little light at the bottom of the ocean, therefore fish at the bottom

of the ocean cannot see colour. However, the deficits may also reflect processing limitations, that is, readers do not have sufficient working memory capacity to deal with all the components of the reasoning chain. Markman and Gorin (1981, p. 320) describe the relationship between this reasoning and error detection as follows:

> *"...in order for children to detect inconsistencies they must encode and store the relevant information, draw the needed inferences, retrieve and maintain the incompatible propositions in working memory, and compare them. Failing to do this, children can be left unaware that there is any problem."*

There is some evidence to show that, as processing demands increase, error detection performance decreases. For example, in Vosniadou et al (1988) children in grade 3 detected more errors when listening to rather than reading a passage, presumably because processing demands are reduced when the need for print decoding is removed. This difference due to input modality, however, was not found amongst adults in Yussen and Smith (1990), perhaps because in adults decoding is more automatic and hence makes minimal processing demands.

In Garner and Kraus (1981/2) intra-sentence contradictions were easier to detect than inter-sentence contradictions. In the latter case, more processing, including a sentence wrap-up episode intervened between the two pieces of contradictory information, thus possibly reducing the capacity available to detect the error. Strangely, Zabrucky and Ratner (1986) found no difference in error detection among either grade 6 or grade 3 children when

contradictory pieces were either in adjacent sentences or separated by four sentences. However, some children who failed to detect the error initially could do so when the pieces of information were presented together in one sentence, so Zabrucky and Ratner conclude that distance does have some effect on error detection, presumably because increased distance increases the processing resources required.

Both Tunmer et al (1983) with young primary school children and Epstein et al (1984) with high school students found that explicit contradictions were easier to detect than implicit ones. In Markman (1979), errors were also detected more readily in the explicit than in the implicit conditions. In the case of the latter, an inference was required before a contradiction could become apparent. For grade 6, but not grade 3 children, the difference between conditions disappeared when they were informed that errors were present. In other words, the older children were capable of carrying out the processing required to detect the errors, although they did not do so spontaneously. As Markman (1979, p. 653-4) states of the relationship between information processing and comprehension monitoring:

> *"Monitoring for consistency is not a single unified act. It is composed of a variety of subprocesses which need to be appropriately organized. As with other cognitive tasks, developmental differences could occur for a multitude of reasons. One or more of the subprocesses may not exist in the child's repertoire. ... Though children may be able to carry out any single subprocess, there may be limitations on how many processes they can*

> *handle. Alternatively, children may be able to cope with the entire set of required processes, but may not spontaneously organize the behaviour themselves."*

It is possible, then, that readers fail to detect errors because they are not engaging in the type of constructive processing necessary for errors to become apparent. As Markman (1979) notes, readers may be unable to carry out the necessary processing, or they may simply not do so spontaneously.

Epstein, Glenberg and Bradley (1984) show that factors which affect the likelihood of two premises being coactivated in working memory also affect the likelihood of error detection and, in particular, the illusion of knowing (IK), defined as failing to find an internal inconsistency but having a high level of confidence in one's comprehension. Among high school chemistry students, IK was higher when contradictory pieces of information were spread over more than one paragraph rather than being in the same paragraph, when one of the contradictory premises was implicit rather than explicit, and when the second premise used different wording rather than the same wording as the first. In the condition combining these three factors, that is an implicit first premise, a paraphrased second premise, and contradictory information spread over more than one paragraph, 47% of subjects showed IK. Thus, factors which decreased the likelihood of the two contradictory premises being activated concurrently in working memory also decreased problem awareness and increased illusion of knowing. Markman (1979) also shows that coactivation is a necessary condition for detection of internal inconsistencies, but is not necessarily sufficient.

The evidence as to whether differences in the extent of constructive processing can account for differences in error detection performance is mixed. Harris et al (1981) observed similar increases in reading time for anomalous sentences among grade 3 and grade 6 children. The authors conclude from this observation that both age groups were striving to integrate new information with their existing text representation, although the grade 6 children found more errors. Zabrucky and Ratner (1986) also observed similar increases in reading time for inconsistent information among grade 6 and grade 3 children. On the other hand, August et al (1984), using longer narratives and children in grade 5, observed greater increases in reading time following a missing page for skilled than for less skilled readers, with skilled readers also finding more errors. In this case, then, the skilled readers appeared to be making greater attempts to integrate information. Grabe et al (1991) produced mixed findings; the gaze durations of adults and children in grade 7 increased to a similar extent following an internal inconsistency, but only the gaze durations of adults increased following a falsehood. Thus, the children did not appear to be checking incoming information against existing general knowledge to the same extent as adults, a difference which was also reflected in poorer error detection performance on the part of the children. However, they integrated successive sentences in the text to a similar extent to adults and detected a comparable number of internal inconsistencies.

So, although the Harris et al (1981) study does not follow this pattern, there is some evidence that the extent of processing, in particular the integration of new information with the existing text representation, has a strong in-

fluence on error detection. In terms of the Winograd and Johnston (1982) critique, then, differences in error detection need to be examined from the point of view of determining whether the differences are due to processing limitations rather than inadequate monitoring per se. Inadequate monitoring can only really be said to be the cause of failure to report errors if the information necessary to detect the error is adequately represented in memory in an integrated form.

5. Finally, Winograd and Johnston (1982) suggest that some readers may adopt reading goals which limit the likelihood of detecting certain types of errors. For example, in Baker (1979) the adult subjects reported reading for gist understanding. The detection of problems embedded in main idea sentences was much higher than for problems embedded in detail sentences. Presumably these readers could grasp the main idea of a paragraph and feel satisfied with their comprehension given the goal of gist understanding, even though one of the detail sentences contradicted the main idea. However, because the problematic detail sentences did not interfere with the goal of gist understanding, they were not reported. Reis and Spekman (1983) propose a similar explanation for the fact that poor readers in grades 6 and 7 failed to benefit from specific instructions on how to find internal inconsistencies. Schommer and Surber (1986) found that the illusion of knowing, a misunderstanding of the gist of a passage coupled with high confidence in comprehension, is more likely when a text is processed only in a shallow way, focusing on individual propositions rather than the construction of a macrostructure. Although no error detection was involved here, these findings support the general point that the reader's goals and

choice of processing strategies influence the level of problem awareness.

Baker (1984b) suggests that many younger and poorer readers see the goal of reading as understanding the words, a view supported by the interviews of Garner and Kraus (1981/2). Such children may monitor their understanding at a lexical level, but fail to detect internal inconsistencies, since their attention is not focused at this level. The point made by Winograd and Johnston is that it is not possible to state on the basis of failure to report errors that readers are not monitoring their comprehension. They may be monitoring, but using different standards of evaluation to the one(s) expected by the experimenter. They may be quite capable of applying these expected standards if asked to do so. Winograd and Johnston's point is a valid one, but nevertheless it should be a matter of educational concern if readers are adopting reading goals which lead them to overlook problems of coherence in a text, especially in expository texts from which students are expected to learn. Epstein, Glenberg and Bradley (1984, p. 355) make a similar point in defending the validity of the error detection paradigm:

> *"... ability to detect a contradiction is very close to what we normally mean by comprehension. Authors intend readers to understand main ideas presented in texts. If, however, a reader cannot detect a contradiction of a main idea (inserted into the text by an experimenter), then it is unlikely that the reader has understood the text as the author intended."*

At the beginning of this section, Brown (1984b) was cited, suggesting that comprehension monitoring is an important aspect of skilled reading because it is

necessary for the detection of comprehension problems, which in turn is necessary for the regulation and repair of such problems. It was then shown that, in general, older and more proficient readers detect more errors in texts, with error detection being used as a measure of the extent to which readers are monitoring their comprehension. However, some criticisms of the error detection paradigm, such as those made by Winograd and Johnston (1982), were discussed and found to have some validity. Failure to report errors may be due, not to inadequate monitoring, but to inadequate subject knowledge, the use of repair strategies, failure to recall relevant information, processing limitations, and reading goals incompatible with finding certain types of errors. This raises the question of what conclusions can be reached about differences in comprehension monitoring between skilled and less skilled readers.

Garner and Reis (1981) showed by means other than error detection that good readers are more likely than poor readers to detect comprehension difficulties and to take appropriate remedial action. Pressley et al (1990) showed that adult readers were not particularly skilled at judging incorrect responses to thematic questions. Thus, although error detection needs to be viewed cautiously as a measure of comprehension monitoring, these two studies confirm earlier conclusions, namely, that skilled readers are better than less skilled readers at evaluating their comprehension, and that the comprehension monitoring skill of adults is still short of perfect. In response to this latter observation, Baker (1979, p. 366) points out:

> *"Comprehension monitoring does not seem to be an ability that automatically develops with maturity but is, instead, highly dependent on knowledge and expertise. Since students do not typically receive*

formal instruction in comprehension monitoring, their skills may not be well developed by the time they reach college. This is one possible reason why many college students experience difficulty in the task of learning by reading."

Pressley et al (1990) also argue that comprehension monitoring is a skill that should be taught, especially since in their study there was no relationship between reading skill and the ability to evaluate correctness of responses to questions. In other words, skilled adult readers were no more able than their less skilled peers to judge whether a response to a comprehension question was correct or not. Pressley et al (1990, p. 247) comment:

"The finding that there is no relation between ability to read and ability to evaluate the validity of general conclusions drawn from text has implications beyond having possibly identified a serious deficiency in monitoring. After all, if even good adult readers cannot recognize their misconceptions of a text, as seems to be the case, it is unlikely that readers can 'outgrow' this deficiency with age or overcome it by improving their general reading skill."

In summary, it was shown in the first part of this section that younger and less skilled readers sometimes set limited and inappropriate goals for reading, placing too much emphasis on word level processing at the expense of integration and understanding of main ideas. Such readers at times also demonstrate inadequate strategy knowledge, especially in relation to dealing with processing difficulties. However, it was pointed out that inadequate metacognitive knowledge about reading cannot account for all the data; in some cases the application of available

knowledge may be limited by the processing demands of the text, for example, its complexity and unfamiliar content. Strategy development and application may also be limited by motivational factors, in particular a self-defeating attributional style among underachieving readers.

In the last part of this section it was shown that older and more skilled readers are more likely to detect errors inserted into texts, and hence show greater evidence of comprehension monitoring. However, several reasons for not reporting errors, apart from inadequate monitoring, were also discussed, including insufficient background knowledge and inadequate constructive processing of the text, possibly due to capacity limitations or shallow processing goals.

The ability to monitor comprehension is an important one when learning from expository texts. Certainly such texts do not normally contain errors, but readers do have lapses of concentration, make incorrect interpretations, or have difficulty integrating and organising information. If readers are to understand a text as intended by the author, then it is important that they be aware of such problems in order to take remedial action.

Metacognitive knowledge and skill is an area which distinguishes skilled and less skilled readers, and therefore would appear to be a worthwhile area to target in an instructional programme. Some aspects are likely to be of greater benefit to poor readers, such as establishing more appropriate goals for reading, adopting effective strategies to deal with difficulties, and understanding the importance of effort and strategic behaviour in bringing about successful outcomes. However, training in comprehension monitoring is likely to benefit all children, since the performance of even adults is still short of perfect in this area. How such skills have been taught in training studies will be the subject of the next chapter.

2.6 Conclusion

This chapter focused on aspects of processing which were identified in the first chapter as important components of reading, with the intention of examining whether these skills distinguished more and less advanced readers. In each of these areas, more advanced readers in terms of both age and skill were found to perform significantly better than less advanced readers. Therefore, potentially at least, skill differences in the four areas examined could contribute to the comprehension differences between more and less skilled readers, and so could be worthwhile targets for a training programme.

Many studies show that more advanced readers have more efficient word recognition skills, which some theorists consider to be essential for effective comprehension (Perfetti, 1984; Stanovich, 1980). Speed of word recognition is one aspect of the processing efficiency influencing working memory capacity and measured by the Reading Span test (Daneman and Carpenter, 1980). Readers with large spans, and hence greater working memory capacity, have been shown to be more efficient at syntactic analysis (Carpenter and Just, 1989), pronoun assignment and title generation (Daneman and Carpenter, 1980), and recovery from inconsistent interpretations (Daneman and Carpenter, 1983). However, Walczyk (1990) questions whether word recognition continues to make a significant contribution to the explanation of comprehension differences once a certain threshold of skill is reached. Beyond this level, word recognition may make only a small unique contribution to overall processing efficiency.

While adequate decoding is undoubtedly a prerequisite for effective comprehension, evidence was presented in this chapter that good readers, even adults in

tertiary study, do not perform perfectly in tasks involving macrostructure formation and error detection (for example, Williams et al, 1981; Hiebert and Englert, 1983; Brown and Day, 1983; Grabe et al, 1991; Baker and Anderson, 1982; Baker, 1985b). Thus, in spite of presumably adequate decoding skills, these adults still had room for improvement in higher-level comprehension skills.

The intention of this study is to improve comprehension of expository text for all children, in this case a mixed ability grade 5 class. Such a class is likely to contain at least some children who are very proficient decoders, and for whom word-level instruction is unlikely to be the most effective way to improve higher-level comprehension processes. In the absence of evidence to the contrary, it will be assumed that all children in such a class are sufficiently skilled decoders to be able to benefit from strategically-oriented comprehension training. Therefore, the possibility of word-level training will not be pursued further in the next chapter. However, it is important to identify any children with below average decoding skills, who may not be in a position to benefit from training. This study will also examine whether there is any association between decoding skill and the ability to benefit from comprehension training.

Bridging inferences are an important component of constructing a coherent text representation (Haviland and Clark, 1974). It was shown that more advanced readers are more likely to use their general knowledge to infer information not stated in a text (for example, Holmes, 1987; Bransford et al, 1982; Oakhill, 1984; Johnson and Smith, 1981). However, in the first chapter it was pointed out that readers cannot and do not construct all possible inferences; those that are of particular importance are ones necessary to explicitly connect separate ideas within a text (McKoon and Ratcliff, 1992). Therefore, to improve comprehension read-

ers need to be able to judge when a text is incomplete because some information is left implicit, to activate relevant background knowledge, and to integrate such knowledge with textual information to create the necessary connections. These skills could be a potentially useful component of a training programme.

Readers who have fully understood a text should be able to state in a condensed form 'what the text is about', that is, to produce a macrostructure (van Dijk and Kintsch, 1983). It was shown that more advanced readers perform better on tasks which depend on such global understanding than less advanced readers. Such tasks include selecting or generating titles (Williams et al, 1981; Bridge et al, 1984), determining the importance of ideas (Brown and Smiley, 1977; 1978; Winograd, 1984), and making use of the text's structure to organise recall (Meyer et al, 1980). In forming a macrostructure, readers use many sources of information, including the explicit signalling in a text, such as titles, topic sentences and lexical pointers, as well as their knowledge of the topic to construct topic sentences and global relationships (van Dijk and Kintsch, 1983). Teaching children to make better use of explicit signalling and to invent their own summary statements to capture main ideas could be important aspects of a comprehension training programme, especially since even the more advanced readers often do not perform perfectly in this area.

In the final part of this chapter it was shown that less advanced readers often have less metacognitive knowledge about reading and may hold beliefs which inhibit their learning. Less advanced readers may see the goal of reading as decoding fluently rather than the construction of meaning, and may skip difficult parts rather than employing repair strategies (Garner and Kraus, 1981-2). Underachieving readers may have low academic self-esteem and may place too little

emphasis on the role of effort in bringing about success or failure (Butkowsky and Willows, 1980; Carr et al, 1991). Less advanced readers also monitor their comprehension less effectively than more advanced readers (for example, Vosniadou et al, 1988; Baker, 1984b), although even advanced readers do not perform perfectly, as indicated previously. Thus, metacognitive awareness and regulation appear to be important areas to target in a training programme. However, it has been emphasised that metacognitive knowledge and regulation cannot be separated from cognitive processing; readers who have difficulty integrating successive ideas are unlikely to detect internal inconsistencies, even if they have a high level of metacognitive knowledge (Baker, 1985b). Strategies encompassed under Flavell's (1981) category of cognitive actions may be particularly useful here. Such strategies are designed to assist readers to construct meaning from the text, but in the process also yield metacognitive information, that is, how well comprehension is proceeding. It will also be important to convince children that such strategies are worth using, and that the additional effort involved will produce better performance (Paris and Oka, 1986a).

Having identified important skills which distinguish more and less advanced readers, the next chapter will examine the extent to which these skills are amenable to instruction and succeed in improving comprehension of expository texts in particular.

Chapter 3

Instructional Programmes to Improve Comprehension

3.1 Introduction

Dole, Duffy, Roehler and Pearson (1991) discuss reading instruction under two broad headings: *What* should be taught? and *How* should it be taught? With regard to the first heading, Dole et al put forward three criteria for determining what should be taught. First, the chosen skills should be identified as important components of reading within a cognitive model. Second, the chosen skills should differentiate good and poor readers. Finally, the skills should be amenable to direct instruction.

The first two criteria were addressed in chapters one and two, and strategies aimed at encouraging the use of prior knowledge to construct important inferences, the formation of a macrostructure, and the application of metacognitive knowledge to monitor and regulate comprehension were identified as possible can-

didates for instruction. This chapter will now examine the extent to which these skills are amenable to instruction and succeed in improving comprehension.

With regard to **how** reading should be taught, Dole, Duffy et al (1991) are critical of the traditional methods of teaching. They point out that in the past instruction was based on the drill-and-practice model, that is *"repeatedly exposing students to tasks such as answering comprehension questions and completing skill exercises until they have achieved mastery"* (p. 249). According to Pearson (1984), this traditional approach results in the rich getting richer and the poor getting poorer. He explains his objections as follows:

> *"The prevailing wisdom concerning comprehension instruction seems to be that if students get enough exposure to a skill or kind of question, they will eventually improve at it. While such a position may be consistent with the engaged-time-on-task argument derived from Rosenshine and Stevens (1984) review, it is not consistent with arguments emanating from the direct instruction or grouping findings. Furthermore, simply on common-sense grounds there is something suspicious about a position whose implicit rationale is that if children have trouble with X, what they need is to practice X more often. Such a position probably works fine for students who can perform the task at a moderate error rate: however, for students who hover near chance level on the task, the additional practice may only reinforce their already misguided strategies."* (p. 226)

Pearson's comments could also be applied to the practice of encouraging students to do more reading of their own choice, for example, by having silent reading periods or personal reading contracts, as has been done in some Australian

schools. Such additional practice may be beneficial to students who are already moderately skilled, rather than those who are poor readers and need further skill instruction in order to make significant progress. Another limiting factor is that most of the reading done in this way is likely to involve fiction read for pleasure, so students are not being prepared for learning from expository text.

Armbruster, Anderson, Armstrong, Wise, Janish and Meyer (1991) refer to evidence from the National Assessment of Educational Progress that American high school students have difficulty learning from and understanding expository text. They hypothesised that this difficulty related to lack of familiarity with the genre due to limited exposure in the primary school years and to instruction which did not emphasise conceptual understanding, hypotheses which were supported by their observations in grade 4 classrooms. In social studies and science lessons there was very little reading of expository texts, no explicit instruction as to how to read and learn from such texts, and not even any encouragement to use strategies which have been shown to be effective, such as periodic summarisation of text segments. When reading did occur, students took turns at reading aloud, and the questions asked by teachers did not encourage integration of segments of texts or the development of deeper conceptual understanding. As Armbruster et al indicate, under these circumstances children may see reading as a performing art with the goal of reading fluently to the class, rather than as the construction of meaning to expand knowledge.

Lorch, Lorch and Klusewitz (1993) found that the college students in their study made clear distinctions between different types of reading and the cognitive demands of these various tasks. The types of reading identified by these students fell into two clusters, those of school reading and personal choice reading. While

there were some differences within these clusters regarding perceived cognitive demands, the two clusters overall differed in the following ways:

> *"As compared with personal choice reading, school reading is perceived as involving less enjoyment, emotion and interest; slower reading, more thinking, memorization and testing of understanding; less anticipation of what is ahead in the text; more use of supports and greater concentration but also somewhat greater distractibility; reading of more of the text; less visualization; more attempts to interrelate text information; more critical analysis of content; more rereading and attention to both major points and details; more attempts to relate the material to background knowledge; and more attempts to apply the reading to specific tasks or problems. In short, school reading is less enjoyable and much more cognitively demanding than reading by personal choice."* (p. 246)

The instruction described by Armbruster et al (1991) does not prepare primary students for this type of school reading. Neither does the drill-and-practice model criticised by Pearson (1984) or the related approach of encouraging students to read greater volumes of personal choice texts. Instead, Pearson proposes a direct instruction model which involves explicit explanation and modelling of a reading skill, guided practice and feedback, and independent practice with some provision for self-monitoring. In this way, the academically engaged time, which Rosenshine and Stevens (1984) identified as an important variable related to reading achievement, is productively spent because even weaker readers know what skills they should be practising. Book, Duffy, Roehler, Meloth and Vavrus (1985, p. 30) explain the steps in the direct instruction model in more detail as follows:

"...(1) an introduction of the skill to be taught (what, how, and why), (2) an explanation of the skill which includes the thinking process modelled aloud by the teacher, (3) teacher interaction with students during which (a) students are given a chance to practise the skill and to explain their thinking of the process and (b) the teacher corrects mistakes and tries to get students to think through the use of the skill on their own, (4) practice in using the skill (e.g. on worksheets), and (5) application of the skill in connected text (e.g. books)."

Book et al (1985) demonstrate that grade 5 teachers can be trained to deliver instruction from basal readers within this explanation model, and that student awareness of what they are being taught, why it is important and how to go about their assigned tasks is increased as a result. Duffy, Roehler, Meloth, Vavrus, Book, Putnam and Wesselman (1986) produced a similar result in low reading groups in grade 5. Although teacher explicitness and student awareness increased following training, there was no significant effect on reading gains as measured by the comprehension subtest of the Gates-MacGinitie Reading Test. However, Duffy, Roehler, Sivan, Rackliffe, Book, Meloth, Vavrus, Wesselman, Putnam and Bassiri (1987) did find greater improvements in reading achievement in grade 3 as a result of greater teacher explicitness and student awareness according to the Michigan Educational Assessment Program reading section and the Stanford Achievement Test word study subtest, but not the comprehension subtest.

In contrast to the approach based on a belief that children's performance will improve with greater exposure to a task, be it drill exercises or personal choice reading, the direct instruction approach emphasises the need for greater teacher input in terms of explicit explanations, modelling and guidance of skill applica-

tion. Although direct instruction is certainly an improvement on the traditional approach, Paris, Wixson and Palincsar (1986) point out that it still gives little attention to how students regulate and control their own reading. To encourage such cognitive self-management, instruction must also incorporate a metacognitive component. Within this metacognitive component, Paris, Lipson and Wixson (1983) make a useful distinction between declarative, procedural and conditional knowledge. Declarative knowledge is propositional in nature and includes beliefs about material, task and personal characteristics. Procedural knowledge involves a repertoire of actions rather than propositions, that is, knowing how to execute various skills. Conditional knowledge involves knowing when and why to apply various actions; it allows readers to make appropriate use of the available declarative and procedural knowledge in particular contexts.

O'Sullivan and Pressley (1984) demonstrate the importance of these three aspects in the context of a simple memory experiment, where children in grades 5 and 6 learned word associations by the key-word method. One group received strategy instruction, that is, explanation of the strategy and how to use it, another group received elaborated strategy instruction, involving the same as the first group plus explanation and demonstration of when the strategy was useful, while the control group received only the practice and test materials without specific instruction. Strategy instruction improved recall of both the practice and transfer lists relative to the control group, but elaborated strategy instruction was more effective than strategy instruction alone. While children needed the most explicit form of instruction, adults performed equally well with both types of instruction, suggesting that these more cognitively sophisticated subjects could infer their own conditional knowledge as a result of using the strategy.

As has been pointed out in the previous two chapters, motivation is an important component of strategic behaviour. Paris et al (1983, p. 304) list as follows some factors influencing whether or not readers will behave strategically: *"... agents' choices of actions and goals are influenced by individual styles such as risk taking, achievement aspirations, self-concept, or fear of failure"*. The beliefs held by individuals as part of their declarative knowledge about their capabilities and the value of achievement outcomes could influence whether or not they are prepared to invest the effort necessary to apply strategies. Paris and Oka (1986a, p. 104) make this point in a more general way when they state:

> *"Children must be convinced that the strategies are important, worth some extra effort, and instrumental for enhancing performance. These beliefs motivate students to use the strategies without prompting or help."*

Reid and Borkowski (1987) demonstrate the importance of attributional style on strategy learning and transfer among children in grades 2 to 4 who were overactive and had difficulty maintaining attention. In preparation for a series of memory tests these children received either strategy training, strategy plus self-control training, or strategy plus self-control plus attribution training. The self-control training involved encouraging children to proceed through a task methodically, guided by a series of self-statements. The attribution training involved linking specific successful and unsuccessful performances by each child to the degree of effort and strategy use involved. On the immediate posttests the attribution training group showed a higher level of strategy use, recall and strategy generalisation than the other two groups, who did not differ significantly from each

other. While the subjects in this study were specially selected hyperactive, underachieving children, normal intact classrooms may well contain some children whose strategy learning is inhibited by a self-defeating attributional style.

Studies carried out to improve reading comprehension have varied greatly in the extent to which a metacognitive component has been incorporated. Brown, Campione and Day (1981) and Brown and Palincsar (1982) distinguish between blind, informed and self-control training on the basis of how much metacognitive teaching is included in a programme. In blind training, students are taught a particular strategy and are helped and directed to apply it by the experimenter. Students are not told, however, why the strategy is important, how it improves performance, or under what circumstances it can fruitfully be applied. While use of the strategy may improve during the experiment, blind training studies generally fail to result in maintenance and generalisation of training effects.

In informed training, students are told the significance of the trained strategy and are given a rationale for its use. If students realise the benefits of a particular strategic approach, they are more likely to make unprompted use of it and hence to show maintenance and at least near transfer. In self-control training students are also taught methods for monitoring the effectiveness of their own strategy application. They are taught the executive skills carried out by the experimenter in blind training, as explained by Brown and Palincsar (1982, p. 6):

> *"Direct instruction of self-control skills is particularly important in the context of transfer. For students participating in blind training, the experimenter does the executive work, telling the learner what to do and frequently how long to do it. Self-control training can be regarded as an attempt to emulate more closely the activity of the spontaneous*

user of the strategy - the trained student is taught to produce and regulate the activity. Telling students to monitor and regulate their activities should also produce the effects sought in informed training attempts; if students monitor their own performances, they can see for themselves that performance is improving - they can provide their own information about strategy effectiveness."

In keeping with their advocacy of self-control training, Brown and Palincsar recommend that training should involve three aspects, namely, skills training to teach task-appropriate strategies, awareness training to teach the significance and range of trained strategies, and self-regulation training to enable students to orchestrate and monitor their own strategy application. The three components are particularly important as the instructed strategies become more difficult for the learner. Brown, Campione and Day (1981), for example, showed that even college students of average achievement needed the full self-regulation training package in order to master one particularly difficult aspect of summarisation, namely the invention of topic sentences.

The metacognitive component of training is important if students are to use their skills flexibly and spontaneously, if they are to generalise their learning to a range of tasks where the skills are relevant, and if they are to maintain their skills over time. To establish whether such learning has occurred, studies need to investigate spontaneous strategy usage, transfer and maintenance. A convincing training study must incorporate such testing. No evidence is provided on this point if students are explicitly directed to use their new skills immediately after training on a task very similar to ones used during training. As Paris, Lipson and Wixson (1983, p. 298) point out, both during training and posttesting re-

searchers should be encouraging strategic rather than purely obedient behaviour. *"Imitations of skilled behaviour in order to be obedient are vastly different from the self-directed use of the same behaviour to enhance learning. This distinction is critical for analyses of learning because many studies have shown that compliance has little enduring effect on children's learning and generalisation."*

In tandem with this metacognitive emphasis, a good training programme must not loose sight of the importance of the trained strategies themselves, the cognitive aspect. As Capelli and Markman (1982, p. 95) note, the monitoring of comprehension, a metacognitive aspect, is closely related to and arises from thorough application of cognitive processes:

> *"To understand something, one must construct a coherent representation by integrating the material, interpreting it, and drawing inferences to bridge gaps. The successful execution of the processes indicates that comprehension is proceeding smoothly; the inability to carry them out signals comprehension difficulty."*

It makes sense, therefore, that strategies chosen as the basis for a training programme should play both a comprehension fostering and comprehension monitoring role, as advocated by Palincsar (for example, Palincsar and Brown, 1984) and described by Flavell (1981) within his category of cognitive actions.

The final point in this overview section concerns the choice of materials to be used in training studies. Rabinowitz (1988, p. 234) has noted:

> *"Strategies are never applied in isolation. They are always applied to specific materials. The ease of accessing the relevant conceptual*

knowledge can subtly influence students' perception of a strategy and, as a consequence, can influence subsequent use of the strategy."

Rabinowitz (1988) used a categorisation strategy with adult subjects to aid recall of exemplars varying in typicality. He found that the strategy was applied more effectively when the material was more typical and hence more easily accessible, and was more likely to be maintained without prompting in these circumstances. With the more accessible material, subjects also considered the strategy to be more useful, less difficult and requiring less effort.

Miyake and Norman (1979) show that when a text is conceptually difficult relative to the prior knowledge of the readers, they are unable to ask many questions, even though the questions could potentially overcome the existing obstacles to comprehension. In the absence of an adequate knowledge base, readers are unable to formulate these potentially useful questions. Thus, strategy usage, in this case asking questions to overcome gaps in comprehension, is limited when the topic is too difficult.

Similarly, Capelli and Markman (1982) suggest that weaker readers will have difficulty mastering new skills when they are struggling to make sense of unfamiliar material, so *"at least in the early stages of training, it would be advantageous to use texts for which appropriate schemata are readily accessible. With more skilled readers it might be appropriate to use less familiar material in order to maintain their interest and increase the generalisability of training"* (p. 94).

When discussing training studies in the sections to follow, the extent to which metacognitive considerations are addressed will be noted, that is, whether the training was blind or also included informed, self-control components. Lysynchuk, Pressley, d'Ailly, Smith and Cake (1989) also draw attention to the need to exam-

ine the methodological adequacy of training studies, since in their review of such studies published in respected refereed journals they found none that met all their 24 criteria for internal validity and 6 criteria for external validity. They discuss internal validity under the four headings of general design, possible confounds, measurement and statistical adequacy. Under general design they include criteria such as having subjects randomly assigned to treatment and control groups, and convincing control groups that they are receiving a treatment to avoid Hawthorne effects. Under possible confounds they include not exposing control groups to the same materials as the experimental groups for the same amount of time. Under measurement they include having checks to determine whether students did as they were instructed and process measures to assess actual strategy use. Lysynchuk et al also establish 6 criteria for external validity, including the use of measures to assess maintenance of strategies over time and transfer of strategies to subjects or materials other than those encountered during training.

Lysynchuk et al's discussion is an important one to encourage high standards in training studies. However, some of their criteria are also difficult to apply. For example, when working with intact classes in schools it could be very disruptive to randomly assign children to treatment groups, possibly jeopardising the goodwill among teachers on which access to classrooms depends. While in principle the use of process measures to assess strategy use is a worthwhile suggestion, it must also be remembered that there can be a poor relationship between reported strategy use or knowledge and actual performance, as was pointed out in the first two chapters (Phifer and Glover, 1982; Kletzien, 1991; Wong and Wong, 1986; Olshavsky, 1976/7). Hare and Borchardt (1984, p. 74), following the teaching of the summarisation rules, found *"negligible and nonsignificant relationships be-*

tween stated and actual rule use." Process measures may not be as useful as they appear on face value.

However, the importance of including delayed and transfer tests was previously emphasised; without such tests it is impossible to establish whether students have acquired not only strategy knowledge but also the metastrategic knowledge necessary to apply strategies effectively (Kuhn and Phelps, 1982).

The sections to follow will examine studies aimed at improving children's ability to construct inferences, to form the macrostructure of texts, and to monitor and regulate their comprehension. Before going on, though, a brief mention will be made of what is not included.

The studies to be discussed are ones attempting to bring about a transferable improvement in reading ability, rather than improving only the comprehension and retention of one particular text. The literature on adjunct questions would be classified in the latter category. As indicated in reviews such as that by Andre (1979), many studies have investigated whether adjunct questions inserted before, after or within a text improve retention of that text. Attention has also focused on the effect of different kinds of adjunct questions, for example, verbatim and inference questions (Friedman and Rickards, 1981). Some studies, rather than asking questions, have given students specific processing directions, for example, for each paragraph to write out the key words, to paraphrase, or to generate a logical extension (Glover, Plake, Roberts, Zimmer and Palmere, 1981). In Doctorow, Wittrock and Marks (1978) children were directed to write a summary statement of what happened in each paragraph of a story. Some studies have investigated different methods of activating background knowledge before reading a text, for example, a teacher directed overview of the material about to be

read, as opposed to more student centred discussion of the topic (Dole, Valencia, Greer and Wardrop, 1991). What all these studies have in common is that they measured only retention of the text being studied.

Paris, Wixson and Palincsar (1986, p. 103) refer to the use of techniques such as adjunct questions, advance organisers, imagery and analogies as an 'activities approach', and point out the shortcomings of such approaches as follows: *"...most of the activities have been designed to facilitate learning of the content rather than learning the activities as self-controlled tactics that students can use. This may be one reason for the limited generalization found in instructional research on these activities."*

The studies mentioned above succeeded in improving learning of content, and potentially could even improve the reading strategies available to students when processing novel texts. However, such transfer was not encouraged or measured. All the studies to be discussed in this chapter at least attempted to bring about a generalised improvement in comprehension.

3.2 Encouraging the integration of prior knowledge and textual information

Some researchers in their training programmes have encouraged readers to engage their prior knowledge to a greater extent, in particular, to use existing prior knowledge to generate inferences when information was not explicit in texts. Among these researchers, Carr, Dewitz and Patberg (1983, p. 1-2) explain the importance of engaging background knowledge as follows:

> *"First, as the reader interacts with the information in the text, his/her prior knowledge affects his/her interpretation of the text and determines the plausibility of subsequent information. Second, the reader uses background knowledge to determine the number and types of inferences needed to comprehend text and to monitor the 'goodness of fit' between background knowledge and new information. New information is, in effect, filtered through the framework of prior knowledge.*
>
> *In this process of filtration, the reader encounters new textual information which is not fully explicit. Consequently, he/she is called upon to infer information beyond that which is contained in the text. Simultaneously, he/she must infer logical relations between propositions in the text and previous knowledge. To infer either unstated information or logical connections requires the reader to rely not just on previous knowledge alone; he/she must also use strategies to find clues in the text which promote and support inferences."*

According to these authors, the process of filling in selectively created cloze blanks mirrors the process of inference generation because syntactic and semantic information has to be combined with background knowledge in order to determine an appropriate word to insert. Hence studies by Carr, Dewitz and Patberg (1983), Dewitz, Carr and Patberg (1987), and Sampson, Valmont and van Allen (1982) used a cloze procedure to encourage students to generate more appropriate inferences while reading. On the other hand, Hansen (1981) and Hansen and Pearson (1983) used prereading discussion and predictions as well as increased levels of inference questions for this same purpose, while Gordon and Pearson (1983) taught children to generate inferences while reflecting on segments of text.

In Hansen (1981), grade 2 students reading at or above grade level were divided into three groups as they studied their normal basal readers over a 10 week period. In addition to their normal instruction, one group discussed background knowledge and predictions based on this background knowledge prior to reading. Students were told that comprehension was improved by relating new information to existing knowledge, with this point made more graphic and concrete by having children write predictions about the text on pieces of paper which were woven into other pieces representing the brain and existing knowledge. This Strategy group and the control group answered literal and inference questions after reading the text in a ratio of 5:1. A third group, the Question group, answered all inference questions after reading to observe the effect of increased practice. At the end of each week of instruction, students completed a worksheet about the text studied, which included literal, inference and strategy (prediction) questions not previously studied. For the immediate posttest, students read and answered questions about two expository texts, and read and recalled a further text.

For the last five of these worksheets, when treatment effects would have had time to become evident, both the Strategy and Questions groups answered more literal and inference questions correctly that the control group. For strategy questions, the Strategy group performed better than the Question group, which in turn performed better than the control group. However, on the posttest there were no treatment effects for the number of questions answered correctly, nor was there a treatment effect on free recall. Thus, the effects of training were not maintained even on an immediate posttest involving the same type of task as in training. Since the activation of prior knowledge and generation of predictions was directed by the teacher during the training, it may be that students did not continue to use this approach when left to their own devices in the posttest.

Hansen and Pearson (1983) used a similar procedure as in the previous study, but included students with a wider range of reading ability in grade 4. These students were either above average or slightly below average according to the Stanford Achievement Test comprehension subtest. As in the previous case, the Strategy group discussed personal experience and knowledge related to the topic of the texts being studied, then predicted how such ideas might appear in the texts. The Strategy group was asked all inference questions after reading, while the control group answered literal and inference questions in a ratio of 4:1. The worksheet at the end of each week included novel literal and inference questions. The first posttest was based on a text at each student's own reading level, while the second posttest involved a common text written at grade 2 level.

For the last five worksheets, the treatment improved the performance of weaker students for inference, but not literal questions, while amongst above average readers the experimental treatment did not produce better performance

for either type of question. On the first posttest, with the test passage written at the appropriate reading level for each child, weaker readers showed a treatment effect for both types of questions, while the better readers did not show an effect for either type of question. On the second posttest, with the test passage written at grade 2 level for all students, both above and below average readers in the experimental group scored higher on inference questions than their peers in the control group, although there was no effect for literal questions.

This treatment involved the activation of prior knowledge and its use to generate predictions about text topics, coupled with increased practice at answering inference questions. The treatment benefited primarily below average readers, who may not normally integrate their knowledge and textual information in this way. Possibly above average readers spontaneously engaged in this type of integration, and so were not assisted by the training. However, on the second posttest above average readers in the control group scored only 68% for both literal and inference questions, so there would still appear to be some room for improvement even for above average readers. In addition, there is a broader question about the value of such relatively unconstrained use of prior knowledge to generate inferences or predictions. A very vast number of ideas could be generated in this way, which may or may not be relevant to the content of the passage about to be read. It was noted previously that the instruction was very much directed by the teacher, so it is not clear how students would assume control of the strategy during individual reading.

Gordon and Pearson (1983) compared two treatments in their attempt to improve the inferencing ability of average and above average readers in grade 5. The experimental treatments were delivered for ten minutes prior to normal

basal instruction over five lessons for each of eight stories. In the Content and Structure treatment, the teacher first expanded on the theme and background knowledge associated with the actions, objects and events in each story, as well as propositions important to the overall structure. In the remaining sessions for each story, children were guided by the teacher to determine the macrostructure and to write important propositions on a diagram representing the major elements of story grammar. In the Inference Awareness treatment, the teacher guided the students to generate inferences after reading a segment and to look for clues supporting the inferences both within the text and prior knowledge. During this time the control group completed a variety of language activities which were deemed to involve a similar level of novelty and teacher interaction.

All groups completed a weekly worksheet which included three common questions and seven treatment-specific questions as additional practice and a measure of how well students were mastering the strategies. In addition, students were asked to free recall the final story. They also completed an immediate posttest and another after 16 days. Both tests involved two stories and 15 questions, one set answered with the text available and one without. Each test included an equal number of text explicit, text implicit and script implicit questions (Pearson and Johnson, 1978).

With regard to the three common questions on weekly worksheets, there was no significant treatment effect for text explicit and text implicit questions, but the Inference Awareness group answered more script implicit questions correctly than the Content and Structure group, who in turn answered more of these questions correctly than the control group. Free recall of the final training story was scored according to the number of complete and accurate main ideas included, and here

the Content and Structure group performed better than the other two, who did not differ significantly from each other. However, improved recall was limited to students who had performed moderately well on a pretest of free recall. In addition, the recall was based on a story studied during training, so it is not surprising that the group whose training focused on main ideas also recalled the most main ideas.

On the posttests there was no significant overall treatment effect. However, the experimental treatments produced superior scores when questions were answered with the text present, but not when they were answered from memory. Gordon and Pearson suggest that the treatments had taught the children to use the text more effectively to answer questions. In addition, the experimental treatments produced higher scores than the control treatment for students who had performed moderately well on similar pretests.

With regard to the aptitude by treatment interactions, Gordon and Pearson speculate that the training was sufficiently complex so that a minimal level of achievement had to be reached before it was beneficial, and also that the stories from basal readers were relatively more difficult for the less able readers *"In a situation where the lower ability students were grappling both with difficult content and novel instructional devices, the cognitive load was simply too much for them to handle"* (p. 37). However, it should be remembered that there were only average and above average readers in this study.

The studies discussed so far have shown only limited transfer of training. Perhaps this result is not surprising given what was said earlier about metacognitive aspects of training. There is no clear indication in the studies that teaching included informed or self-control components.

Carr, Dewitz and Patberg (1983) used a more constrained procedure to generate inferences than in the Hansen studies discussed earlier. In this study students in grade 6 were taught to generate appropriate words to fill in cloze blanks over a period of eight weeks. One group was taught how to look forward and backward to find clues to fill cloze blanks, created by selectively deleting content words. This group was also taught a self-monitoring checklist to test the appropriateness of inserted words. After completing each practice passage, answers were discussed with the class and the teacher asked further inference questions about the passage. Another group received the cloze training and practice, as well as a structured overview of each passage, setting out text content in a hierarchical order. The control group read the same passages but without any specific instruction. Tests were conducted at the end of each unit during training, using texts on the same social studies topics as in training. An immediate posttest and a 6 week delayed posttest used different topics.

In response to literal questions, both experimental groups exceeded the control group for all tests. The same result was obtained for inference questions, except that group differences did not appear until the second test during training, that is, four weeks after the beginning of training. Hence, training effects were transferred to different topics and maintained for 6 weeks. Effects were larger for below average than for above average readers.

Dewitz, Carr and Patberg (1987) used a similar procedure with students in grade 5, this time including a group receiving the structured overview only, without cloze training. In addition to the 6 week delayed test, there was also a test 20 weeks after the completion of training.

Broadly speaking, the cloze training, whether coupled with the structured overview or not, improved responses to both literal and inference questions. For literal questions, a treatment effect did not appear until the second test, four weeks into training, while for inference questions the effect did not appear until the third test, at the end of the 6 week training period. Also, while for literal questions the training effect was maintained for both the 6 week and the 20 week delayed tests, for inference questions the effect was maintained only for the 6 week test. Unlike in the 1983 study by these authors, there was no ability interaction. Giving students an experimenter designed overview of passage content without cloze training did not improve comprehension.

Finally, Sampson, Valmont and van Allen (1982) also used a quasi-cloze procedure to attempt to improve the comprehension of grade 3 students reading at or above grade level. In this case, though, there was no direct instruction of strategies to fill in blanks. Rather, students practised, over 27 lessons, filling in words selectively deleted according to the topic being studied in the reading class at the time. For example, if the topic was adjectives, then adjectives would be deleted from the practice passages. The teacher then discussed the appropriateness of words chosen with each group of students. The control group completed only normal basal reading activities. A quasi-cloze test was included in both the pretest and the immediate posttest, as well as alternate forms of the Gates-MacGinitie Reading Tests vocabulary and comprehension subtests. Pretest scores were used as covariates in the analysis of posttest scores.

The quasi-cloze practice resulted not only in more correct responses in a quasi-cloze posttest, but also a greater diversity of correct responses. In addition, the treatment resulted in superior scores on the comprehension subtest of the Gates-

MacGinitie test.

To summarise, the use of cloze exercises to encourage students to draw on their prior knowledge within the constraints provided by textual information seems to be an effective approach to improve inferencing skill. The studies using this method were relatively lengthy ranging from 20 lessons (Carr et al, 1983) to 27 lessons (Sampson et al (1982). Dewitz et al (1987) indicate that ample opportunity to practice is essential, since training effects in their study did not become evident for 4 weeks in the case of literal questions and 6 weeks in the case of inference questions. These studies also included a metacognitive component in the form of questions to monitor performance. There could be several reasons why the Hansen approach of prior knowledge activation and prediction was not as effective, including the less focused nature of the predictions, the shorter duration of training (only 10 strategy lessons), and the lack of self-monitoring guidance so that the strategy could be effectively applied without teacher assistance.

The cloze method, when used as in these studies with selectively deleted content words, has the advantage of forcing students to integrate successive ideas in a text, and to integrate textual information with prior knowledge in order to insert appropriate words. Carr et al (1983) and Dewitz et al (1987) obtained significant improvements on tasks not identical to those used in training, and the improvements were largely maintained on delayed tests. It appears, then, that in the course of training students learned to process text more thoroughly and continued to do so, even when reading texts without cloze blanks in the posttests. These studies avoided one possible problem with training focused on inference or prediction generation, and that is the problem of developing in students an awareness of when inferences are necessary, or at least likely to enhance

comprehension of the textual information. Not all studies in this section found interactions between treatment and reading ability, but where such interactions did occur they indicated that this type of treatment had a greater effect on less skilled readers.

3.3 Macrostructure formation

In the first chapter, the theoretical importance of generating a text's macrostructure and the processes used to do so were discussed. The second chapter showed that more advanced readers in terms of both age and reading skill are more adept at tasks involving the use or construction of macrorelevant information, tasks such as selecting or inventing titles and topic sentences, constructing coherent paragraphs, deciding whether or not sentences are consistent with the paragraph topic, and summarising. However, even more advanced readers do not perform perfectly in these areas, so teaching strategies to improve macrostructure formation to an intact class in upper primary school could potentially enhance the global comprehension of all children. As Pressley, Johnson, Symons, McGoldrick and Kurita (1989, p. 24) in their discussion of effective strategy instruction point out, comprehension improvement is not an issue relevant only to poor readers. *"Strategy instruction is not simply remediation. Rather it is an attempt to develop sophisticated comprehension skills in all students, skills that are generally more effective than ones that even good child readers use on their own."*

Being able to understand important ideas and their interrelationship in the overall structure is an important skill in learning from instructional text. It is a prerequisite for effective study, as shown for example by Brown and Smiley (1977; 1978). It is also vital when students consult reference books in preparation for projects or essays. Yet it is a skill not well mastered by upper primary school children, as pointed out by B. Taylor (1986, p. 197).

> *"Many middle grade students form almost no mental gist, even of unimportant information, after reading several pages in a textbook.*

> *Consequently, these students are unable to write or tell much at all when asked to summarize the important ideas of what they have just read. Many other middle grade students form a gist comprising unconnected ideas and unimportant details as opposed to important information."*

Taylor's comment is supported by her observation of students in grade 6 reading at or above grade level. When asked to summarise three or four pages from a grade-level social studies text, these students produced very few important ideas. In other words, they had not retained the gist of their reading. It would appear, then, that even relatively skilled readers in this age group could benefit from training which improves their ability to determine macropropositions. As pointed out by Meyer (1984, p. 136), such training could take a variety of forms.

> *"The ability to identify a text's top-level structure will facilitate their (students') identification and memory for macropropositions. Generating summaries, brief outlines, abstracts and titles for text should increase their ability to identify macropropositions and to construct them when they are not explicitly stated in the text."*

The diversity of possible strategies is also acknowledged by Pressley et al (1989) when they refer to summarisation as a 'family of strategies'. Students could be encouraged to focus on macrorelevant information by generating headings, topic sentences or passage outlines, or summarising texts according to Kintsch and van Dijk's (1978) macrorules.

This range of possibilities is evident in the studies to be discussed in this section. In the first group of studies to be examined students were taught to

make use of explicit text features, such as headings, subheadings and conventional text structures. In the second group students were taught to generate their own headings, topic sentences and passage outlines. In the third group students were taught to focus on similar information by asking themselves questions, while in the fourth group students were taught to use summarisation rules.

3.3.1 Using explicit text features and conventional text structures

Holley, Dansereau, Evans, Collins, Brooks and Larson (1981) and Brooks, Dansereau, Spurlin and Holley (1983) instructed undergraduate students to focus on headings and passage outlines while processing a text. In Holley et al, students in the first session received either input training or output training instructions. In input training they were directed to attend to headings and to link passage contents to headings as they read. In output training they received a passage outline, that is, a hierarchical listing of all the headings, and were directed to use the outline to help them recall details and to memorise the outline. They then applied these directions to the study of a practice passage containing headings. Two further groups were told only to use their normal study technique. One group read the practice passage with headings, the other read the text without headings. For these groups headings were similarly present or absent in the test passages.

In the second session students read and recalled a longer text. In the third session they read a second text, which they recalled five days later. The training groups did not differ significantly from each other in terms of recall scores, nor

did they differ from the headings-no training group. Hence scores for the three groups reading texts with headings were collapsed and found to be significantly higher than those for the no-headings group, especially on the delayed test. The presence of headings clearly increased recall, although the training instructions conferred no additional benefit. Holley et al suggest that one practice session was insufficient time for students to adapt their normal study strategy to the directions.

Brooks et al (1983) conducted a similar study with undergraduate students. In the first experiment students received no specific instructions, but read texts either with an outline only, headings only, an outline and headings, or neither. The three posttests involved writing an essay incorporating an organised summary of the passage, creating an outline of the passage, and answering a set of multiple-choice questions. For one text the tests were taken immediately, while for the other tests were taken after a delay of two days. There was no significant treatment effect on the immediate tests, but the headings-only group outperformed the others for all three delayed tests.

In a second experiment two groups read texts with headings, while a control group read the texts without headings. One of the groups with headings was told only to use normal study techniques, while the other was instructed to use the headings to develop expectations about the content to follow, to determine how each heading was related to its section, to memorise the headings, and to use them as retrieval cues. All groups then studied a practice passage, and two days later studied the test passage after being reminded of instructions. Five days later they took the three tests. The trained group performed significantly better on the essay test than the control group, which did not differ significantly from

the headings-only group. Results were in the same direction for the outline test, but did not reach significance, probably due to the large variance in scores in the trained group. There were no significant differences in the multiple-choice tests.

Thus, the experiment by Holley et al (1981) and the first experiment by Brooks et al (1983) suggest that adults are able to make use of headings to improve recall of macrorelevant information even without training, as evidenced by the fact that headings-only groups performed better than control groups reading texts without headings, and as well as the headings-plus instructions groups. However, in the second experiment by Brooks et al headings improved recall only when students were directed to attend to them.

In Taylor (1982) classroom teachers taught children in grade 5 to write skeleton outlines of passages by first numbering each heading, subheading and paragraph, and then writing their own statement about each one, incorporating its gist. Students practised this technique once a week for seven weeks, while the control group received traditional instruction based around pre-reading discussion of background knowledge and answering questions set by the teacher after reading. One week after the conclusion of instruction students read the first test passage, and two weeks later the second. In both instances they were instructed to continue using the study strategy, that is, either outline generation or question answering. They then recalled each passage and answered a set of short-answer questions.

In both tests in this first experiment, the outline generation group recalled more ideas than the traditional instruction group, and in a more organised fashion, although there was no significant difference between the groups on the short-answer question test. There was no interaction between reading ability and treat-

ment. For the two tests, 55% and 84% of the information recalled in the two tests respectively had been included in the passage outlines generated by the experimental group, so recall was based on, but not limited to information which had been summarised. However, in both recall tests in a second experiment there were no significant treatment effects, and the traditional instruction group answered more questions correctly than the outline generation group. Taylor examined the outlines and answers produced for the test passages and found that while the groups in both experiments answered an average of 79% questions correctly, the outlines produced by the experimental group was 60% correct in the first experiment, compared to only 43% in the second experiment. Thus, the experimental group in the second experiment had not mastered the outlining technique sufficiently in the seven hours of instruction.

Taylor and Beach (1984) conducted a similar experiment with students in grade 7. In this case an untreated control group was included, which completed only the pre- and posttests. The pretests involving free recall and short-answer questions revealed no significant differences between the groups. For the posttest, administered the week after training, students studied one of two passages. Questionnaire responses revealed that the topic of one of these passages was considerably more familiar to the children than the other. For the more familiar passage, both the passage outline group and the conventional instruction group recalled more than the untreated control, but did not differ significantly from each other. For the less familiar passage, the outline group outperformed both the conventional instruction and the control groups. With regard to question-answering, the two instructed groups did not differ from each other, but both performed better than the untreated control. Thus, the outlining treatment was particularly bene-

ficial when the topic was less familiar and in all instances was at least as effective as conventional instruction.

Although these two studies suggest that the outlining procedure is a useful one, it should be noted that there was no test of 'spontaneous' strategy use, since students were required to explicitly use their particular study technique for each test. There is also no indication that training included the informed, self-control components considered important by Brown, Campione and Day (1981). In addition, only Taylor (1982) included a delayed test to assess the maintenance of trained strategies.

Several studies have taught students to use knowledge about conventional text structures to generate outlines of texts. Brooks and Dansereau (1983) established empirically, through a survey of college students, that texts explaining scientific theories were expected to have a relatively predictable structure, including the following sections: Description, Inventor, Consequences, Evidence, Other theories and Xtra information, forming the acronym DICEOX.

College students learned these categories, practised using them as a framework to summarise passages, and used them as an aid to retrieval in six hours of instruction. The control group learned only general study support techniques. The test passage was read immediately after training and was recalled five days later. The experimental group recalled significantly more main ideas than the control group, and in a more organised form, but there was no difference between the groups in terms of recall of detailed information.

In a second experiment, training was reduced to one and a half hours, the test passage was either read in the original form or rearranged such that it no longer followed the DICEOX sequence, and a cued recall test was included as well as

free recall. Cuing involved questions reminding students of DICEOX categories, such as: 'Write a brief account of the theory's history, including the name of the inventors, important dates, and historical background information'.

In this experiment there was no significant difference between the free recall of the experimental and control groups, although the experimental group performed significantly better for both main idea and detail recall in the cued test. This result applied both to the conventional DICEOX text and the non-DICEOX version. This second experiment suggests that a lengthier training period was necessary for the experimental group to master application of the strategy such that they could use the DICEOX categories to retrieve encoded information without explicit cues.

In a similar vein, Cook and Mayer (1988) taught junior college chemistry students to recognise and use three top-level text structures in short science texts, these being generalisation (a main idea followed by supporting details), enumeration (a list of facts, often attributes of an object or process), and sequence (a connected series of events extending over time). Students were taught to recognise these structures and learned two or three steps appropriate to each to help them use the structure to summarise important points. The students practised applying these strategies over a period of about nine hours by completing strategy worksheets. The instructor ensured in each session that the worksheets were correctly completed and that each student could explain from memory the information included in the summary. During this time the control group completed normal laboratory exercises.

Pretests and immediate posttests included free recall and short-answer questions, some literal and some involving application of the information learned.

On pretests there were no significant differences between groups. However, the experimental group made significant pre-posttest improvements in free recall of high-level ideas and answers to both literal and application questions. The control group made no such improvements on any measure, and the experimental group also did not improve on recall of low-level information. While these results were very positive, Cook and Mayer note that they used only texts which clearly exemplified the three text structures, so it is not clear whether the training effects would transfer to less well organised texts.

Armbruster, Anderson and Ostertag (1987) taught a similar procedure to grade 5 students of mixed reading ability, although without students of markedly below average ability. In this case only the problem-solution text structure was used. Students were provided with a problem-solution frame, that is, boxes into which to write a short statement of the problem, the actions taken, and the results, and then to summarise the main points relating to each statement. The children were taught explicit rules for analysing texts in this way according to the direct instruction model, and were also given a very general checklist to help them assess their work. Instruction took place over eleven lessons. During this time the control group answered and discussed questions related to the training passages. Immediate posttests involved answering an essay question calling for an outline of the problem-solution information, as well as short-answer questions, the majority of which did not relate directly to problem-solution information. Students also summarised two short passages. Prior to reading test passages they were reminded to use the trained strategy.

The experimental group included a higher number of ideas relevant to the topic in their essays than the control group, and also included more important and less

unimportant information in their summaries. However, there was no difference in scores for the short-answer questions. Reading ability did not interact with treatment for the essay question, although the training was of less benefit to lower ability readers for the summarisation task.

In chapter one it was pointed out that explicit structural markers in a text are one important source of information for macrostructure formation. However, studies by Brooks, Dansereau, Holley and Spurlin (1983) and Brooks, Dansereau, Spurlin and Holley (1983) suggest that even competent adult readers do not always make effective use of explicit structural cues such as headings. In the study by Williams, Taylor and Ganger (1981), discussed in the previous chapter, adults did make use of explicit topic sentences in simple texts, although for children in grade 5 these topic sentences had to be highlighted to encourage them to make greater use of this summary information in selecting or generating paragraph titles. It would appear, then, that especially younger readers could benefit from training which focuses attention on explicit structural markers to guide abstraction of the most important information from a text. Taylor (1982) and Taylor and Beach (1984) produced higher recall of passage information among children by encouraging them to generate outlines interrelating headings, subheadings and topic sentences. Retention of important ideas has also been increased by teaching adults (Brooks and Dansereau, 1983; Cook and Mayer, 1988) and children (Armbruster et al, 1987) to use conventional text structures to select and organise important information. This kind of training clearly has the potential to improve macrostructure formation. However, it was pointed out in chapter two (for example, Winograd and Bridge, 1986) that natural texts are usually not as well structured and considerate as the materials used in these studies, and hence

may not provide all the explicit structural markers the students have been taught to attend to.

3.3.2 Generating macrorelevant information

The previous section discussed studies where readers were encouraged to use explicit structural cues such as headings, or their knowledge of conventional text structures to identify and organise information central to the topic. However, as has been noted previously, texts are not always well written; they may contain insufficient headings or not conform to 'pure' text types. Hence, readers need to be able to generate their own macrorelevant information, such as headings, topic outlines and topic sentences for paragraphs.

Brooks, Dansereau, Holley and Spurlin (1983) taught undergraduate students to generate their own set of hierarchically structured headings when none were present in the text. In one training session students were given notes in which the importance of headings was explained, the attributes of good headings were described, and guidelines for developing and using headings were given. Students then applied these guidelines to a practice text, while another group read the same text with headings but no specific instructions, and a further group read the text without headings. Prior to reading the test passage, students in the experimental group were given a review of their instructions. Immediate posttests included an essay question requiring production of an organised summary of the text, a passage outlining task, multiple-choice questions and short-answer questions.

The heading-generation group scored higher than the others on all measures, but only three of the differences reached significance, these being essay content, essay organisation, and outline content. The headings-only group was not signif-

icantly different to the no-headings group on any measure, suggesting that even for adults the mere presence of headings does not always guarantee that they will be used by readers to organise information.

In a study by Berkowitz (1986), classroom teachers taught grade 6 children a mapping technique whereby they wrote the title of a passage in the centre of a page, then skimmed the text to find four to six main ideas to write around the centre. Finally the text was examined more closely and two to four important details were written under each main idea. Students then used the map to test their recall of the passage. Lessons took place once a week over a six week period. Practice texts were taken from grade-level social studies books and contained no headings or subheadings. Three other groups also read the practice texts. One group received a map constructed by the experimenter, which was explained by the teacher and used by the children to test their recall. Another group answered questions, half related to main ideas and half to details, while the third group was instructed to reread each text and to rehearse the information. Three test passages were read immediately after training, each one tested by free recall and short-answer questions. For the first two passages students were directed to apply their instructed strategy in writing, while for the third text (transfer test) only covert strategy use was allowed. Tests for the second passage were administered again after a delay of two weeks without rereading of the text (delayed test).

Scores on both the immediate tests and the delayed recall test, but not the transfer test, were highest for the map-construction group, although only the difference for immediate recall of passage 2 reached significance. Similarly, only on the second immediate test did both the map-construction and question-answering groups answer significantly more questions correctly than the other two groups.

When treatment effects did occur, there was no interaction with reading ability.

Given these rather unimpressive results, Berkowitz determined the extent to which strategies had been mastered by examining the maps and answers to questions generated while reading test passages 1 and 2. Only students who met the criterion of 80% correct were included in the subsequent reanalysis, which included 81% of the map-construction group and 64% of the question-answering group. In this reanalysis, the map-construction group had higher total recall than the question-answering group for both immediate tests and the delayed test, but not the transfer test. There was no difference between the two groups in terms of the number of questions answered correctly. Training in this study involved only six hours, which appears not to have been sufficient for all students to master the map construction strategy. However, even those students who had reached the criterion of mastery did not show improved performance on the transfer test where they were not directed or even allowed to overtly construct maps, suggesting that the training had not significantly affected students' normal covert reading strategies. Berkowitz also notes that the treatment had a greater effect on passage 2 than passage 1 because some texts are easier to map than others. This would also appear to be a factor limiting the general usefulness of this approach.

Baumann (1984) taught children in grade 6 to generate passage outlines in eight lessons. Initially children were shown how to select the most important idea in a paragraph when a topic sentence was present, then they practised generating topic sentences for paragraphs without such sentences. Next they practised identifying an explicit summary sentence for a whole passage and generated one when it was not explicit. Finally they generated a total outline encompassing the passage summary sentence and paragraph topic sentences. Students were

taught according to the direct instruction model, and they were informed of the usefulness of the procedure and how it would help them become better readers. During this time, the basal group continued to study exercises related to main idea comprehension in their basal readers, while the control group received vocabulary training. The immediate posttests involved multiple-choice questions on identifying explicit and implicit main ideas in paragraphs and short passages, and on linking details to explicit main ideas. Students also generated passage outlines for two texts and recalled two further texts.

The experimental group performed better than the other two groups on all measures except free recall. Differences in recall scores were, however, in the same direction. There were no interactions between treatment and reading ability. The reason for the failure to find a significant training effect on free recall is not clear. However, the recall test was the only genuine transfer test; the other tasks were similar to ones the experimental group had practised during training. It may be that the children had not incorporated the outlining strategy into their normal reading, but only used it when required to produce a written product, a possibility that was also raised with regard to the transfer test in Berkowitz (1986). It may be that lengthier training, perhaps culminating in oral rather than written outlining, may be necessary to encourage children to incorporate the strategy into their normal reading.

Several studies have attempted to teach children to generate topic sentences for paragraphs. Stevens (1988) taught remedial readers in grades 6 to 11 who were reading at least two years below grade level. From children meeting these criteria those were selected who showed good literal comprehension but relatively poor main idea comprehension according to the experimenter-designed pretest.

Training involved paragraphs of varying lengths and structures on the topics of geology and weather. Some contained no topic sentences, and where topic sentences were present they occurred in various positions in the paragraph. Students learned to identify the topic of a paragraph and then what was said about the topic to formulate a summary statement. They were taught according to the direct instruction model from a computer terminal. Students were also informed of the usefulness of the strategy and were taught how to check the adequacy of their summary sentence. During this time the control group read the same paragraphs and for each answered a multiple-choice question asking for the best summary statement. The immediate posttest required students to answer one literal, one inference and one main idea question for each paragraph, with half the paragraphs on the same topics of geology and weather, and half on a variety of other topics.

On the pretest there was no significant difference between the groups. On the posttest the experimental group answered more main idea questions correctly than the control group for both the same- and different-topic paragraphs. However, there was no reliable treatment effect on inference questions, so the training was relatively specific and had not transferred to other aspects of reading. A similar study by Stevens, Slavin and Farnish (1991) using intact classes of grade 3 and 4 children also produced an improvement in responses to main idea questions, but not inference questions.

The Stevens (1988) study is one of the few in this area to incorporate the informed, self-control training advocated by Brown, Campione and Day (1981). However, there was no delayed test, so it is not clear whether training effects were maintained over time.

Kurtz and Borkowski (1987) taught children in grades 4 to 6, some of whom were classified as impulsive. Two groups were taught paragraph summarisation skills over three sessions, and were told how such summarisation would make them better readers. First the children classified words and supplied superordinate terms for groups of exemplars. Then they were taught to identify the topic of a paragraph and the most important thing that was said about the topic so that they could first select explicit topic sentences and later invent their own. One of these groups also received executive training in each session, where the instructor emphasised the importance of using appropriate strategies for a task, of working slowly and carefully, and of monitoring performance. A control group read the same practice paragraphs and selected or generated a summary sentence without instruction.

Pretests and posttests included writing summary sentences for paragraphs and the comprehension subtest of the Iowa Test of Basic Skills, the former administered two weeks after the completion of instruction, the latter a further month later. No mention was made of training strategies during the posttesting. There were no significant differences between the groups on pretests. However, on the summarisation posttest the executive training group exceeded the strategy-only group, which in turn exceeded the control group. There was no interaction between treatment and cognitive tempo. Only a subset of the children took the Iowa posttest, however here too both strategy groups exceeded the control group, so the training had a transferable effect on reading comprehension.

Borkowski, Weyhing and Carr (1988) taught paragraph summarisation skills in a similar manner to learning disabled students, aged 10 to 14 years and reading at least two years below grade level. Some students received only strategy train-

ing, while others received attribution retraining as well. In the latter treatment the instructor periodically made an error when modelling topic sentence selection or invention, which was related to not trying hard enough and not using the strategy. The error was then corrected by applying the strategy. The control group read and summarised the same paragraphs, but without strategy or attribution training.

Pretests and posttests included a paragraph summarisation task, an attribution questionnaire and the Stanford Diagnostic Reading Test. One posttest was taken the week after training, the other two weeks later. Analysis of both posttest scores using pretest scores as the covariate indicated that children who received both strategy and attribution training performed significantly better on the paragraph summarisation test than the other two groups, who did not differ significantly from each other. For these learning disabled students, then, strategy training alone was not sufficient to improve performance. Training effects were maintained on the delayed posttest. However, the groups were not significantly different on the attributions questionnaire or the Stanford test overall. More detailed analysis revealed that while treatment did not affect performance on the literal comprehension section of the test, where questions tapped detailed information, children receiving strategy plus attribution training performed significantly better than others on the inferential section, where questions tapped main idea information.

The studies in this section generally indicate that teaching students to generate macrorelevant statements, such as headings or topic sentences, improves their global comprehension. In Brooks et al (1983), adults who generated headings while reading subsequently wrote higher quality essays in terms of organisation

and content. In Kurtz and Borkowski (1987) and Borkowski et al (1988), children taught to invent topic sentences subsequently wrote better paragraph summaries and improved their performance on standardised comprehension tests, while children receiving topic sentence instruction in Stevens (1988) subsequently showed improved main idea recall. Results from the studies by Berkowitz (1986) and Baumann (1984) are less clear-cut. In Berkowitz the mapping technique, even when mastered, did not improve recall on the transfer test. In Baumann training improved the ability of children to outline passages, but not their recall. The reason for this lack of transfer is not clear, although it was suggested that some oral practice of the instructed outlining strategies at the end of the training period may encourage the children to incorporate the strategies into their normal reading, even when no written outline is required. There was no indication of ability interactions in studies including children with a range of reading abilities, and some studies succeeded in improving the comprehension of below average readers. On the basis of the studies reviewed thus far, it is suggested that teaching children to invent topic sentences, coupled with attention to explicit structural markers as discussed in the first section, are useful strategies to improve main idea comprehension among readers of various abilities.

3.3.3 Self-questioning about main ideas

As noted by MacDonald (1986), teachers have often been advised to encourage students to ask themselves questions while reading in order to improve comprehension and recall. However, there are both empirical and theoretical difficulties involved in this line of research. Davey and McBride (1986) describe research findings from instruction based on self-questioning as 'meagre and contradictory'.

Pressley et al (1989, p. 15) note that: *"The instructional worth of this strategy in elementary schools has not been evaluated as completely as would be desired."* Wong (1985) in her review of research in this area points out there are several theoretical perspectives underlying the advocacy of self-questioning, these being schema theory (questions to activate appropriate prior knowledge), metacognition (questions to regulate and monitor comprehension) and active processing (questions to shape and focus thinking about the text). Especially with regard to the active processing approach, Wong (1985, p. 228-9) notes a lack of conceptual clarity, and she comments as follows:

> *"I submit that in self-questioning instructional research we need to conceptualize what specific cognitive processes are manipulated and mobilised by the type of self-questions used ... Such needed conceptualization would clarify what psychological processes mediated the instructional efficacy of self-questioning."*

This section will examine one subset of the self-questioning research, where questions are focused on main ideas. Other aspects of the research will be referred to again in the next major section dealing with the development of greater metacognitive awareness.

In Adams, Carnine and Gersten (1982), teachers instructed children in grade 5 to ask questions about headings and subheadings. Students were selected on the basis of decoding ability adequate to read grade-level social studies texts but poor performance on two short-answer pretests. Study skills instruction took place over four lessons in which students were taught six steps to follow in learning from a text. First they previewed the text by reading headings and subheadings,

then recited each subheading, and then asked a question about each subheading to set a purpose for further reading. In the fourth step students were directed to locate information to answer the questions set for the section plus any other important information, then to reread the subheading and recite the important ideas, and finally to reread all subheadings and recite important ideas related to each. Adams et al note that they could not teach any rules for determining important information, but instead taught this process by modelling. Students were informed of the purpose of each step, and by means of the final two steps were able to monitor the effectiveness of their own study. In addition, at the end of each training session students completed a short-answer quiz which gave them further feedback on the effectiveness of their study and the value of the study strategy. During the four sessions one further group, the independent study group, read the same social studies texts and completed the daily quizzes, but received no instruction. The control group completed only normal classroom activities.

Recall and short-answer posttests were completed immediately and again after two weeks. While there were no significant differences between the groups on the pretests, the instructed group answered significantly more questions correctly on both posttests than the other two groups, who did not differ significantly from each other. Differences in free recall were in the same direction but did not reach significance.

Adams et al make several comments regarding these findings. First, they note that independent seatwork is not an effective method of increasing study skills; the independent study group who read the same texts and took the same tests as the instructed group was not significantly different to the untreated control.

Second, although the training improved performance, Adams et al suggest that the training period was not sufficient for students to master independent application of the strategy; while average scores for the daily quizzes were around 78%, posttest scores for the short-answer tests averaged only around 47%. Nevertheless, Adams et al (1982, p. 49-50) did observe some changes in the way the skill instruction (SI) group approached social studies texts.

> *"Between the pretests and the posttests, one of the notable differences in the study behaviour of students in the SI sample was that instead of reading nonstop through the passages, the students were overtly using strategies that increased their attention to the material. Students, to a varying degree, attended to the subheadings, stopped to re-read, reviewed after reading, or took notes."*

Singer and Donlan (1982) note that if students are directed to ask questions without instruction, they often focus on explicit information of minor importance. To avoid this problem, Singer and Donlan provided students in grade 11 with a structure to enable them to focus on important points. The authors taught the students to ask general questions based on five main aspects of story grammar, namely, leading character, goal, obstacles, final outcome and theme. Students were taught to ask general questions related to these headings, such as: What action does the main character initiate? and then to instantiate the general question with specific content from the story being read. After an introductory lesson, each one of the headings and associated general questions were taught in one separate lesson.

In each lesson, the class stopped reading part way through a story and each student wrote three questions. At the end of the story students wrote any further

questions which had occurred to them. The control group answered questions posed by the teacher. After each of the six reading sessions, all students answered multiple-choice questions about the main points of the story. There was no separate posttest following the training sessions.

In the first two sessions there was no significant difference between the experimental and control groups in their response to the multiple-choice tests, but from the third session on scores of the experimental group exceeded those of the control. According to Singer and Donlan (1982, p179), the treatment lead to improved memory for important ideas because *"the schema-general and story-specific questions enabled the reader to select specific events from the content of the story and assimilate them under specific knowledge structures instead of only identifying and storing words in memory as they occur in the story."*

However, the procedure used by Singer and Donlan is similar to that in adjunct question research; there is no test of spontaneous strategy usage because there is no transfer test where students are not directed by the experimenter to apply the instructed strategy. The authors attempt to counter this criticism by pointing out that only one aspect of story grammar was taught in each lesson, but that all aspects were tested in each lesson, so there was some measure of transfer. In addition, each lesson involved a different story. However, this point does not overcome the criticism that prior to each test the students were directed to formulate questions, and hence did not have to choose to apply the strategy spontaneously.

There was also no indication of any metacognitive component in this study. Students do not appear to have been given a rationale for the strategy, and the treatment would not qualify as the informed, self-control training advocated in

the introductory section to this chapter. Finally, on a point of interpretation of results, Singer and Donlan comment that the training effect they achieved may have been due to the emphasis on the story schema as a means of organising information, or the generation of questions, or both. There is no way of determining which factor was most important.

Some similar problems are evident in the study by Andre and Anderson (1978/9), where students in grades 11 to 12 were given written instructions on how to formulate questions about main ideas in introductory psychology texts. In a first session, students were taught how to identify explicit main idea statements in text, and then how to formulate a question about these main ideas, preferably a question asking for a new instance of central concepts or else a paraphrase of the main idea as the basis for a question. In the second session, the steps for generating questions about main ideas were again given to students in the experimental group, and they were told to use the self-questioning technique while reading the two test passages. A control group was told only to read and reread the texts in order to prepare for a test. A short-answer test followed, with half the questions based on main ideas, such as asking students to name novel instances of a concept, and half requiring technical details, such as names or dates.

There was no overall treatment effect, but treatment did interact with verbal ability. Students lower in verbal ability, as measured by the Wide Range Vocabulary Test, benefited from the treatment, while students high in verbal ability did not.

Andre and Anderson followed the same procedure in a second experiment, although a third group of students was created, one which was given no training but was told only to generate one question of the type teachers might ask for

each paragraph.

Both the question-training and the question-no training groups exceeded the control on the short-answer posttest. Treatment effects were again greater for students lower in verbal ability. Both questioning groups spent more time studying the test passages than the control group, but the question-training group formulated significantly more good questions (novel instances, paraphrases) than the question-no training group. The last finding is rather surprising given that there was no significant difference between the two questioning groups with regard to test scores; if good questions played an important role in producing good test performance one would have expected the question-no training group to perform more poorly on the test. Nevertheless, a regression analysis revealed that the percentage of good questions was a reliable predictor of test scores, while study time was not.

Andre and Anderson attribute their treatment effects to two main causes, namely deeper processing and greater metacognitive awareness. Students were required to identify main ideas and to think about them more deeply in the process of formulating a question, and they were assumed to be more aware and consciously in control of their study activity, although there is no direct evidence of the latter.

However, as in the case of Singer and Donlan (1982), there was no test of spontaneous strategy usage or its maintenance over time. There is no indication of transfer, in particular, whether the treatment affected aspects of reading other than identifying instances of central concepts. Andre and Anderson note that students were told the purpose of the study, but there is no information to suggest that these students were fully informed learners according to the criteria set out

earlier.

Wong and Jones (1982) used a modified version of the Andre and Anderson procedure to teach average students in grade 6 and learning disabled students in grades 8 and 9, the latter reading three to four years below their grade level. Students were first taught to identify and underline explicitly stated main ideas to a criterion of 80% correct. Then in two further sessions they were taught a self-questioning procedure by means of the direct instruction model. As in Andre and Anderson, students formulated questions about the main ideas in the text. Unlike in Andre and Anderson, though, students were also given a rationale for the self-questioning procedure, with the instructor explaining the usefulness and purpose of the strategy. The control group read the same texts, but was told only to judge the quality of the writing. Four days of testing immediately followed the training. One each day, the experimental group was directed to write questions for each of two test passages, and then all students answered comprehension questions, the nature and content of which is not specified. Free recall was measured using a third passage on each day.

There was no treatment effect on comprehension scores, although the treatment by ability interaction approached significance. Learning disabled students tended to benefit more from the treatment than normally achieving students. There was no treatment effect on recall scores.

Thus, the results of this study were rather poor, even though tests were conducted immediately after training and students were prompted to use the instructed strategy before comprehension tests, so no maintenance or transfer was involved. It could be that two training sessions were not enough to teach the questioning procedure to the children. Although Wong and Jones produce a graph to

show the number of good questions generated for the test passages, it is difficult to judge from this how well the children performed, for example in comparison to a skilled adult. (Presumably a maximum score as a reference point would not be appropriate in this case.) Chan and Cole (1986) noted that children found underlining easier than self-questioning, so even though children in Wong and Jones were able to underline main ideas, they may not have mastered the questioning technique. If this were the case, however, one might expect an improvement in performance with practice over the four days of testing. Such an improvement did not occur; on the contrary, scores declined over the test period.

Wong, Wong, Perry and Sawatsky (1986) individually taught eight learning disabled students in grades 7 and 8 to summarise social studies texts. Because the students were learning disabled, the training extended over a period of several months. Students were taught to identify explicit topic sentences in single paragraphs, to write this main idea in their own words as much as possible, and then to add some important details. Once this strategy was mastered, the students began working on social studies texts. They were issued with a summarisation grid on which they wrote headings, subheadings, main idea sentences and important details. While completing the grid they were instructed to set themselves a series of questions or self-guidance statements, such as: 'Is there anything I don't understand in this paragraph? What is the main idea sentence? Let me rewrite the main ideas sentence and add important details. How does the summary statement link up with the subheading? Let me review this section. Do my summary statements link up with each other? Do they all link up with the subheading?' Once students reached a criterion of 80% correct on passage summarisation, they took an immediate and a delayed posttest one month after instruction. Both

tests involved social studies and science texts, the latter as a measure of transfer to another subject. Students first summarised and the recalled the texts. During this time they were not given strategy prompts or summarisation grids, indeed no mention was made of the summarisation instruction.

Prior to instruction these students had performed poorly on summarisation pretests. However, over the course of the intervention all students improved dramatically and much of this improvement was maintained on the delayed and transfer tests. Free recall was limited largely to information included in summaries, but this limitation was not considered a problem since the summaries included the most important points.

MacDonald (1986) taught children in grades 7 and 8 to generate questions about main ideas over nine lessons. Students were taught to recognise main ideas in paragraphs according to the direct instruction procedure, then to rank the importance of these ideas, and finally to generate questions about information identified as most important. Students applied the four-step procedure to their independent reading of a novel; first they identified the main idea of each paragraph, after four paragraphs they decided which of the main ideas was most important, they formulated a question about what still needed to be known about this idea, and read on to find the answer. At the end of each lesson the purpose of the skills which had been taught was explained, and suggestions were made as to where these skills would be useful. The control group continued with a normal reading class, where they practised using a dictionary and thesaurus and completed book reports.

Pretests included the SRA Achievement Series reading comprehension subtest, the Test of Reading Comprehension (TORC) general vocabulary subtest, and one

of two social studies passages for which students were to generate good questions prior to a free recall test. Immediate posttests included the TORC syntactic similarities subtest, the TORC paragraph reading subtest, and question-generation and free recall for the alternate social studies passage.

There was no significant difference between the two groups on any tests at either time of testing. Similarly, there were no significant differences between the groups in terms of pretest-posttest gains. The quality of questions generated in the experimental group improved only for children whose pretest recall score was above the 75th percentile for this sample. Question quality had a significant and positive relationship to recall, so for this subset of students the treatment also improved recall. As noted, though, such an improvement was not evident in the experimental group score overall.

A more carefully controlled study of question generation was carried out by Davey and McBride (1986). Unlike Andre and Anderson, they did not use texts containing explicitly stated topic sentences, nor did they restrict questioning to single paragraphs. Grade 6 students were taught to abstract main ideas for themselves, and to formulate questions about them in texts describing people with interesting jobs. They were also instructed to formulate questions linking different sections of text, so the questions encouraged both macrostructure formation and integration. The authors verified that students in the training group had mastered these techniques in the five lessons provided. In addition, this group was given a rationale for the questioning technique, and also a checklist of five questions according to which they could evaluate their own success in applying the strategies. Hence this study qualifies as informed, self-control training.

In addition to the question training group, Davey and McBride also had a control group which only read the training passages, a literal question group which answered literal questions about each training passage, an inferential question group which answered inferential questions, and a question practice group which was given a definition of 'good' questions and asked to generate two good questions per passage without any further instruction.

During two immediate testing sessions, all students were directed to generate two good 'think-type' questions focused on main ideas for each of the test passages. After each passage they answered a set of literal and inferential short-answer questions, and predicted how well they would perform on the test.

Comprehension scores and quality of generated questions were analysed using pretest comprehension scores from the California Achievement Test as a covariate. According to this analysis, both the question-training and the question-practice groups answered more literal questions correctly than the inference questions and control groups, but not the literal questions group. However, in the case of inference questions, requiring the integration of information from different sections of text, the question-training group performed better than all others. In terms of both the quality of questions generated while reading and the prediction of level of performance, the question-training group exceeded all others. Unlike in some other studies discussed in this section, there was no interaction between treatment and ability.

This study indicates that detailed training is required to improve students' ability to answer inferential questions. However, in the case of literal questions, practice with literal questions or a direction to generate questions also improves performance. In the latter case, the ability to answer literal questions probably

improved because this is the type of question students ask themselves when told to generate questions without training, as already noted by Singer and Donlan (1982).

This study is an improvement on others in this section in several ways; the covariance analysis takes into account initial reading ability, there is not such a large overlap between training and test questions since the question-training group was not instructed specifically to generate the type of literal and inferential questions posed during testing, and students in the question-training group received informed, self-control training. Another advantage of this study is that the authors did not use artificial texts where each paragraph contains an explicit topic sentence. Rather, students were required to actively construct their own main idea statements upon which to base a question. However, even in this study there is no measure of spontaneous strategy usage or its maintenance over time.

Although not specifically devised within the self-questioning framework, Salomon, Globerson and Guterman (1989) did teach students in grade 7 self-guiding questions. They used a computer programme called Reading Partner as a more capable peer in the Vygotskian sense to guide students to use strategies they would not use on their own. Students were introduced to three general questions, presented as 'ways to be a good reader who thinks'. These questions included: 'What is the message? What thoughts come to mind? Do I understand the text?' Students were periodically reminded to ask themselves these questions, with the intention that they become internalised in the course of training. Students were also taught and given examples of specific reading strategies, namely, inferences about titles, identifying key sentences, creating images for segments of text, and formulating intermediate summaries. While reading training texts from the com-

puter screen the use of these strategies was prompted by questions such as: 'What does the title tell me about the text I am about to read? What kind of image do I conjure up about the text?' The strategies were taught and practised over three sessions. During this time another group, the content group, read the same texts and answered literal and inference questions, while the control group only read the texts.

At the end of each training session all students completed a questionnaire on how much mental effort they had expended on reading. Time spent reading each text was automatically recorded by the computer. Alternate forms of the Israeli Standardised Reading Test were used as a pretest and posttest, the latter given ten days after training. As part of the posttest and as a measure of metacognitive knowledge, students were asked to give advice to fellow students on how to improve their reading, in one case to a student who was already reading well and in the other to a student who was a poor reader. One month after training students were given an essay writing test as a measure of transfer.

The experimental group spent significantly longer reading the training texts, and reported exerting greater mental effort than the other two groups. The experimental group also performed significantly better on the metacognitive knowledge test. Most of the advice given by this group was specific and strategic, compared to the general and mainly non-strategic advice given by the other two groups. Covariance analysis of the posttest comprehension scores, using pretest scores as the covariate, showed that the experimental group performed significantly better than the other two groups, who did not differ significantly from each other. The experimental group also produced better quality essays in terms of clarity, organisation and development of ideas, although there was no pretest of writing from

which to assess whether this difference was due to the treatment. A path analysis among variables suggested that greater mental effort while working on the strategy programme enabled students to internalise the self-guidance and regulation questions, which in turn lead to improved comprehension and writing quality.

Some of the studies in this section produced very limited improvement in comprehension (Andre and Anderson, 1978-9; Wong and Jones, 1982; MacDonald, 1980). The training period in some of these studies was also very short (Andre and Anderson, Wong and Jones). Many of the studies do not include a test of 'spontaneous' strategy use, or strategy maintenance over time, or transfer to tasks other than those used during training (Singer and Donlan, 1982; Andre and Anderson, 1978-9; Wong and Jones, 1982; Davey and McBride, 1986). The effect of this approach on students of different ability levels is rather variable, since in Andre and Anderson (1978-9) and Wong and Jones (1982) lower ability students tended to benefit from the treatment, while in MacDonald (1986) it was higher ability students, and in Davey and McBride (1986) there was no interaction between treatment and ability.

Although some of the studies suffer from the methodological problems discussed at the beginning of this section, other studies have shown that self-questioning can be a useful technique under some circumstances. In Wong, Wong, Perry and Sawatsky (1986), training had a marked and durable effect on the ability of learning disabled students to retain important information from social studies texts. In Salomon, Globerson and Guterman (1989), the self-questioning training lead to a generalised and durable improvement in comprehension, strategy knowledge and writing ability. What these studies have in common is that the questions both focused on important comprehension enhancing skills (for ex-

ample, 'What does the title tell me about the text I am about to read?' or 'How does my summary statement link up with the subheading?') and provided a basis for students to regulate and monitor their own processing (for example, 'Let me review this section. Is there anything I don't understand in this paragraph?'). Wong (1985, p. 231) draws attention to these two important functions of questions.

> *"Applying metacognitive theory to self-questioning instructional research entails two instructional implications: (a) teaching students to be sensitive to important parts of the text by asking questions such as, What is the main idea in this paragraph? Can I summarize the important points in this paragraph? (b) teaching students to monitor their state of reading comprehension by asking questions such as, Is there anything I don't understand in this paragraph? Such self-questioning is designed to increase awareness of when students encounter a reading comprehension difficulty."*

Self-questioning is a useful technique if it fulfills the dual functions of encouraging better macrostructure formation and better metacognitive regulation. To realise these benefits, children need to understand the purpose of their questions and to have sufficient time and practice to master the skills and make use of them in a strategic way.

3.3.4 Teaching summarisation rules

In chapter one, the macrorules developed by Kintsch and van Dijk (1978) and van Dijk (1977) were outlined, these being deletion, generalisation, selection (in-

tegration) and construction. Brown and Day (1983) elaborated on these rules and used slightly different nomenclature, such as 'superordination' for what Kintsch and van Dijk refer to as generalisation, and 'invention' for what they refer to as construction. Several studies have investigated the effect on comprehension of teaching students to use summarisation rules of this type.

In Rinehart, Stahl and Erickson (1986) regular classroom teachers instructed children in grade 6 to summarise social studies texts over five lessons according to the direct instruction model. The children were taught to delete trivial and redundant information and to select the main idea of each paragraph. An additional rule was also taught, that of relating the main idea to the most important supporting details. However, Rinehart et al decided not to include the superordination and invention rules in the direct instruction because they could prove too difficult for the children. They note, though, that this omission may have limited the effectiveness of their treatment. Students were first taught to construct single paragraph summaries, then to summarise several related paragraphs, and finally to summarise a whole passage without first summarising individual paragraphs. Students were informed as to why and where summarising would be useful, and were given a checklist to monitor their application of the rules, the use of which was also modelled by the teachers. Children in the control group continued with their normal work from basal readers, which did not involve main idea identification or summarisation.

Pretests included the Gates-MacGinitie Reading Tests comprehension subtest and the Wisconsin Design for Reading Development outlining subtest, the latter also being used as a posttest. Immediately after training all students studied a social studies passage for which they were directed to take notes, and then

answered a series of main idea and detail short-answer questions. They also summarised five paragraphs, two of which had no explicit topic sentences, one of which had an explicit topic sentence in the final position, and two of which had an explicit topic statement but spread over more than one sentence. Gates-MacGinitie pretest scores were used as a covariate in the analysis of these posttest scores.

The experimental group produced better quality summaries for only two of the paragraphs, those where a topic statement was explicit and at the beginning of the paragraph. Rinehart et al note that if the invention rule had also been taught, the experimental group may have performed better on the remaining three paragraphs. The experimental group answered more main idea questions correctly than the control group, but there was no significant group difference with regard to detail questions. Rinehart et al suggest that because students had learned to identify and attend to important information, this lead to both more time spent taking notes and to better recall of main ideas. There was no difference between the groups on the outlining test; both groups performed near ceiling level.

This study indicates that summarisation training can improve main idea comprehension, as measured by short-answer questions. The experimental group had not previously practised answering questions, so this test did not replicate training exercises. The training also included the informed, self-control components advocated in the introduction. However, there was no delayed test to measure maintenance of the trained strategies, nor was there any test of 'spontaneous' strategy use since all students were directed to take notes while studying the test passage.

Hare and Borchardt (1984) taught summarisation rules to junior high school students. These authors taught all the rules as elaborated by Brown and Day (1983) and added two further rules of their own, these being, combining more than one paragraph into a single summary statement when the paragraphs were an elaboration of the same topic, and polishing the final summary by rewriting it in a more flowing style. Students were taught the rules and some guidelines for checking their summaries. The teacher modelled the application of the rules and discussed as examples summaries of varying degrees of proficiency which had been prepared as 'models' by the experimenters. Students practised writing their own summaries, first on more considerate texts, then on naturally occurring texts dealing with a variety of subjects. Written feedback was given on each student's work, explaining how they had or had not applied the rules correctly. Training took place over six hours, during which time the control group was involved in what are described as 'usual summer activities'.

Pretests and immediate posttests were of an identical format, and included summarising two passages and listing the summarisation rules students would use if teaching someone else to summarise, an outline of a further passage, a cloze test, and writing out the main ideas of another passage. Pretest-immediate posttest gain scores were used for the analyses. Two weeks after training students again summarised a passage and noted the rules they would teach to others.

In terms of gain scores for both test passages, the experimental groups exceeded the control group in summarisation efficiency (the number of main ideas divided by the total number of words in the summary) and in the number of main ideas produced in one of the two passage summaries. There were no significant group differences in terms of outlining, the cloze test, or writing out main ideas.

With regard to evidence of rule use in the summaries, the experimental groups made significantly greater use of the superordination, deletion, paragraph collapsing and polishing rules; inventing topic sentences, however, was still difficult even for the trained students. In terms of rules the students stated they would teach others, the experimental groups exceeded the control for superordination, deletion and topic sentence rules, but Hare and Borchardt note that a correlational analysis revealed *"negligible and nonsignificant relationships between stated and actual rule use"* (p. 74).

On the delayed test, the experimental groups showed greater summarisation efficiency, greater actual and stated use of deletion, paragraph collapsing and polishing rules, and greater stated use of the topic sentence rule. Thus, the gains made in summarisation were maintained, although these gains did not transfer to the other posttest measures.

Day (1986) taught summarisation rules to college students, including average students in a rhetoric class and poor writers in a remedial class. Four groups were taught over two class periods, with three of these groups receiving explicit strategy training. In each of these groups, the instructor first explained the rules and then demonstrated their application on one of the passages previously summarised by the students as a pretest. In doing so, she physically marked the text, crossing out trivial and redundant information, circling exemplars and substituting a superordinate term, underlining explicit topic sentences, and inventing topic sentences when none were present. Under the guidance of the instructor, students followed the same procedure on the second pretest passage. Finally they completed independent summaries of a training passage. In the second session, students received written feedback on their first training summary, watched

the instructor again demonstrating the rules by physically marking this training passage, and then independently summarised a second training passage.

In the fourth group, the self-management only group, the instructor also modelled the application of the rules, but without marking the text and without giving students the opportunity to do so. After her oral demonstration, students independently wrote another summary of the first pretest passage before summarising the training passage. In the second session they also received written feedback, listened to the instructor orally demonstrating the rules on the first training passage, and then wrote summaries of the second pretest and second training passages. They received general self-management instructions to pay attention and check their work. The same general instructions were also given to one of the explicit training groups (rules plus self-management), while for a second group the instructions were specifically related to the summarisation (rules integrated with self-management). The third group received no self-management instructions at all (rules only).

The pretest, immediate posttest and two week delayed posttest each involved summarising two passages. Day analysed the results in terms of the application of each rule. The use of deletion rules by all students was near ceiling level and did not differ between groups or times of testing. The four groups did not differ in their use of the superordination rule on the pretest, but the three explicit rule groups performed significantly better than the self-management group on training passages, and only they showed significant improvement from the pretest to both the immediate and delayed posttests. Similar results were obtained for the use of the selection rule, however here, unlike in the previous analyses, treatment also interacted with ability, and there were differences between the three

explicit training groups. While both average and poor writers benefited from explicit training in the use of the selection rule, average writers benefited more. The rules training plus self-management groups showed better use of the selection rule on both posttests than the rules training only group. With regard to the invention rule, there were no significant differences between groups on either the pretest or the training passages, but on the immediate posttest explicit rules training, especially when integrated with self-management, resulted in greater use of this rule. As in the case of the selection rule, average writers benefited more than poor writers. Only slight training effects were maintained on the delayed posttest. Thus, explicit rules training generally improved the use of summarisation rules, although the importance of adding the self-management component varied with the relative difficulty of the rules. In this case, the superordination rule was relatively easy to learn, the selection rule was moderately difficult, and the invention rule most difficult. Day (1986, p. 208) notes:

> *"Students do not require self-monitoring instructions when the specific strategy being taught is fairly easy; they profit from the pairing of self-monitoring and specific strategy instruction when the strategy is of medium difficulty; and they fail to benefit (at least initially) from self-monitoring when the trained strategy is difficult to implement."*

Although Day achieved improvements in the use of these rules through her training, the study provides no evidence as to whether any other aspects of comprehension and recall improved as a result. In Hare and Borchardt (1984), the summarisation training did not lead to improvements in a cloze test, outlining or writing out main ideas, while in Rinehart et al (1986), training did produce

better responses to main idea questions. From these studies, the effect of formal, rule-based summarisation on comprehension is not clear. However, in chapter two a study by Winograd (1984) was referred to. He concluded that while both summarisation and comprehension involve sensitivity to the structural importance of ideas, summarisation also involves specialised skills in addition to those used for comprehension. In order to enhance comprehension, therefore, it would appear to be more useful to teach children to determine main ideas and their contribution to the overall topic rather than the more specialised summarisation rules. It should also be noted that in their model of comprehension, van Dijk and Kintsch (1983) revised the original macrorules into the form of macrostrategies, according to which the reader uses several sources of information to form a macrostructure while reading, or while rereading in the case of a difficult text. In this case the reader makes use of explicit structural markers, as discussed in the first section, but also uses knowledge of the topic to construct macropropositions from detailed information, as discussed in the second section. A combination of these approaches, together with the type of self-questioning discussed at the end of the third section should provide a solid foundation for improving comprehension.

3.4 Improving metacognitive awareness

In chapter two it was shown that more advanced readers generally have higher levels of the metacognitive knowledge needed to plan, regulate and evaluate comprehension. For example, more advanced readers have a more sophisticated concept of the purposes of reading; older and better readers see their goal as understanding important ideas, whereas the younger and weaker readers see their goal as decoding fluently (Myers and Paris, 1978; Gambrell and Heathington, 1981; Garner and Kraus, 1981-2). Less advanced readers do not necessarily report or demonstrate fewer different strategies, although they use available strategies less effectively to answer questions, recall a text, or tutor younger children to answer questions (Kavale and Schreiner, 1979; Paris and Myers, 1981; Garner, Wagoner and Smith, 1983). In addition, many studies have shown that less advanced readers are less likely to detect problematic information in texts.

The studies to be discussed in this section have focused on setting more meaning-based purposes for reading, increasing understanding of different types of questions and how to answer them, and improving comprehension monitoring.

Some studies within the self-questioning framework have focused on the objective of improving goal-setting (Manzo, 1969; Helfeldt and Lalik, 1976; Singer, 1978; Chan and Cole, 1986). In thinking of questions and checking that they can be answered, the reader has a purpose for reading which involves construction of meaning, not just fluent decoding and single word interpretation.

In an early study, Helfeldt and Lalik (1976) used a reciprocal questioning procedure somewhat similar to that described by Manzo (1969). The teacher asked a question about the section of text being read, and if the student answered cor-

rectly, then s/he could ask a question of the teacher. Helfeldt and Lalik state that students were expected to imitate the teacher's questioning behaviour, however, there is no indication of the type of questions asked by either the teacher or the students. The reciprocal questioning group of grade 5 students scored significantly higher on an immediate posttest using the Van Wagenen Analytic Reading Scales than a control group who had only answered and not asked questions. However, the groups were initially matched only on gender and general intellectual ability, not reading ability. Hence it is impossible to know whether the final difference was purely a result of the training.

Chan and Cole (1986) also do not adequately describe the type of questions asked by children. They compared normally achieving children in grade 3 with learning disabled children matched to them on reading ability, but three years older in chronological age. The experimental conditions were presented to the children as a way the toy robot present had found to improve reading. The conditions included asking the robot two questions of unspecified nature to check understanding, underlining two words that the robot might find interesting, underlining two words and asking questions about them, and reading and rereading the passage. Training took place over four days. On the third day students answered multiple choice questions about the practice texts with the texts present as a measure of comprehension. On the fourth day they answered questions without the text present as a measure of recall. In both cases, though, all groups were directed to use their instructed strategy. On the fifth day students again answered questions about two texts, but since no mention was made of the instructed strategies, these tests were considered to be a measure of transfer.

Training did not produce a significant effect on any of the tests for normally

achieving children. However, among the learning disabled groups all three conditions involving question formation and/or underlining resulted in higher achievement than the read-reread control conditions on all three tests. The underlining and self-questioning treatments were equally effective in enhancing performance on the comprehension and recall tests, but the underlining strategy was more effective in the transfer test. Interviews conducted after the transfer test revealed that very few students used the self-questioning strategy when left to their own devices, while the underlining strategy continued to be used even without teacher directions.

Apart from the mixed results, Chan and Cole also fail to report information important to making a decision about the effectiveness of this treatment. Singer and Donlan (1982) note that without training students often ask poor questions, focused on explicit information which may be of minor importance. In the Chan and Cole study there appears to have been no explanation or modelling of strategies for asking more thoughtful questions or for underlining important words. There is also no information on the content of test questions; if they were focused on explicit information of minor importance, then even poor self-questions could have enhanced performance by focusing extra attention on this information.

In the previous section, the studies using self-questioning taught children to focus their questions on important information; the questions were a means of encouraging children to guide their application of effective comprehension strategies and to monitor their comprehension. In the studies referred to above, the purpose of the questions is less specific and illustrates some of the theoretical difficulties with this approach discussed in the previous section. Salomon, Globerson and Guterman (1989, p. 621) make an important point with regard to their teaching

of metacognitive self-questioning when they state:

> *"It should be noted, though, that metacognitive regulation always applies to particular, task-related cognitive strategies, and thus calls for the latter together with the training of metacognitions."*

Thus, if the function of questions is to set a purpose for reading, the instructor should also ensure that this purpose is focused on important information and incorporates important comprehension skills. In other words, children need to be taught how to frame comprehension enhancing questions. As Salomon et al note, learning such cognitive strategies is an essential part of learning more sophisticated metacognitive regulation and monitoring.

Raphael and her colleagues attempted to improve knowledge about the task and strategy aspects of metacognitive knowledge (Flavell, 1981). Her approach sprang from the body of research on question answering, for example, the literature on adjunct questions mentioned earlier, as well as more recent theorising about metacognition. Adjunct question research attempted to encourage readers to focus on and process more thoroughly the information targeted by questions, often leading to greater retention of such information. However, this type of research did not attempt to enhance the repertoire of strategies readers could spontaneously apply to novel texts. Raphael's approach went beyond adjunct question research by incorporating a metacognitive element to foster such transfer. She reasoned that the ability to answer questions about any text would be improved if students had more metacognitive knowledge about task demands and strategies for meeting such demands. As Raphael and McKinney (1983, p. 84) explain:

> *"Knowledge about sources of information available for answering comprehension questions is assumed to be an important aspect of metacognition. Knowledge that more than one source of information is available is important for readers who rely too heavily on text for answer information, readers who rely too heavily on background knowledge, or readers who fail to recognize the importance of integrating textual information with background knowledge."*

In order to make task demands clearer, Raphael and colleagues taught children to identify three types of questions as classified by Pearson and Johnson (1978) and which the researchers called Right There, Think and Search, and On My Own (Raphael, 1982). In the first type, the text explicit question, the question stem and the answer could be found within the one sentence. The appropriate strategy to answer such questions, therefore, is to locate the question stem in the text and then look for the answer in the same sentence. In the second type, the text implicit question, the answer is still stated in the text, but not in the same sentence as the stem. An appropriate strategy, therefore, is to scan more than one sentence and to look for an answer that can be combined with the question stem. In the final type, the script implicit question, the answer is not stated explicitly in the text. The appropriate strategy in this case is to consult prior knowledge for an answer which makes sense in the context of the passage.

Raphael and Pearson (1982) established with a group of undergraduate students that adults do indeed possess the above knowledge about task demands and strategies. After only a brief introduction to the task, the adults were able to identify and answer the different question types with a high degree of accuracy. Subsequently Raphael and colleagues investigated the teaching of these

skills to children by a method incorporating aspects of the direct instruction model described at the beginning of this chapter (Book et al, 1985). The three question types and their corresponding answering strategies were explained and demonstrated. Children practised the strategies in guided practice sessions on increasingly longer segments of text, and were given immediate feedback. Finally they practised on their own on whole texts.

In Raphael and Pearson (1982), average readers in grades 4, 6 and 8, as well as low and high ability readers in grade 6 worked their way through three booklets incorporating the above sequence of instruction and practice. In the first experiment in Raphael and Wonnacott (1985) average readers in grade 4 did the same. In the latter case there were no treatment effects evident in immediate posttest question answering performance. In Raphael and Pearson, the experimental groups were more able to identify the type of question, but they did not consistently produce better quality answers.

Raphael and Pearson (1982) suggest that these rather unimpressive results were due to insufficient training and a failure to fully master the strategies. Hence Raphael and McKinney (1983) and Raphael and Wonnacott (1985) in their second experiment included spaced practice after the initial training period. In Raphael and McKinney, children of low, average and high reading ability in grades 5 and 8 completed a practice passage once a week for 6 weeks, identifying the three question types and giving answers from the text to match the type of question. Raphael and Wonnacott worked with average readers in grade 4 and continued the maintenance practice for 8 weeks.

Treatment effects were again mixed. In Raphael and McKinney, trained students were superior in their ability to name the question type, but there was no

overall treatment effect on quality of response due to a number of interactions. Training improved the responses of low and average ability students, but not high ability students. Children in grade 5 benefited more from the training than the older children, with the response quality actually lower in the experimental group in grade 8 than in the control group. In Raphael and Wonnacott there was a training effect on response quality on only one of the two posttests. In this posttest no mention was made of the training procedure, whereas in the other posttest, when students had to first name the question type before answering the question, there was no treatment effect on response quality.

Raphael (1984) suggests that the training period has to be adjusted to suit the age of the students. She recommends an intensive week of instruction plus six to eight weeks of maintenance practice in grades 4 and 5, an intensive week of instruction only for grade 6, and a brief orientation period to explain the question types and strategies for grade 8. These suggestions represent relatively modest investments of time. Also, as shown by Raphael and McKinney (1983) and Raphael and Wonnacott (1985), the programme can be implemented by classroom teachers.

Against these positive features, though, it has been pointed out that improvements in the quality of responses to questions did not always occur. Raphael and colleagues included no delayed tests in their work, so there is no indication as to whether any training effects were maintained. In addition, only question answering and naming ability were examined. No other measures of comprehension were included. Thus, it is not clear whether students' comprehension *while reading* improved at all, for example, the ability to notice and rectify any obstacle to comprehension, or whether students merely applied the steps they were in-

structed to follow in response to questions *after reading.* From the way Raphael's strategies are described, it would appear that the text was still available to students while answering questions, hence students were not required to form an integrated memory representation to support later question answering. Pressley et al (1989, p. 17) also note some reservations about this line of research, even though it has produced some improvement in question-answering for younger and less able readers.

> *"It should also be pointed out, however, that the effects produced by question-answer relation training have never been very large. In addition, it is not possible to make a case that improved responding to adjunct questions improves comprehension or long-term memory of the passage since there has been no test of this possibility to date."*

Similar limitations are also evident in the study by Garner, Hare, Alexander, Haynes and Winograd (1984) where children were taught to use text lookbacks to answer questions. The children were remedial readers aged between nine and thirteen years whose decoding skills were average, but whose comprehension skills were at least one year below grade level. In the three training sessions the children were taught why, when and where to use the lookbacks, that is, when they could not answer a question from memory, when the question was text- rather than reader-based, and where they skimmed the text to find the relevant section for rereading. These behaviours were modelled by the instructor, and the children then practised the focused lookback strategy with guidance from the instructor. During this time the control group read the same texts and was taught text processing strategies other than lookbacks. Five days later the two groups were

videorecorded as they read another two short expository texts and answered questions.

Both groups answered about one-third of the text-based questions correctly from memory. For the remainder of the text-based questions, the experimental group used lookbacks more than twice as often as the control group and answered over three times as many questions correctly following lookbacks, with both differences being significant. The experimental group used the text sampling method they had been taught on each lookback, while only half the control group used this more focused approach.

The study clearly achieved its aim in three short lessons with remedial readers, and so demonstrates a useful strategy for improving question-answering. However, Pressley et al's point still remains; there is no evidence of improved comprehension ability, especially not since the two groups recalled a similar amount of information from memory.

In chapter two it was shown that less advanced readers detect fewer of the errors deliberately inserted into texts, although even advanced readers do not necessarily perform perfectly. Several studies have attempted to improve error detection performance.

Markman and Gorin (1981) showed that children detected more errors when they were first given examples of the types of problems to look for. Baker and Zimlin (1989) expanded this instruction further to examine the effect on error detection. Children in grade 4, defined as average or above average readers according to the California Achievement Test reading subtest, were taught to monitor comprehension according to three standards. One group was taught to use microstructure standards, that is, standards which did not require integration across

large segments of text. These standards were lexical (nonsense words), external consistency (prior knowledge violations) and propositional cohesiveness (for example, ambiguous anaphoric expressions). The macrostructure group was taught the standards of structural cohesiveness (thematically consistent but irrelevant information), internal consistency (logical contradiction within the text) and information completeness (information signalled by the text is not forthcoming). In one training session the instructor modelled the application of each strategy separately, and students then applied each standard separately to paragraphs, with prompting and corrective feedback from the instructor. Students then applied all three instructed standards to two further training passages. Finally they were asked to apply all six standards to two immediate posttest passages. Two to three weeks later they examined another two passages containing all six types of errors. The control group received no instruction but completed only the tests. To receive credit for error detection, the children had to both underline and explain a problematic segment.

The instructed groups detected more errors than the control group on both the immediate and delayed tests, so the training effect was maintained. Above average readers detected more errors than average readers, but both benefited from the treatment. The microstructure group detected more microstructure than macrostructure errors, and vice versa for the macrostructure group, indicating a specific training effect. However, the two trained groups still detected more of the errors they were not instructed to detect than the control group, so there was also some generalisation of training. Although there was a clear training effect, Baker and Zimlin note that students still had not mastered the simultaneous application of all six strategies. During training, when students were applying

only single standards, they detected 90% of errors, compared to 77% when they were applying three standards to the final practice passages. However, on the test passages involving all six standards they detected only 54% of errors. There is also no indication in this study as to whether the improved monitoring also lead to improved comprehension, for example, greater retention of important ideas.

Miller (1987) taught children in grade 5 a set of self-statements with the intention of improving their comprehension monitoring. The children were identified as average or above average readers according to the Stanford Diagnostic Reading Test. The texts used in the study contained explicit logical inconsistencies, where a concept appearing in two sentences was replaced on its second occurrence by its opposite. Pilot testing showed that 100% of the errors were detected by adults, but less than 20% were detected by grade 5 children. Instruction took place in one session. In the self-instruction condition the children were taught a set of self-statements to define the task, guide their monitoring, self-reinforce, evaluate progress and obtain feedback. For example, the first self-instruction stated: 'I will read the first two sentences and then stop to think about what kind of problem to look for'. The instructor modelled the application of the self-instructions while reading a text, then the instructor and the child did the same together, and finally the child repeated the task alone. During this time the didactic control group received directions in the second person, such as: 'You should read the first two sentences and then stop to listen about what I want you to do'. These children went through the same sequence of steps of listening to the instructor reading the directions and text, repeating with the instructor and repeating alone.

The pretest, immediate posttest and a posttest delayed by one week each involved reading a set of expository paragraphs, half of which contained a contra-

diction. To receive credit for error detection the two contradictory concepts had to be underlined and explained. In the delayed test the children were also asked to recall the self-instructions and to give a definition and example of opposite concepts.

On the pretest there was no significant difference between the groups. However, on both posttests the self-instruction group detected more errors than the didactic control group. On the pretest there was no significant difference between average and above average readers, but on the posttests the above average readers detected more errors than average readers. When pretest scores were used as a covariate, above average readers in the self-instruction group performed significantly better than those in the didactic control group, while for average readers this difference only approached significance. Thus, the self-instruction treatment was more beneficial to the more highly skilled readers.

Miller notes that after one week all the self-instructions were recalled by only 46% of the average readers, compared to 78% of above average readers. Also, all the above average readers correctly defined and gave an example of opposites, compared to 76% of average readers. Therefore, she suggests that average readers need longer training, and that the training programme should explain the value of the strategy and where it is useful. In addition, she included no measure of children's comprehension and recall other than the error detection.

A further study by Miller, Giovenco and Rentiers (1987) attempted to remedy these problems. The children were in grades 4 and 5 and included above and below average readers according to the Metropolitan Reading Test. The procedure was similar to the previous study, but training was extended to three sessions and each immediate posttest passage was followed by a multiple-choice question to

assess recall of factual information. The children were also informed of the value of the self-instruction procedure and where it was useful. The experimental group again detected more errors on the posttest than the control group, but in this study there was no interaction between ability and treatment; both above and below average readers benefited from the lengthier treatment. The multiple-choice questions were not a useful means of assessing comprehension since the children in both groups scored near ceiling level.

Short and Ryan (1984) taught a self-questioning procedure designed to improve comprehension, but also investigated whether this training would generalise to error detection. The children in grade 4 included three groups of below average readers and a reference group of highly skilled readers, defined according to the Stanford Diagnostic Reading Test. The three groups received either strategy plus attribution training, strategy training only, or attribution training only. In the strategy training the students were taught to ask themselves five questions focused on the most important components of story grammar, questions such as: 'Who is the main character? When and where did the story take place?' The instructor modelled the application of these questions, underlining information in the story relevant to each question. Students practised the procedure over three training sessions. Attribution training involved using five self-statements while reading, among them statements such as: 'Praise yourself for a job well done', 'Try hard', and 'Think how happy you'll feel when you do well on the test'. The instructor also modelled internal attributions and explained the relationship between effort, strategy use and successful performance.

The pretest and a posttest delayed by three days involved free recall of a story and short-answer questions focused on important story components. Six

days after training, students completed an error detection test and the Reading Concept Inventory, which assesses awareness of the functions of reading. The skilled reference group completed only the posttests.

While there was no significant difference between the three groups for free and probed recall on the pretest, on the posttest the two trained groups performed significantly better than the attribution training only group. The two trained groups were not significantly different to the skilled reference group for free and probed recall, or for awareness of the functions of reading. The two trained groups tended to perform better on the reading concepts test than the attribution training only group, but the difference did not reach significance. With regard to error detection, the skilled reference group performed better than all other groups, who did not differ significantly from each other. The skilled group detected about half the errors, compared to only 12% to 22% in the other groups. This result is rather surprising since the trained groups had improved markedly in their ability to retain information related to main story grammar components, yet this improvement did not appear to generalise to the detection of inconsistencies in the same components. Contrary to expectation, the addition of attribution training to the strategy instruction did not improve performance. Possibly the attribution training here was too general, in contrast to the studies by Borkowski, Weyhing and Carr (1988) and Reid and Borkowski (1987), for example, where the importance of effort attributions was linked to specific errors made by the experimenter and the children, and it could be demonstrated that better strategy application and effort would lead to correction of the error.

The effect of error detection training on comprehension overall is difficult to assess, since the first three studies referred to provide no such measure, and the fi-

nal study improved comprehension, but not error detection. It is unfortunate that not all studies included a measure such as retention of important ideas, because it cannot be assumed that training students to search for deliberately inserted errors will automatically improve their overall understanding of a text. It may be that an integrated approach, such as that adopted by Salomon, Globerson and Guterman (1989) would be more beneficial. Here self-guiding questions focusing on important cognitive skills and comprehension monitoring were taught in combination, with good results, and without the use of unnatural training materials.

In summary, in this section various methodological or conceptual limitations to several of the studies have been pointed out. It is concluded that regulation and monitoring of comprehension has been more effectively incorporated into studies discussed in previous sections, for example, Kurtz and Borkowski (1987), Wong et al (1986) and Salomon et al (1989). In each of these studies there was a close connection between the cognitive and metacognitive skills taught, resulting in substantial improvements in the comprehension of important ideas.

3.5 Composite packages: Palincsar

Palincsar, in a series of studies, has taught a combination of the comprehension fostering strategies discussed in previous sections, in a context which encourages students to regulate and monitor their own comprehension. In selecting the strategies to be taught, Palincsar and Brown (1984) refer to six processes which they consider have been shown to be important in theoretical discussions of reading difficulty and the reading education literature, these being:

1. to set a purpose for reading
2. to activate relevant background knowledge
3. to allocate attention to major ideas
4. to evaluate text for internal consistency and prior knowledge compatibility
5. to monitor the level of comprehension
6. to draw and test inferences of various kinds

Palincsar (Brown and Palincsar, 1982; Palincsar and Brown, 1984) focuses on four strategies which incorporate these processes, namely:

- summarise the main idea of each paragraph (processes 3 and 5)
- formulate a question about important ideas (process 3 and 5)
- ask for clarification about any points not fully understood (process 4)
- predict the content of sections to follow (process 6)

Palincsar and Brown (1984) comment that processes 1) and 2) are involved to some extent in all the strategies. In addition, the teacher introduces each lesson by discussing relevant background to the selection about to be read (process 2), and emphasises the purpose of learning from text by setting comprehension questions at the end of each session (process 1). Thus, one purpose for students is to read in such a way that they can answer the teacher's questions, but other purposes include being able to summarise information while reading and checking whether predictions were correct.

As noted in the previous section on macrostructure formation, summarising in the form of selecting or inventing topic sentences for paragraphs is an excellent way of focusing attention on important information in the text (Palincsar and Brown, 1986). The strategy is also a comprehension monitoring technique, since the inability to extract the gist of what has been read signals that comprehension is not as thorough as it might be, and some remedial action is called for (Palincsar and Brown, 1984). Palincsar (1984) notes that materials in the initial stages of her training programme contained a lot of explicit topic sentences so that students could engage in the easier process of selecting rather than inventing main idea statements. Later in the training, however, students were required to synthesise their own main idea statements in the absence of explicit topic sentences.

In formulating a question to ask about each paragraph, the second of the four strategies, students were instructed to focus on important ideas, thus there was some overlap between question generation and summarising, as noted by Palincsar (1984). Like summarising, question generation was assumed to encourage the focusing of attention on central concepts and more active monitoring of comprehension.

The strategy of predicting was used only when appropriate, for example, when there were headings, subheadings, embedded questions in the text, or explicit signals to what lay ahead. Palincsar and Brown (1986, p. 772) describe the function of predicting as follows:

> *"To do this successfully, they must activate relevant background knowledge that they already possess ... They now have a purpose for reading - to confirm or disprove their hypotheses. Furthermore, the opportunity has been created for the students to link the new knowledge they will encounter in the text with the knowledge they already possess. The predicting strategy also facilitates the use of text structure. Students learn that headings, subheadings and questions embedded in the text are useful means of anticipating what might occur next."*

Clarification, the final strategy, was also used only as required. Students were instructed to be alert for possible obstacles to comprehension, such as unfamiliar vocabulary or concepts, unclear referents, or omitted information. When students encountered such obstacles, they were encouraged to take remedial action, such as rereading or asking questions to pinpoint and clarify the difficulty (Palincsar and Brown, 1986).

In addition to having these four strategies explained to them, students were also informed of their purpose and the context in which they were useful (Palincsar 1984; 1985). The strategies and their significance were reviewed at the beginning of each session in the initial stages of training (Palincsar, 1984; 1986). Students were also told to use the strategies to monitor their comprehension, as well as to evaluate their effectiveness in using the strategies. The strategic approach to

comprehension was explained via the metaphor of a football game; to achieve its goal, a team has to plan and review its actions, and to make amendments if certain actions are deemed not to be working successfully. Hence, the Palincsar package was based on informed, self-control training.

In all her studies, Palincsar also charted the students' scores on the daily comprehension quizzes on a personal chart. Thus, students could monitor their own progress and the success of the intervention. In this way, the students should be convinced of the value of their strategic behaviour, and hence be more likely to make continued unprompted use of the strategies when appropriate, as indicated at the beginning of this chapter.

Turning now to **how** the strategies and their monitoring were practised, Palincsar used a dialogue she called reciprocal teaching, modelled on the best interactions between adults and children. In this type of instruction, the expert serves as a model and a guide for the responses of novices, who, because of the guidance given, can immediately participate in the learning dialogue at whatever level they are capable of. The expert's role is to prompt and shape their responses through corrective feedback. Palincsar and Brown (1984, p. 123) describe this general type of instruction as follows:

> *"Children first experience a particular set of cognitive activities in the presence of experts, and only gradually come to perform these functions by themselves. First an expert (parent, teacher, master-craftsman, etc) guides the child's activity, doing most of the cognitive work herself. The child participates first as a spectator, then as a novice responsible for very little of the actual work. As the child becomes more experienced and capable of performing more complex as-*

pects of the task, aspects that she has seen modeled by adults time and time again, the adult gradually cedes her greater responsibility. The adult and child come to share the cognitive work, with the child taking initiative and the adult correcting and guiding where she falters. Finally, the adult allows the child to take over the major thinking role and adopts the stance of a supportive and sympathetic audience. Initially, the supportive other acts as the model, critic and interrogator, leading the child to use more powerful strategies and to apply them more widely. In time, the interrogative, critical role is adopted by the child, who becomes able to fulfill some of these functions for herself via self-regulation and self-interrogation. Mature learners are capable of providing the interrogative, critical role for themselves."

When applied to the reciprocal teaching of reading, all participants in a class silently read a paragraph. The instructor then begins to model the four comprehension strategies, 'thinking aloud' as s/he does so. A new 'teacher' from among the students is assigned for the next paragraph, and the process is repeated. The instructor prompts students who are unable to respond, for example, suggesting how to start a question. The instructor also responds to and shapes the students' contributions, for example, suggesting that a summary statement would be better if shortened. Thus, through the process of observing the instructor's proficient performance and formulating their own responses with the guidance and support of the instructor, the students gradually approximate the instructor's performance.

With the support of the instructor, each student can immediately participate at their own level. Even though only one student functions as the teacher at

any one time, all students are expected to be involved, reading each paragraph, applying the strategies for themselves, and adding any clarifications or predictions not thought of by the 'teacher'. Thus, the method is designed to encourage a high proportion of academically engaged time (Rosenshine and Stevens, 1984). Because of the instructor's intervention, the method should also avoid the problem described by Pearson (1984) of the rich getting richer and the poor getting poorer because they do not know how to productively spend their practice time.

Palincsar (1985, p. 2) suggests that the reciprocal teaching method itself, in addition to the strategies actually being taught, may play a large part in the success of this programme because of the characteristics of this method.

> *"We espouse this form of instruction because (1) the teacher models the activities rendering the underlying processes overt, explicit and concrete; (2) the skills are exercised in the appropriate context of reading, not as isolated, decontextualised skills; (3) the dialogue provides the opportunity for the students to evaluate their facility with the skills and to evaluate the effect employing the skills has on comprehension; (4) the teacher is constantly engaged in diagnostic teaching, providing feedback, instruction, and modeling that will move the student from one level to the next level of competence employing the strategies."*

Certainly the results of the package, comprising the four strategies, the metacognitive component and the reciprocal teaching method, have been impressive. In the initial study (Brown and Palincsar, 1982), four grade 7 students were taught individually. As in most of the studies to follow, children of this age were selected on the basis of adequate decoding, at least 80 words per minute with

less than two errors per minute, coupled with poor comprehension, in this case achievement that was two to three years below grade level on the Metropolitan Reading Comprehension Test. Before, during, and after the intervention students completed daily comprehension quizzes. The ten questions on each quiz were based on the classification by Pearson and Johnson (1978), that is text explicit, text implicit and script implicit questions.

Initially the four students averaged only 15% correct on these tests. After the intervention their scores soared to 85%-90% correct. After six months students were still scoring 60%, which rose back to 90% after four sessions of reciprocal teaching.

Based on these initial results, Palincsar and Brown (1984) conducted a larger study. Six grade 7 students, based on the same criteria of adequate decoding and poor comprehension, received reciprocal teaching in groups of two for a period of 20 days. Another group was helped to locate information, in a manner similar to that used by Raphael and colleagues. Students answered the three types of comprehension questions, and the instructor guided students in finding the answer in the same sentence as the question stem (text explicit), across more than one sentence (text implicit), and from prior knowledge (script implicit). A third group took the daily comprehension quizzes but received no instruction, and a fourth group completed the tests only.

In addition to the daily tests before, during and after the intervention, students also took a series of delayed tests after eight weeks. For the reciprocal teaching and untreated control groups, generalisation tests of the same form were presented in science and social studies with no mention of or overt connection to the reading instruction. In addition, there were four transfer tests, which required

a written summary of two texts, construction of sets of comprehension questions of the type teachers would ask, detection of incongruities in texts, and rating the importance of idea units.

In the pretesting period there was no significant difference between the four groups; all averaged between 30% and 40% correct on the experimenter-designed comprehension tests. The reciprocal teaching group steadily improved to reach an average of 80% correct during posttesting. This gain was maintained on the delayed tests 8 weeks later. While the locating information and practice control groups improved slightly, they were still not significantly different to the untreated control group by the end of the intervention. Thus, practice in answering questions did not in itself improve comprehension.

The reciprocal teaching group was also superior to the untreated control group on the generalisation tests in social studies and science, and on three of the four transfer tests. On summarising, question generation and anomaly detection the reciprocal teaching group improved from the pretest to the posttest, while the control group did not. On the fourth transfer test, rating the importance of idea units, there was no significant difference between the two groups.

Three months after the intervention the reciprocal teaching group showed an average gain of 15 months on the Gates-MacGinitie Reading Tests comprehension subtest. No comparative information on the control group is given. Nevertheless, it is clear that the reciprocal teaching group made considerable gains as a result of the intervention.

This study also raises a number of interesting questions. One of the six students stood out by showing considerably less gain than the others. This student did not progress beyond 50% on the comprehension tests, showed little improve-

ment on the generalisation tests, and made no gain on the Gates-MacGinitie test. A reason for this difference is not evident in the information given, for example, in decoding ability or entry scores for comprehension.

Another question concerns the choice of students for intervention. One student who was reading only a matter of months below grade level on the Gates-MacGinitie test improved by 17 months following the intervention. This observation raises the question whether normally achieving students would also benefit from this programme.

As a final comment, the impressive gains were achieved as a result of an intensive intervention. There were only six students in the reciprocal teaching group, and they were taught in groups of two, a luxury which few teachers would have. Such small groups allowed the experimenter to interact closely with each student to diagnose and remediate problems in the reciprocal dialogue. The case of Charles is a good example. This student had a low IQ score, was reading at the equivalent of grade 3 level, and could not formulate a question. Palincsar and Brown (1984, p. 139) describe the teacher's initial interaction with Charles: *"The teacher, estimating that he is having more than usual difficulty with the task, opens her interaction by stating the main idea. She continues to lead him, asking for a 'why' question but, receiving no response, she resorts to forming the question for him to mimic. Even imitating a fully formed question is difficult for Charles. Again, on Day 4, the teacher formulates the question, but this time she waits until Charles comes very close to an adequate question by himself."* Although this treatment is highly beneficial for Charles, in much larger groups it is possible that other students would become distracted during interactions of such a basic nature. The question of group size and maintaining the engagement

of all students in larger groups will be discussed further below in relation to other Palincsar studies and the experimental work for this thesis.

In a number of studies Palincsar has investigated the teaching of larger groups. In the second study reported in Palincsar and Brown (1984), teachers of remedial reading received three training sessions on the reciprocal teaching method before teaching their existing groups. The groups contained between four and seven students, with no group scoring more than 50% correct on pretest comprehension tests. After 15 days of intervention all groups were scoring averages of between 70% and 80% on posttest comprehension tests. This level was maintained on the set of tests eight weeks later. On three of the four transfer tests the groups showed modest but statistically significant gains, enough to raise them to the level of an average grade 7 student. While the gains were not as great as in the first study reported by Palincsar and Brown (1984), the results indicate that reading teachers can conduct the intervention successfully with groups of up to seven students meeting the criteria of adequate decoding but poor comprehension.

Palincsar (1987) further investigated the teaching of larger groups by six classroom teachers, only some of whom were reading teachers. Two classes taught by each teacher were chosen for study, one as a control group receiving traditional reading instruction, and one as a reciprocal teaching group. The control groups completed the same daily comprehension tests as the experimental groups. The size of groups ranged from 7 to 19, with the largest reciprocal teaching group containing 16 students. Pretest achievement levels in the groups were variable, ranging from a class average of 34% to 60% on the pretest comprehension tests. On the Gates-MacGinitie test the groups showed average delays of between 0.5 years and 3.5 years in comprehension level.

The teachers were trained on one full day and two further sessions. During the intervention contact with three of the teachers was limited largely to observation and discussion, since they were applying the method well. However, the other three teachers required more extensive assistance and coaching.

Students in the experimental groups practised the strategies for 5 days on worksheets, and then received 20 days of reciprocal teaching. Prior to the intervention, students in the reciprocal teaching groups did not differ significantly from those in control groups in terms of comprehension, summarising and question generation. However, on the daily comprehension tests the reciprocal teaching groups improved steadily while the control groups in general did not.

In the comprehension posttests, a criterion of 70% correct on four of the five tests was set. Within the reciprocal teaching groups, 70% of the students reached this criterion, compared to 25% in the control groups. In the reciprocal teaching groups a further 26% of students had made a gain of at least 20% on their pretest scores, compared to none of the control students. In the control groups 60% of the students made no substantial gains, compared to 3% in the reciprocal teaching groups. The control groups once again indicate that regular practice with comprehension tests is not a particularly successful way of improving comprehension.

Regarding the transfer tests, the reciprocal teaching groups scored higher than the control groups on posttest summarisation, but not on question generation. Palincsar notes that pretest scores for summarisation reliably predicted pretest comprehension scores. Improvement in summarisation and question generation predicted improvement in comprehension. Decoding scores predicted comprehension only early in the intervention, not in the second half.

This study indicates that the reciprocal teaching method can be used in groups of up to 16 students. These students were also more diverse in terms of decoding and comprehension ability than those in previous studies, where students were selected according to specified criteria. However, only group means are reported, not the range of students within each group, so it is not completely clear just how diverse the students were. Palincsar notes only that she attempted to choose those classes where the greatest number of students met her adequate decoding but poor comprehension criterion. This study again indicates, like Palincsar and Brown (1984), that classroom teachers can implement the reciprocal teaching method, although in some cases quite intensive training is needed because some teachers find the method difficult.

Another option for dealing with larger groups is to use peer tutoring, as investigated by Palincsar, Brown and Martin (1987). In this study, three teachers familiar with the reciprocal teaching method selected tutors on the basis of their pretest comprehension summarising and question generation ability, as well as teacher judgement of suitability. In two classes, each tutor was assigned two tutees, while in the third class each tutor had only one tutee. The tutors received training in the reciprocal teaching method for 10 days, during which time their tutees completed individual worksheets. Both groups completed daily comprehension quizzes. The tutors then conducted reciprocal teaching sessions with their tutees for 12 days.

During their training period, the tutors improved on their pretest comprehension scores by an average of 11%. This gain was maintained for the remainder of the study. However, their future tutees completing worksheets and comprehension tests during this same period did not improve, indicating again that practice

on its own does not lead to improvements in comprehension. During the reciprocal teaching period, the tutees steadily improved to reach the level of normally achieving grade 7 students and the entry level of their peers selected as tutors. Of the 15 tutees, 8 gained 20% or more in the course of the intervention, but 3 gained less than 10%. There was generally no improvement on the transfer tests of summarising and question generation; only the tutors improved on question generation, and neither group improved on summarisation. This finding does not accord with the finding of Palincsar (1987) that improvement in summarising and question generation predict improvements in comprehension.

Thus, peer tutoring can be used quite successfully to implement the reciprocal teaching method. However, the training of tutors depends on teachers being thoroughly familiar with the method themselves. Also, in spite of the overall success of the approach, 20% of tutees made no substantial gain in their comprehension scores. This difference between tutees in their response to the intervention again raises the question of why some students do not benefit, a question already raised in regard to Palincsar and Brown (1984).

In all of the studies mentioned thus far the whole reciprocal teaching package was implemented. In order to assess which particular parts of the package were responsible for gains in comprehension, Palincsar (1985) conducted two experiments in which the strategies and the interaction patterns were systematically varied. In the first experiment one group received reciprocal teaching for all four strategies for 20 days, while a second group received summarisation training only for 8 days, then question generation only for 8 days, and finally all four strategies for 4 days. For a third group the order was reversed, with question generation first and summarisation second.

In pretesting, the Question-first group averaged around 45% correct on the comprehension tests, compared to slightly above 50% for the Summary-first and Four-strategies groups. The Four-strategies group made more rapid gains, especially initially, and achieved a higher comprehension level than the other two groups in the immediate posttesting period. However, it should be noted that the difference between the Four-strategy and Question-first groups in the immediate and 3-month delayed posttests were of about the same magnitude as in the pretesting session. Thus, the Four-strategies package was not hugely more effective than the sequence of Question only-Summarising only.

The comparison between the Question-first and Summary-first groups suggests an interesting difference in the efficacy of the two strategies. In the first block of 8 days both groups improved their comprehension. However, in the second block, the group that switched from summarising to question generation made no further gains, while the group that switched from question generation to summarising showed quite dramatic improvement. These findings would suggest that summarising is a more powerful strategy than question generation. Both groups made further progress with all four strategies, suggesting that the focus on comprehension monitoring and engagement of prior knowledge encouraged by the clarification and prediction strategies also have a role to play in improved comprehension.

In her second experiment, Palincsar (1985) compared different methods of delivering instruction. One group received the traditional reciprocal teaching package for 12 days. A second group received 4 days of reciprocal teaching as an introduction to the four strategies, and thereafter wrote their responses to paragraphs. The teacher marked good responses with a star. In this condition

students practised the strategies, but received minimal feedback or modelling. In a third group the teacher demonstrated the application of the four strategies as the class read each paragraph, but students did not practice the strategies, at least not overtly. Their only involvement was answering traditional teacher questions on the reading passages. Thus, while there was modelling of the strategies in this case, there was no practice of the strategies and hence no feedback on the success of their application. The fourth group completed worksheets on each of the four strategies separately, so there was no opportunity to integrate the use of all strategies in the context of reading. However, there was opportunity to practise each strategy and to receive feedback from the teacher as s/he assisted students with the worksheets. A fifth group served as an untreated control and completed only the assessments.

In the pretest and the first half of the intervention all groups were comparable with regard to comprehension scores; in the pretest, judging from the graphs presented, all groups averaged around 40% correct, while during the first half they averaged around 50%, with the control group somewhat lower. However, during the second half of the intervention the reciprocal teaching group improved to slightly over 60%, while the reciprocal teaching plus written response group and worksheet group were slightly over 50%.

This experiment suggests several important features a successful teaching method should incorporate. Although students should have models of the strategies being effectively applied, they must also have opportunity to practice applying the strategies themselves, since observation alone was not enough to bring about substantial improvement. Students need to receive specific corrective feedback since practice without this type of guidance produced only limited gains.

Students need to integrate the strategies in the context of actual reading, since isolated worksheet practice produced limited gains.

The Palincsar experiment does not rule out the possibility that these features could be incorporated into a methodology other than the classic reciprocal teaching. One such modification is referred to to in Palincsar and Brown (1986). In this study, students read from science textbooks. For each section heading, students wrote two statements predicting what they thought the section would be about. Predictions were then discussed. Next the students read the section, usually about four paragraphs, and then wrote two questions and one summary statement about the section. Finally they noted any points that needed clarification. All written responses were discussed in the class. Comprehension was assessed by weekly tests, but these tests covered material studied as part of the training and so were not independent posttests.

In the pretest these students averaged 36% correct, but after 4 weeks of the instruction students generally scored 20% higher. It should be noted that greater gains were achieved by this group than by the group in Palincsar (1985) which wrote responses but received very little interaction and feedback. It would appear, then, that responses can effectively be written provided that instruction also involves discussion and more extensive feedback. This finding is an important one since a written format could make it easier to work with large, intact classes.

Although the results reported have been impressive, Paris, Wasik and van der Westhuizen (1987, p. 23) draw attention to what they see as limitations in the investigation of reciprocal teaching.

> *"It is important to note that all the evidence on reciprocal teaching has been accumulated with relatively few teachers and students from*

seventh-grade classes of poor readers. It is not known if reciprocal teaching improves the reading comprehension of students across a wide age range, whether regular teachers can be trained easily to use it, or whether it promotes comprehension of students who have average or above average reading abilities."

Some evidence relating to these points can be found in the Palincsar studies. Palincsar and Brown (1986) indicate that the reciprocal teaching dialogue can be used with children as young as grade 1, provided that the text is read aloud by the teacher. In Palincsar and Brown (1984) and Palincsar (1987), classroom teachers delivered the instruction, although in Palincsar (1987) in particular the in-service training was relatively time-consuming and half the teachers still found the reciprocal teaching dialogue difficult. In Palincsar, Brown and Martin (1987), the students selected as tutors were more capable than their peers, but they benefited from training as well as their tutees. Nevertheless, Paris et al's comments are valid in that it would be useful to have more information on the points they raise, in particular, to study children of different ages and abilities, and to investigate a format of instruction that may be easier for teachers to implement.

3.6 Composite packages: Paris

The reading programme devised and tested by Paris and his colleagues is, like that of Palincsar, a composite package, involving a variety of comprehension strategies and a strong metacognitive component. According to Paris and Oka (1986b), the programme involves strategies to enhance construction of meaning, including elaboration, inference, integration, activation of prior knowledge and summarisation, as well as strategies to foster comprehension monitoring and regulation, including rereading, self-questioning, checking consistencies and paraphrasing.

In addition, the Paris programme aims to increase 'reading awareness' and to test the relationship between such awareness and comprehension performance. Paris and Jacobs consider three aspects of reading awareness, namely, evaluation, involving appraisal of the task and one's own abilities, planning, involving selection of actions to accomplish the task, and regulation, involving the monitoring of actions and redirection of effort in case of difficulties. The Paris programme, named Informed Strategies for Learning (ISL), explicitly teaches these aspects of reading awareness.

The list below, taken from Paris and Oka (1986b), describes in more detail the strategies taught in the 20 modules of the ISL programme. This list is an elaboration of the set of strategies taught in earlier studies, for example in Paris, Cross and Lipson (1984) and Paris and Jacobs (1984).

I. Awareness of reading goals, plans and strategies

1. Goals and purposes of reading: "Hunting for reading treasures"
2. Evaluating the reading task: "Be a reading detective"
3. Comprehension strategies: "A bag full of tricks for reading"

4. Forming plans: "Planning to build meaning"

5. Review

II. Components of meaning in text

1. Kinds of meaning and text content: "Turn on the meaning"

2. Ambiguity and multiple meanings: "Hidden meaning"

3. Temporal and causal sequences: "Links in the chain of events"

4. Clues to meaning: "Tracking down the main idea"

5. Review

III. Constructive comprehension skills

1. Making inferences: "Weaving ideas"

2. Preview and review of goals and tasks: "Surveying the land of reading"

3. Integrating ideas and using context: "Bridges to meaning"

4. Critical reading: "Judge your reading"

5. Review

IV. Strategies for monitoring and improving comprehension

1. Comprehension monitoring: "Signs for reading"

2. Detecting comprehension failures: "Road to reading disaster"

3. Self-correction: "Road repair"

4. Text schemas and summaries: "Round up your ideas"

5. Review: "Plan your reading trip"

The strategy within each module was described by a metaphor, such as 'Be a reading detective', in order to help children understand the cognitive process involved. The metaphor was illustrated on a bulletin board, with the reading detective, for example, illustrated by a magnifying glass. The metaphor also appeared on all worksheets. It enabled students to discuss abstract cognitive processes in a concrete way, and also served as a memory cue to recall strategies in the review modules.

Inspection of the list above shows that sections I and IV relate largely to the evaluation, planning and regulation involved in reading awareness, while the eight strategies in the remaining two sections relate largely to the construction of meaning. In comparison to the Palincsar programme, ISL is differentiated into many more explicit strategies. In the studies of Palincsar, some of these strategies emerge from the dialogue amongst teachers and students. For example, the 'teacher' for a particular paragraph always considers whether anything needs clarifying. If a point is deemed unclear or confusing, this can provide an opportunity to discuss possible remedial action. Such a dialogue can encompass most of what is taught in section IV of ISL, raising the question as to whether the Palincsar approach is not a parsimonious one for improving reading comprehension. Students need to internalise and use only four main strategies, plus the consideration of headings and sub-headings to activate prior knowledge, and this small number of strategies is continually practised and refined in the course of the intervention. It may be that Paris and colleagues chose to use their more de-

tailed approach with two sections focused on reading awareness because of their specific interest in increasing such awareness and investigating its relationship to comprehension performance. This point will be discussed further below when examining the results achieved with the ISL programme.

Before discussing these results, the method of teaching the ISL programme will be briefly discussed. As already mentioned, each strategy was described by a metaphor; the metaphors are set out within quotation marks on the above list of strategies. The aim of the metaphors was to help children understand, discuss and remember relatively abstract cognitive processes. Paris and Oka (1986b, p. 31) give an example of how a metaphor might be used to enhance understanding of strategies. *"In order to describe comprehension monitoring, for example, we drew analogies between traffic signs (e.g., stop, yield, caution, dead end) and similar cues in text that might signal occasions to paraphrase, to look up new words in the dictionary, to slow down the reading rate, or to reread."* Metaphors were visually represented on bulletin boards, which were a reference point for each lesson. This representation was repeated on worksheets.

The first four modules in each section of the ISL programme each taught one strategy. Every module involved three 30-minute lessons. Within this sequence, the strategy and its metaphor were introduced by the teacher, who also explained when and how to use the strategy, in keeping with the requirements of informed training. The teacher then modelled the application of the new strategy, after which students completed worksheets to provide immediate practice and feedback. With each successive worksheet, instructions became less explicit so that students had to take more responsibility for selecting appropriate strategies, thus encouraging greater self-control over learning. In the third lesson of a module

students read content area material, for example, from social studies or science books, to encourage transfer of the strategies to contexts other than reading lessons. The fifth module in each section revised and integrated the use of all strategies taught in that section.

Prior to each worksheet, students discussed when and why to use particular strategies. After each worksheet they discussed what they had done, including the ease or difficulty of using certain strategies, and the benefits of adopting this strategic approach. Paris and Jacobs (1984, p. 2091) emphasise that children must be convinced that strategies are worthwhile, and group discussions play an important role in this persuasion process. *"The strategies must be regarded by children as useful means to desired ends that have personal significance and utility. When these skills are perceived in this manner, they become rational plans for problem-solving that can be recruited and self-controlled."* In other words, students must not only internalise the strategies, but must also be motivated to invest the time and effort involved in their use if the strategies are to be used spontaneously and transferred to other contexts. This point has also been strongly emphasised in the work of Borkowski discussed in previous chapters, for example, Carr, Borkowski and Maxwell (1991).

Turning now to the effectiveness of the ISL intervention. In the study by Paris and Jacobs (1984), the children in four intact classes, two at grade 3 and two at grade 5 level, received two 30-minute lessons per week for 14 weeks. Four comparable classes served as controls. The ISL programme was a shorter and earlier version of that that described at the beginning of this section, although embodying the same principles and methodology.

Prior to, and again one month after the conclusion of the intervention, all

children completed comprehension and awareness tests. The comprehension tests included the Gates-MacGinitie Reading Tests comprehension subtest, a cloze test with every fifth word deleted, and an error detection task. For this task, the children read two passages, each containing two nonsense words and two clauses with scrambled word order. Scores represented the proportion of errors correctly underlined.

To assess reading awareness, Paris and Jacobs conducted structured interviews based on 19 open-ended questions, of which only 15 provided useful data. These 15 questions were arranged into three subscales relating to the areas of evaluation, planning and regulation respectively. The first question from each section is set out below as an example.

Evaluation: 'What's the hardest part about reading for you?'

Planning: 'If you could read only some sentences in a story because you were in a hurry, which ones would you read?'

Regulation: 'Do you ever go back and read things over? Why?'

On the pretest, there were modest but significant correlations between reading awareness scores and performance on the three measures of comprehension, ranging from r = .24 to r = .40. However, as will be discussed later, such correlations need to be treated with caution when investigating the contribution made by reading awareness to comprehension performance (cf McLain, Gridley and McIntosh, 1991).

With pretest scores taken into account, the experimental classes performed significantly better than the control classes on reading awareness, cloze and error detection posttests. However, there was no treatment effect on Gates-MacGinitie test scores.

On the basis of the pretest scores, students were divided post hoc into high, average and low awareness groups. While at the pretest there were approximately equal numbers of experimental and control class children in each group, at the posttest there were significantly more experimental than control class children in the high awareness group, while the situation was reversed in the low awareness group. In both the pretests and the posttests, children in the high awareness group performed better on all comprehension tests than children in the average and low awareness groups.

Thus, the ISL intervention clearly increased reading awareness, as measured by the structured interview, and also improved cloze and error detection performance. No age or pretest awareness level interactions were found, suggesting that the treatment benefited all groups of children.

Paris, Cross and Lipson (1984) similarly studied the effects of the ISL programme on four intact classes, two in grade 3 and two in grade 5. The intervention continued for four months, with two 30-minute lessons per week. In addition to the Gates-MacGinitie Reading Tests comprehension subtest and the cloze and error detection tests as measures of comprehension, Paris et al also included the Paragraph Reading subtest of the Test of Reading Comprehension (TORC). Unlike the Gates-MacGinitie test, the TORC is not timed. In the TORC, five multiple-choice questions follow each paragraph length story, with questions focused on the best title for the passage, details about the story, and two inferences. Instead of the interview to assess reading awareness, Paris et al constructed a 20-item multiple choice test to measure the extent of learning of the strategies taught in ISL. Posttests were once again conducted one month after the conclusion of the intervention.

Analysis of posttest scores with pretest scores as covariates revealed no significant treatment effect on either the Gates-MacGinitie test or the TORC. However, the experimental classes were superior to control classes on both the cloze and error detection tests. Once again grade did not interact with treatment, so both grade 3 and grade 5 children benefited.

The strategy knowledge test, administered as a posttest only, was used to divide children at each grade level into three groups. Children in the top third were designated 'good learners', those in the bottom third 'poor learners'. A greater number of children from the experimental classes were in the top category than children from the control classes, with the position reversed for the bottom category. Nevertheless, a few children in the experimental classes had not learned the strategies well (9 children), and some children in the control classes knew them without being taught (12 children).

Learner group did not have a significant effect overall on either the Gates-MacGinitie or the TORC scores, but did have an effect on error detection and cloze scores. In other words, children who knew the ISL strategies well made greater pretest-posttest gains on the cloze and error detection tests than children who did not know the strategies well. However, strategy knowledge had no such effect on the Gates-MacGinitie test and the TORC.

In discussing the lack of treatment effect on these two comprehension tests, Paris et al (1984, p. 1249-50) reject as a possible explanation that ISL does not bring about a general improvement in reading comprehension, and instead question the appropriateness of the instruments.

> *"The GATES and TORC measure student achievement that is primarily determined by general intelligence and background experi-*

ences. They are good tests for discriminating students according to ability, but the instruments seem ill-suited as measures of variation in classroom curricula. One might argue that successful performance on standardized tests depends on comprehension strategies, but these seem dissimilar to many ISL strategies and, indeed, unlike many other ordinary reading activities of elementary school children. For example, rereading and paraphrasing in a deliberate fashion may be too time consuming while taking a standardized test."

While these criticisms may apply to the Gates-MacGinitie test, it is not clear how they apply to the TORC. First, according to the information given by Paris et al, only the Gates-MacGinitie test was timed, so children should have had sufficient time to apply reflective strategies to the TORC. Second, three of the five questions for each passage on the TORC involved inventing a passage title and inferencing. The ISL programme included lessons on inferencing and summarising main ideas, strategies which should have had the potential to enhance performance on the above question types.

Paris and Oka (1986b) report a large scale study using 46 intact classes as experimental groups and 25 comparable classes as control groups. All classes were taught by their regular teachers, in contrast to the previous two studies described, where the experimental classes were taught by a researcher. The ISL programme was expanded to the 20 modules described at the beginning of this section, with the expansion largely due to the addition of several more strategies for constructing meaning. Strategies for increasing awareness and comprehension monitoring remained much the same. The duration of the intervention was also expanded, to a total of 30 weeks, by the addition of more review modules where

previously taught strategies were revised and integrated, and more practice in applying strategies to content area reading.

Teachers were provided with extensive materials prior to the intervention, including the bulletin board displays, worksheets and practice passages. The materials and methodology were explained to the teachers in four workshops, as well as periodic meetings throughout the intervention.

Measures of comprehension were again the Gates-MacGinitie test, a cloze test, and an error detection test. Reading awareness was measured with a 22-item multiple-choice test, the Index of Reading Awareness (IRA), based on the structured interviews conducted by Paris and Jacobs (1984). A written test could be administered to large groups of children, but also aimed to overcome some of the problems associated with verbal reports of strategy knowledge (Jacobs and Paris, 1987). The IRA included four sections, namely, evaluation, planning, regulation and conditional knowledge. In addition to the IRA, children rated a series of statements on a Likert scale to reveal their perceptions of the value, enjoyment and difficulty of reading, a measure referred to as Self-Perceptions about Reading.

Using pretest scores as covariates, the analysis of posttest scores showed that children in experimental classes performed significantly better than children in control classes on the cloze, error detection and IRA tests. In the case of error detection, improvement resulted largely from a reduction in the number of false alarms (underlining unproblematic sections of text), rather than from an increase in the number of hits (underlining anomalous sections).

The ISL intervention did not significantly improve scores on the Gates-MacGinitie test or Self-Perceptions about Reading. Paris and Oka (1986b) con-

sider that self-perceptions are a relatively enduring characteristic, not easily altered by 30 hours of instruction. They also reiterate that the Gates-MacGinitie test is closely related to intelligence and is not sensitive to curricular interventions. The authors indicate that they have been unable to locate a single instructional programme which significantly enhanced Gates-MacGinitie scores, a point which will be discussed further below.

Paris and Oka also examined which factors best predicted Gates-MacGinitie scores. At grade 3 level, cloze scores explained 39% of the variance, with other factors adding small but significant contributions, including error detection (+5%), self-perceptions about reading (+3%) and reading awareness (+3%). At grade 5 level, the cloze test explained 33% of the variance in Gates-MacGinitie scores, with other factors making the following additional contributions to explained variance: self-perceptions about reading (+6%), reading awareness (+4%) and error detection (+2%). Thus, while skills associated with reading accounted for most of the explained variance, reading awareness and self-perceptions about reading made small, but still significant contributions.

The ISL programme devised by Paris and his colleagues places a heavy emphasis on reading awareness, that is, evaluation, planning and regulation activities, as well as conditional knowledge about when and why to apply strategies. However, several questions remain regarding this experimental work, including the reliability and validity of the Index of Reading Awareness (IRA) instrument, the extent of the contribution of reading awareness to comprehension performance, and the effectiveness of the ISL approach as a whole in bringing about a general improvement in comprehension.

McLain, Gridley and McIntosh (1991) investigated the reliability and valid-

ity of the IRA. Regarding one measure of reliability, Jacobs and Paris (1987) reported a test-retest correlation of .55 for the 544 children in control groups in the study reported by Paris and Oka (1986b), with pre- and posttests separated by 8 months. McLain et al further investigated the reliability of the IRA by examining internal consistency within the four subscales after administering the instrument to children in grades 3, 4 and 5. Cronbach's alphas ranged from .15 to .32, while for the total test score they were .61 using item scores and .56 using subscale scores. McLain et al (1991, p. 84) comment that the IRA must be used cautiously as a measure of metacognition and add: *"In no case is use of the subscales of evaluation, planning, regulation and conditional knowledge justifiable. The subscales were not internally consistent, nor did primary factor analysis yield interpretable factors. The reliability of the IRA in its present form is marginally acceptable."*

McLain et al also examined the relationship of IRA scores to reading performance. They individually administered the Woodcock Reading Mastery Test (WRMT), which includes four subtests, including Word Identification, Word Attack, Word Comprehension and Passage Comprehension. Word Identification involved orally reading isolated words, while Word Attack involved orally reading nonsense or low frequency words. These two subtests could be combined into a Basic Skills cluster. Word Comprehension consisted of supplying antonyms, synonyms and analogies, while Passage Comprehension involved supplying a missing word (cloze) in short passages. These two subscales could be combined into a Total Comprehension cluster.

Correlations between the IRA scores and these various subscales and clusters were modest, ranging from .22 to .35. In contrast, intercorrelations between the

subscales and clusters of the WRMT ranged from .72 to .96.

McLain et al also carried out multiple regression analyses to determine the best predictors for Total Comprehension and Passage Comprehension. With regard to the Total Comprehension cluster, the Basic Skills cluster explained 68.7% of the variance when it was entered first, while IRA scores added only another 1.9%. If the IRA scores were entered first, they explained only 11.8% of the variance in Total Comprehension, while the Basic Skills cluster added another 58.8%.

In the analysis of Passage Comprehension scores, Word Attack explained 46.9% of the variance, while both Word Identification (+11.4%) and Word Comprehension (+8.8%) made further significant contributions. The IRA added a non-significant 0.2%. As a result of these regression analyses, McLain et al recommend that simple correlations should be treated cautiously because there is an overlap of information between the IRA and basic reading skills. Use of only correlations can lead to an overestimation of the contribution of reading awareness to comprehension performance. McLain et al (1991, p. 86) conclude that: *"At this point, the value of information collected through the extra time and expense connected with the IRA is somewhat doubtful if we do not add any information to predicting reading comprehension beyond that available with the usual Word Attack, Word Identification, etc., scores."*

With regard to the effectiveness of the ISL programme, Paris and his colleagues improved performance on cloze and error detection tests in two smaller scale studies (Paris and Jacobs, 1984; Paris, Cross and Lipson, 1984) and one large scale study (Paris and Oka, 1986). However, in none of these studies did ISL improve comprehension performance on the Gates-MacGinitie test, or on the TORC in Paris, Cross and Lipson (1984). The question then is: What type of

reading skills were improved by the ISL programme?

The error detection tests included two nonsense words and two scrambled clauses per passage. Thus, to detect the errors children needed to apply lexical and syntactic standards to their comprehension monitoring. These standards can be applied without integrating information across sentences. Such integration is necessary to produce a coherent representation of text, but the error detection tasks as devised by Paris and his colleagues give no information as to whether ISL improved this important skill. A more stringent test of integration would be the detection of internal inconsistencies.

A cloze test with every fifth word deleted may or may not test the ability to integrate ideas within a text. In some instances a correct word may be inserted by considering only very localised semantic and syntactic constraints, as in an example such as: The cat was chased __ the dog. Kintsch and Yarborough (1982) consider that cloze tests of this type assess micro- rather than macroprocesses. Their conclusion is based on the observation that poor rhetorical form in a text (randomly organised paragraphs) lead to reduced ability to answer macrolevel questions, but did not affect cloze performance. Shanahan, Kamil and Tobin (1982) go further and question whether cloze tests with every nth word deleted even measure intersentence integration. They found that undergraduates performed just as well on cloze tests when sentences were randomly scrambled as when they were in their intended sequence, even under severe time constraints where readers could not infer the relationships between sentences. There was also no significant difference in responses to cloze blanks in a set of target sentences when they occurred in a normal passage or in a list of unrelated sentences. Henks (1982) advises caution and further research before accepting Shanahan et

al's conclusion that cloze tests are not reasonable classroom measures of global comprehension. Henks points to some factors which could limit the generalisability of the findings, for example, the use of only adult subjects and a limited range of texts, but adds: *"There is a decided possibility that many of the major assertions of the study have merit ..."* (p. 594).

In Paris, Cross and Lipson (1984) there were no treatment effects on the TORC. In this test, the ability to construct a good passage title would involve macroprocesses, that is, the ability to abstract from detailed information the gist or main theme. The inference questions on the TORC may or may not involve macroprocesses, depending on the type of inferences required. It may be, then, that the ISL intervention improved only microprocesses operating largely at the level of the sentence, rather than macroprocesses involving integration across sentences and larger segments of text.

Turning now to the lack of treatment effects on the Gates-MacGinitie test, a timed test could certainly make it more difficult for students to apply a strategic, reflective approach to their reading. Paris et al (1984) comment that no training programme has been able to bring about an improvement on this test. However, Sampson, Valmont and van Allen (1982), using a quasi-cloze procedure as part of normal reading instruction in grade 3, produced significant gains relative to a control group. Palincsar and Brown (1984) reported impressive gains by some of the six students in their reciprocal teaching group. Unfortunately, though, no information is given on gains made by the other three groups not taught according to the Palincsar approach. Hence there is no way of evaluating the effect of the intervention as opposed to other factors, including a possible test-retest effect, with children becoming more familiar with the specialised format of a timed,

standardised test. Nevertheless, the points made in this concluding section do at least suggest that the ISL programme brought about an improvement in only a restricted range of reading skills, and that it may be for this reason, not just the insensitivity of standardised tests, that no gains were found on the Gates-MacGinitie test and the TORC.

3.7 Conclusion

To conclude this chapter, evidence from the preceding chapters will be drawn together to provide a theoretical and empirical foundation for the training programme undertaken in this study.

Armbruster et al (1991) refer to a survey by the National Assessment of Educational Progress in the United States, showing that high school students have comprehension skills which are not adequate to cope with tasks they will be faced with on leaving school. There is no reason to assume that Australian students are more skilled in this area. The intention in this study is to improve the ability of upper primary school children to learn from expository texts, in preparation for the type of reading they will be expected to undertake in high school. As explained previously, the work of Palincsar (for example, Palincsar and Brown, 1984; 1986; Palincsar, 1987) provides a sound framework for improving content area reading. The value of the strategies in her programme, and the metacognitive context in which they are taught, has been supported by other studies discussed in this chapter, for example, predicting from titles (Salomon et al, 1989), generating paragraph summaries (Baumann, 1984; Stevens, 1988; Kurtz and Borkowski, 1987; Borkowski et al, 1988; Davey and McBride, 1986), and being alert to any points which require clarification (Salomon et al, 1989; Wong et al, 1986). However, as was previously indicated, partly in response to comments by Paris et al (1987), there are several questions arising from the Palincsar studies, including the following:

1. How successfully can such a programme be taught to children younger than the grade 7 students primarily taught by Palincsar?

2. Can the programme be as beneficial to average and above average comprehenders as to the poor comprehenders primarily studied by Palincsar?

3. Can any specific characteristics be identified to distinguish those students who benefit from the programme and those who do not?

4. Can the programme be implemented in a written format so as to still retain its interactive nature and a gradual transfer of responsibility to students, while at the same time providing a more structured guide for teachers to follow?

Palincsar (for example, Palincsar and Brown, 1984) and Paris (for example, Paris and Jacobs, 1984; Paris, Cross and Lipson, 1986) included a range of tests to assess the effectiveness of their treatment, including also delayed tests and tests where students were not prompted to use trained strategies, but had to do so of their own volition. It was pointed out in chapter three that many studies do not incorporate these aspects; of the studies concerned with improving global comprehension only Salomon et al (1989), Kurtz and Borkowski (1987) and Borkowski et al (1988) include all three aspects to some extent. A further question to be addressed, then, is:

5. Will any training effects be evident without prompting, on a range of comprehension tests other than those used during training, and will training effects be maintained over time without further intervention?

Turning now to the selection of children for the intervention, several studies show that children in grade 7, the age group mainly taught by Palincsar, are beginning to approach adult levels in tasks involving decisions about macrolevel information, in comparison to grade 5 students, who are only beginning to develop these skills. For example, with regard to categorising the idea units of a story into four levels of importance, Brown and Smiley (1977) showed that grade 7 students

were approaching an adult level of performance, while grade 5 students could distinguish only the most important ideas and made no reliable distinction between the remaining three levels. On a related point, grade 7/8 students improved their recall of important idea units following a study period, whereas grade 5 students did not, presumably because these younger children did not isolate the important units for additional study (Brown and Smiley, 1978). In Brown, Day and Jones (1983), grade 5 students again showed only a rudimentary ability to include the most important and delete the least important idea units from their summaries. Garner et al (1986) and Garner and Gillingham (1987) found that in verbal reports, grade 7 students demonstrated an understanding of unity of meaning, or topical coherence, as the basis of paragraphs to a significantly greater extent than grade 5 students, who were more likely to define paragraphs in terms of physical cues such as indentation. This difference was also evident in performance, since grade 7 students were significantly more able to reject sentences not conforming to the specific topic of a paragraph than children in grade 5. Since children in grade 5 are beginning to develop skills associated with macroprocessing, and so could potentially benefit from training, but at the same time are less skilled than grade 7 students, children in grade 5 were chosen for investigation in this study.

On the basis of some of the successful studies referred to in this chapter, it was anticipated that grade 5 students could benefit from training focused on macrostructure formation. For example, Adams et al (1982), Taylor (1982) and Armbruster et al (1987) taught children of this age to make use of explicit structural cues to identify important information. Kurtz and Borkowski (1987) taught children in grades 4 to 6 to select and invent topic sentences. It was also anticipated that children of all ability levels could benefit, not only those identified

as poor comprehenders. Studies by Palincsar herself, as well as Stevens (1988), Borkowski et al (1988) and Wong et al (1986) showed that below average readers improve their comprehension as a result of training focused on the identification of main ideas. However, studies using children of varying levels of reading ability have shown no interaction between treatment and ability (Taylor, 1982; Armbruster et al, 1987; Berkowitz, 1986; Baumann, 1984; Davey and McBride, 1986), suggesting that average and above average comprehenders can also benefit.

In Palincsar and Brown (1984) a small group of students was selected according to criteria of adequate decoding and poor comprehension. In spite of the relative homogeneity of this sample, some students made impressive gains in comprehension but one student in particular did not, for no apparent reason. An intact class is likely to contain a more heterogeneous group of students, varying along many dimensions relevant to reading, so it becomes particularly important to understand which factors may limit students' ability to benefit from training. In addition to pretest comprehension scores according to a variety of measures, a number of other characteristics will be assessed, based on the theory to be outlined below, for use in a discriminant analysis of those students who benefit from training and those who do not.

In the second chapter, Daneman (1984, p. 368) was quoted as asserting *"...that individual reader differences influence the integration processes of reading comprehension and that one crucial source of reader differences is processing efficiency in working memory."* Sources of difference in processing efficiency include slow word encoding and lexical access, which are hypothesised to reduce the reader's ability to integrate successive segments of the text and to construct a macrostructure. Perfetti (1984) and Stanovich (1980) also link processing in-

efficiency at the word-level to reduced working memory capacity available for higher-level comprehension processes. The studies referred to in the second chapter show a clear trend for poor readers according to standardised tests to have less efficient word encoding skills. However, Walczyk (1993) considers that beyond a certain threshold of skill readers can compensate for slower processing and comprehension is not affected. To examine the possible effect of word encoding efficiency on the ability to benefit from comprehension training, oral reading speed and accuracy will be recorded as an approximate measure of this efficiency. In addition, children's knowledge of words used in test passages will be assessed in order to examine whether poor vocabulary development could be a reason for failing to benefit from training.

Several studies in the first chapter indicated that the comprehension and recall of texts is greater for children who have a high level of background knowledge about the text topic than for children who have a low level of such knowledge (Pearson et al, 1979; Lipson, 1983; Recht and Leslie, 1988; Schneider et al, 1989). Fincher-Kiefer et al (1988, p. 426) explain this influence of subject knowledge as follows: *"... domain knowledge enables the individual to organize relatively large amounts of domain-related text, and via such organization retrieval structures are developed."* If knowledge makes possible this organisation of information, it is not surprising that in the absence of domain knowledge the reader is unable to make use of explicit structural cues in the text (Roller, 1990); in spite of the cues, the reader will not understand how the ideas are interrelated. Vosniadou et al (1988) also suggest that without adequate background knowledge readers cannot establish a coherent mental representation, in the context of which internal inconsistencies would become apparent. As a result, the ability to monitor

comprehension is impaired. Therefore, knowledge of the topic of each comprehension passage will be assessed in order to examine whether poor knowledge about these domains could limit gains from a training programme which emphasises the macrostructure formation and comprehension monitoring mentioned above.

Motivational factors are also considered an important influence on reading performance (Paris et al, 1983; Borkowski and Kurtz, 1987), in particular, whether students will be motivated to make the effort involved in learning and applying strategies. Several authors have considered attributional style in relation to underachieving students (Butkowsky and Willows, 1980; Johnston and Winograd, 1985; Carr et al, 1991). However, Paris and Oka (1986b) investigated motivational characteristics more generally in normal classrooms by assessing attitudes to reading in terms of personal enjoyment and perceived difficulty, the importance of reading, and personal perception of reading ability. A revised version of this survey will be used in order to examine whether negative attitudes associated with reading could limit training gains.

The intention in this study is to improve children's ability to retain important ideas from expository texts and to determine 'what a text is about' in a global sense. To this end, children will be taught to attend to explicit structural cues in the text and to construct their own macropropositions. There are sound theoretical and empirical reasons for focusing training on macrostructure formation. Kintsch and van Dijk (1978, p. 365-6) emphasise the importance of a macrostructure to establish a global interpretation of a text when they state: *"Relating propositions in a local manner is not sufficient. There must be a global constraint that establishes a meaningful whole, characterized in terms of a discourse topic."* The importance of thematic titles in establishing a coherent interpretation of a

text is demonstrated in studies such as Bransford and Johnson (1972), Dooling and Mullet (1973), Schwarz and Flammer (1981) and Schallert (1976). Lorch and Lorch (1985, p. 137) describe the importance of a macrostructure in another way when they state: *"Because most texts are organized around a small set of topics, an important component of text processing involves identifying a text's major topics and their interrelations."* Loman and Mayer (1983) have shown that when readers are assisted to identify important ideas and their connections through the use of signalling, they retain more important information and are more able to creatively solve problems related to the topic of the text.

However, B. Taylor (1986) has shown that when students in grade 6 read from their content area texts, they form almost no mental gist. While adults in some studies have been shown to attend to headings and subheadings (Holley et al, 1981) and spend longer processing topic sentences than detail sentences (Lorch, Lorch and Matthews, 1985), children do not always make use of such explicit structural cues. For example, Williams et al (1981) found that grade 5 students only made use of topic sentences to help them select or generate paragraph titles when the topic sentences were highlighted, leading these authors to conclude that *"...readers have to learn what to pay attention to and this has not been accomplished by the middle grades"* (p. 863).

The studies referred to in the second chapter demonstrate a clear pattern; more advanced readers in terms of age or skill perform better on tasks which involve global comprehension of text. Such tasks involve organising sentences into coherent paragraphs (Danner, 1976); selecting or generating paragraph titles (Williams et al, 1981), generating paragraph titles and topic sentences (Bridge et al, 1984), rejecting sentences not related to the specific topic of the para-

graph (Garner et al, 1986; Williams, 1984) and generating topic-specific follow-on sentences (Hiebert et al, 1983; Englert and Hiebert, 1984). In addition, more advanced readers are more able to assess the importance of idea units (Brown and Smiley, 1977; 1978; Smiley et al, 1977; Grabe, 1980; Ohlhausen and Roller, 1988), include more important information in their summaries (Brown et al, 1983; Winograd, 1984), make better use of summarisation rules (Brown and Day, 1983) and make better use of conventional text structures to organise their mental representation of a text (Meyer et al, 1980; McGee, 1982). However, while these differences have been clearly demonstrated, it is not the case that only younger or weaker readers have room to improve their global comprehension. Even adults do not perform perfectly, for example when selecting or generating titles (Williams et al, 1981), rejecting sentences not consistent with the specific paragraph topic (Hiebert et al, 1983), or applying the full range of summarisation rules (Brown and Day, 1983).

In view of this theoretical and empirical evidence, the first two major strategies to be taught in this study focus on encouraging macrostructure formation through the use of both explicit structural cues and constructed macropropositions. In the first strategy, described as Collecting Ideas, children will be taught to pay special attention to headings and subheadings by asking themselves the questions: What do I know about this topic? What might this section be about? Prediction was not identified as an important component of reading within the theoretical models discussed in chapter one. Also when college students rated strategies associated with different types of reading in Lorch, Lorch and Klusewitz (1993), anticipation of what is ahead in the text was identified more with personal choice reading than school reading. However, these questions are not used primarily to generate

predictions, but rather to suggest to students a specific way of devoting additional processing to important structural cues. If students were given a more general direction, such as: Pay special attention to the headings, they may have little idea of the kinds of mental processes to engage in. The questions suggest a specific way of thinking about headings, as well as activating background knowledge associated with the topic.

In the second major strategy, described as Putting Ideas Together, the children will be taught to generate a summary sentence for each paragraph and to link this sentence to the section heading. Children will again be taught to ask themselves questions: What is the main idea of this paragraph? How does this fit in with what I've already read? These strategies provide a means of forming the mental gist which B. Taylor (1986) found to be lacking in the reading of grade 6 students.

While these cognitive strategies can directly enhance global comprehension, they may also less directly enable children to develop more sophisticated concepts of the purpose of reading instructional text and what it means to understand. Brown (1980) notes that 'understanding' is not an all or nothing concept, but involves a matter of degree. Knowing all the words in a text is one level of understanding, making sense of each individual sentence is another. Through the explanation, demonstration and repeated guided use of the macrostrategies described above, children are presented with a broader model of understanding and have the opportunity to expand their own concepts of understanding to include comprehension of global relationships. Several studies have shown that poor readers have limited goals in reading, focused largely on decoding skill and word knowledge (for example, Gambrell and Heathington, 1981; Garner and Kraus, 1981-2). The questions readers are taught to ask themselves in this study focus

on understanding important ideas and their interrelations and, if internalised, may provide the basis for children to set more sophisticated goals for their own reading.

The importance of reading goals has been emphasised several times in the preceding chapters. Within the Just and Carpenter (1980) model of reading, goals influence how often and how thoroughly wrap-up operations are carried out, that is, they set standards which have to be met before the reader moves on to the next segment. The following observation from Adams et al (1982, p. 49-50) suggests that the children in their study had not only acquired new strategies, but may also have set more stringent goals for their reading.

> *"...instead of reading nonstop through the passages, the students were overtly using strategies that increased their attention to the material. Students, to a varying degree, attended to the subheadings, stopped to reread, reviewed after reading, or took notes."*

If students set more stringent goals for reading, they are also more likely to become aware of comprehension problems. Baker (1984b) notes that being able to evaluate the state of one's comprehension is an important component of effective reading. Yet many studies were referred to in chapter two to show that less advanced readers monitor their comprehension less effectively; younger and weaker readers are less likely to detect deliberately inserted errors which interfere with a full and complete understanding of a text. Even adult readers are short of perfect in this area (Baker, 1979; Baker and Anderson, 1982; Baker, 1985b; Grabe et al, 1991; Zabrucky and Commander, 1993).

The paragraph summarisation strategy provides an initial means of checking

comprehension. As noted by Flavell (1981), cognitive actions can lead to metacognitive experiences. Thus, children in this study will be taught that inability to summarise (a cognitive action) is a sign that a paragraph has not been well understood (a metacognitive experience), and so they should reread the paragraph carefully. However, additional questions to encourage comprehension monitoring will also be included under the heading of Checking and Clearing Up. As the final major strategy, the children will be taught to ask the questions: Are there any words I don't understand? What can I do about it? Do all the ideas make sense? Is there anything that isn't explained?

As Markman (1985) notes, adults do not generally ask themselves the question: Do I understand? Their feelings of puzzlement usually arise as a by-product of active attempts to understand. However, in the case of children, explicit questions are included to encourage them to think more carefully about their level of understanding, with the intention that eventually the monitoring of comprehension becomes a normal part of their reading. Questions are deliberately focused on both words and ideas. Garner (1980; 1981) indicates that poor readers tend to focus on long words as a source of comprehension difficulty, so separate questions about ideas are included to encourage children to also think about unclear or incomplete ideas as obstacles to comprehension.

As described by Palincsar and Brown (1984), effective instruction should make overt the cognitive processes carried out by skilled readers and carefully guide learners in approximating these processes for themselves. However, as learners become more skilled, this guidance should be systematically reduced so that the learners are eventually able to regulate their own processing. In the work of Palincsar, this support and gradual ceding of responsibility is provided in the

reciprocal teaching dialogue. However, it was pointed our earlier (for example, Palincsar, 1987) that considerable in-service training is involved in teaching the method to teachers, and even then some find it difficult.

There is some evidence that the type of skills involved in the Palincsar programme can be taught in a written format (Palincsar and Brown, 1986a), or also via a computer programme (Stevens, 1988; Salomon et al, 1989). In the Australian context, some teachers work in isolated rural schools where lengthy in-service training is not practical. Hence it would be desirable to produce simple material which could be used by teachers to guide students in the manner described.

For the purposes of this study, students will be provided with a workbook divided into three phases to accompany their reading book. In the first phase, all the questions to be asked are set out, and those related to Putting Ideas Together and Checking and Clearing Up are followed by multiple-choice options. Initially the instructor models the decision making process in choosing the best option, for example, a paragraph summary statement, explaining why the chosen option is a good one and why the others are rejected. All students mark what they think is the best option for each question, so all students are actively involved in the task, even in a large class. Once the process is understood, students take over the role of justifying their choices and explaining why other options are inadequate. In the second phase, students are still prompted by written questions, but they now all generate their own responses. Again the instructor models and explains the process, until students can take over this role for themselves. In the final phase, neither questions nor answers are printed in the workbook; students are required to supply both once they have seen the instructor modelling the process

of asking and then answering questions. In the course of the procedure, students will hear many good responses to the questions and many explanations as to why some responses are not ideal, so they have the opportunity to develop a concept of good responses from these positive and negative instances. They also have the opportunity to evaluate their own performance in applying the strategies in the discussion following each paragraph.

It was mentioned in the introduction to this chapter that informed, self-control training is important to encourage transfer and maintenance of strategies. Students need to know not only what the strategies are and how they work, but also when and why they are useful. Students also need to be able to judge how effectively they are using the strategies. Paris and Oka (1986a, p. 104) draw attention to the importance of motivating students when they state: *"Children must be convinced that the strategies are important, worth some extra effort, and instrumental for enhancing performance. These beliefs motivate students to use the strategies without prompting or help."*

Prior to the intervention, the strategies will be explained to students in the context of concrete examples, where students can see not only what the strategies are, but also how they enhance comprehension. To further emphasise their usefulness, students will complete regular comprehension quizzes, the results of which will be entered on each student's individual progress chart, as well as on a class average progress chart, as in the Palincsar studies. Students should be able to see improvements in their performance, which the instructor can link to the students' efforts and use of the strategies in order to motivate them to continue their strategic behaviour even in the absence of prompting.

To fully assess the effectiveness of a training programme, tests should require

spontaneous strategy use, transfer to tasks other than those used during training, and maintenance over time. Hence, students will not be prompted to use the trained strategies during testing, and posttesting will be undertaken both immediately and again ten weeks after training. Tests to be used include comprehension tests of the same format as those used during training, a standardised reading test, a cloze test and an error detection test, the last three to evaluate the generalised effect of training.

The comprehension tests as used by Palincsar are a common form of assessment in classrooms; teachers frequently set questions to assess how well students have understood a passage, so this task has some ecological validity. Questions are based on the three types identified by Pearson and Johnson (1978), namely text explicit, text implicit and script implicit questions. To answer the three types of questions, respectively, students have to understand single sentences, to link more than one sentence, and to integrate prior knowledge with text information. A standardised reading test is included as a measure with known reliability and validity, and also as a test with different content and a different format to the training materials. A cloze test with every fifth word deleted is included as a measure of intersentence integration and also as a comparison to the results of Paris (Paris and Jacobs, 1984; Paris, Cross and Lipson, 1984; Paris and Oka, 1986). As will be explained in the results section, in hindsight this type of test was not a wise choice. Finally, an error detection test is included to assess comprehension monitoring. The errors involved are explicit internal inconsistencies, with the inconsistencies spread across more than one sentence. Markman (1979) notes that constructive processing is required in order for an internal inconsistency to be noticed; successive sentences have to be integrated and compared. Several studies

cited in chapter two show that internal inconsistencies are particularly difficult errors to detect, even for adults (Zabrucky and Ratner, 1989; Grabe et al, 1991; Baker, 1985b; Garner and Kraus, 1981-2; Reis and Spekman, 1983). Thus, the use of an error detection test based on internal inconsistencies is a particularly stringent test of constructive processing and comprehension monitoring.

Chapter 4

Design of Materials and Method

4.1 Design and Selection of Materials

Since Australian students do not use anything equivalent to American basal readers for reading instruction, all materials for this experiment were specifically written for the purpose. Further reasons for writing material were to avoid problems with copyright and to ensure that the texts were at at a level appropriate to grade 5.

The information for these materials was obtained from the children's section of the State Library of South Australia. The topics chosen were among those most frequently discussed in the reference books for children in the upper primary grades. All materials were then written to slot into the grade 5 range on Fry's Readability Graph. All texts were expository and explained various natural phenomena.

The following materials were required:

1. a reading booklet containing texts to be read during class instruction;

2. a work booklet for students to record their responses during class instruction;

3. nine texts for individual comprehension practice during training;

4. nine texts for comprehension tests (three tests at each of three times of testing);

5. three cloze tests (one at each of three times of testing);

6. three sets of error detection tests (one set at each of three times of testing);

7. a survey of opinions about reading (the same survey at each of three times of testing).

4.1.1 Reading booklet

As explained in the conclusion to the previous chapter, instruction of the class was divided into three phases according to the amount of guidance given to students, with each phase intended to last one week. Two passages were written for each phase to ensure that sufficient material would be available. However, time permitted only four of these texts to be used during instruction, these being *Life on the Great Barrier Reef, Natural Disasters, Animal Defence Systems* and *The Age of Dinosaurs.* These four texts are reproduced in Appendix A.

Table 4.1 sets out the titles of the texts which were written, their length, and the number of sentences and syllables per hundred words, the last two figures indicating that the passages were at grade 5 level according to Fry's Readability Graph.

Table 4.1: Teaching Texts

Title	Words	Sent/100w	Syll/100w
1. Life on the Great Barrier Reef	1031	8.0	129.1
2. Natural Disasters	975	7.8	130.3
3. Animal Defence Systems	1368	7.9	127.5
4. Endangered Animals	969	7.9	130.8
5. Migration in Nature	1348	7.9	128.8
6. The Age of Dinosaurs	905	8.0	129.8

In neither the teaching material nor any of the individual tests was any attempt made to write paragraphs according to a particular pattern. In other words, paragraphs were not always of the same length, they did not necessarily begin with a topic sentence, nor were they organised according to the same text structure. For example, many paragraphs in the passage *Life on the Great Barrier Reef* were descriptive, while many paragraphs in *Animal Defence Systems* were of a problem-solution type, explaining what different animals do to avoid the problem of being attacked by predators.

In the previous chapter it was noted that a majority of texts written for both adults and children do not contain topic sentences in the initial position (Winograd and Bridge, 1986). Hare et al (1989) observed that even children as advanced as grade 11 found it difficult to summarise paragraphs which did not contain topic sentences in the initial position and which were not of a simple listing (enumeration) structure. These authors suggest that if training texts include only well structured paragraphs with an enumeration structure, students will not learn

to transfer their skills to naturally occurring types of texts, so the latter should be included in training.

The teaching texts were produced in larger print, 14-point print instead of 12-point, to ensure that the type was not difficult for grade 5 students to read. This amendment was suggested by the teacher of the experimental class prior to the experiment. The teaching texts were stapled into booklet form. Each paragraph was numbered for easy reference and to assist coordination between the reading and the work booklet, to be described below.

4.1.2 Work booklet to accompany teaching texts

In the previous chapter, reasons were given for producing a written work booklet rather than conducting an oral dialogue as in the work of Palincsar. The first of these reasons was to produce a framework to guide instruction which could potentially make the method easier to use for classroom teachers. The second was to ensure a maximum of academically engaged time (Rosenshine and Stevens, 1984), since all students would produce a response for each text segment.

An introduction to the teaching method and the worksheets for each instructional text were printed in booklet form. Material relating to the four texts used during training is reproduced in Appendix B.

Three main strategies and associated questions were taught to children, these being:

1. **Collecting Ideas:** What do I know about this topic? What might this section be about?

2. **Putting Ideas Together:** What is the main idea of this paragraph? How

does this paragraph fit in with what I've already read?

3. **Checking and Clearing Up:** Are there any words I don't understand? What can I do about it? Do all the ideas make sense? Is there anything that isn't explained?

The first strategy was used in relation to all headings and subheadings, while the other two were used after each paragraph.

As was explained in the previous chapter, if students are to independently use the trained strategies, they need to understand not only what the strategies are, but also why they are useful. Therefore, the first part of the work booklet (see Appendix B) introduced the nature of and reason for each strategy via a series of examples and questions. The intention was for the instructor to work through this introduction with the children and to encourage discussion of the strategies and their value in relation to the concrete examples provided. Of course, this could only be an introduction, and it was anticipated that understanding of these points would be consolidated during the course of the experiment through continued use and discussion of the strategies. As will be explained in more detail in the *Procedure* section, this introduction proved to be somewhat long and complex for grade 5 children, and would need to be modified for future use.

The work booklet was organised into three sections, with the amount of guidance given to students decreasing from one section to the next. Each section was intended to cover one week of teaching. For the two passages in the first section, students were provided with the questions to prompt strategy use, and with possible responses in the case of two of the three strategies. Lines were provided beneath each heading and its associated questions for students to write

their predictions (Collecting Ideas). For each paragraph, students were provided with multiple-choice options from which to choose the best paragraph summary statement and how the paragraph related to its preceding section (Putting Ideas Together). Students circled what they thought was the best option. The distractors included errors thought most likely to occur, that is, a focus on only one salient point, sentences that were rambling, or lists of facts rather than summaries. For the final strategy, the work booklet listed words and ideas from each paragraph that could potentially require explanation (Checking and Clearing Up). Students underlined any of these suggestions that they considered unclear, and wrote any additional unclear words or ideas on the line provided.

The initially high level of guidance was provided for several reasons. The structured format provided a simple framework to follow in guiding students and encouraging them to discuss their responses. The correct alternative in the multiple-choice sections provided a model of what good summary and integrating sentences should include. The distractors provided an opportunity for discussion of common errors. Thus, the intention was to develop in grade 5 students, not accustomed to dealing with general summary statements, a clearer notion of the goal they were aiming for through careful guidance for one week. The list of potentially unclear words and ideas was intended to accustom students to a closer examination of their understanding.

For the two passages in the second section, the work booklet still presented prompting questions to remind students of the strategies, but students wrote their own responses on the lines provided. As indicated previously, time permitted only one of the passages to be studied by the class (*Animal Defence Systems*). For the two passages in the third section, the work booklet included only headings

and paragraph numbers with lines on which to write responses. Again only one of these passages was used during instruction (*The Age of Dinosaurs*). Students were now required to ask themselves the questions practised during the previous two weeks, and to write down their own responses. It was anticipated that by this time they would be capable of using the reading strategies efficiently and independently since the same questions had recurred continually during weeks one and two and students had been presented with models of good responses and had practised writing their own.

4.1.3 Comprehension tests and practice passages

Twenty-four texts were written as a pool from which to choose the eighteen passages required for comprehension tests and individual practice pieces. The final selection was made on the basis of the pilot testing, to be described in the next section. Texts eventually chosen as practice pieces are reproduced in Appendix C, and those chosen as comprehension tests are reproduced in Appendix D.

Material for these texts was based on similar sources to the training texts, but did not cover the same topics. They varied in length from 361 to 371 words, and were written within the grade 5 range according to Fry's Readability Graph. The length was reduced for these grade 5 children from the 400 to 450 word texts used by Palincsar for students in grade 7. Three comprehension tests were planned for each time of testing in order to reduce possible effects on overall performance due to specific passage content or style.

Ten comprehension questions were written for each of the twenty-four texts. The questions were selected to cover all sections of the text and to include four text-explicit, four text-implicit and two script-implicit questions (Pearson and

Johnson, 1978). Below is a paragraph from one of the passages, followed by an explicit, implicit and script question, in that order.

Each nest has one queen who lays all the eggs. Most ants in a nest are workers. They are females who can't produce eggs, but who do all the housekeeping jobs. The most important job of the workers is to collect food and to look after the queen and the young ants. Workers have several ways of making sure that as much food as possible gets back to the nest.

1. *Who do the workers look after?*

2. *What sort of ants are the workers?*

3. *Why do you think ants store food, but most other insects don't?*

All the information necessary to answer the first question can be found in one sentence, while the second question bridges two sentences. To answer the third question, students need to reflect on what they have read and any prior knowledge they have in order to make an inference.

For each text, ten questions dealing with word meanings and ten dealing with prior knowledge of the topic were written. As explained at the conclusion of the previous chapter, the intention of these tests to accompany each comprehension test was to examine whether poor vocabulary knowledge or poor subject knowledge could be factors limiting the effectiveness of training. Questions accompanying texts eventually chosen as test passages are included in Appendix D.

For the vocabulary test, words which intuitively seemed the most difficult ones in the passage were chosen for the test. Technical terms explained in the text were not included, and there was no overlap between words included on different tests. The following is an example from the passage on ants:

She liked the scent. Scent means:

a) present

b) parcel

c) smell

d) I don't know

For prior knowledge tests, some questions were based on information later to be presented in the text, but between three and five questions on each test questioned information associated with, but not mentioned in the text. The following are examples from the test on ants, the first one covering information later to be read from the text, the second covering only associated information.

How do ants tell each other where food is?

a) they have special dances to tell other ants where to go

b) they make a smelly trail for other ants to follow

c) they send out special sounds from their antennae

d) I don't know

Why don't queen ants get their own food?

a) they are too big to run around

b) they are too busy giving orders

c) they think queens shouldn't have to work

d) I don't know

Both the vocabulary and prior knowledge tests were written in the above multiple-choice format. This format was chosen to provide a simple and objective way of scoring responses. It was considered that definitions of words and the explanations sought in these knowledge tests could be difficult for grade 5 students to state succinctly. Poor expression in responses would then make them difficult to score. The multiple-choice format was preferred, even though it introduces the possibility of guessing. For this reason, the *I don't know* option was included, and it was stressed to students that they should only select a specific option if they were fairly certain it was correct. The *I don't know* option was always option d), while the correct option was randomly assigned to one of the other three positions.

A simple practice example was written at the beginning of both the vocabulary and prior knowledge tests, and this same example was printed on all the tests. The examples were used to explain the procedure to students, and on later tests to remind them of the procedure.

4.1.4 Transfer tests

At the end of the previous chapter, the importance of assessing comprehension gains on tasks other than those used during training was noted. A cloze test and an error detection test were included for this purpose, as explained in the previous chapter. All transfer tests eventually chosen for use in the main experiment are reproduced in Appendix E.

Five passages were written for the cloze tests, with the three to be used for the main experiment determined by the pilot testing to be described in the next section. Passages were based on the same sources as other materials and

included topics about plants, animals and physical phenomena. Passages ranged in length from 120 to 128 words and were at the grade 5 level according to Fry's Readability Graph. The first sentence in each text was left intact, and from there every fifth word was deleted until twenty spaces were created. Each space was marked by an unbroken line on which students were to write a single word which was semantically and syntactically congruent with its context.

Each cloze test began with the same three practice examples. They were designed to show students that they must read whole sentences, look back to previous information, and look forward to subsequent information. On later tests the examples were intended to remind students of this procedure.

For the error detection tests, fifteen passages ranging in length from 125 to 132 words were written. The topics again included plants, animals and physical phenomena, but there was no overlap between specific topics on any of the test material. For the pilot test, the passages were divided into groups of three in such a way that each group included a range of different topics, for example, flowers, polar bears and fossils. As in the case of the comprehension tests, a set of three different topics was used in an attempt to reduce any effects due to specific passage content and, in the case of the error detection test, to increase the number of items on the test.

Each passage contained two text-based errors which contradicted information given earlier. A variable amount of information intervened between the fact and its contradiction, but the two never occurred in the same sentence. The passages were written primarily to read well; the relative difficulty of detecting the various errors was to be established during the pilot testing. All errors were of a substantive rather than syntactic nature. The following is an example of a

contradiction occurring in a paragraph about starfish.

Starfish have lots of small tubes and suckers on their arms to let them crawl along the ocean floor. This is the only way they move about. If they lose one of their arms, they can grow another one in its place. They use these arms to swim through the water. Starfish come in many different colours.

Each group of error detection passages was preceded by the same two practice examples to illustrate the procedure. Both examples consisted of two sentences containing one explicit text-based error to alert students to the kind of problem they were looking for. The same examples were included on all tests to remind students of the procedure.

4.1.5 Standardised test of comprehension

The use of a standardised comprehension test allows any comprehension gains as a result of training to be measured with an instrument of known reliability and validity. Since its content and format is different to the training tasks it is also another measure of transfer of learning. However, as pointed out by Paris, Cross and Lipson (1984), since such tests are usually speeded, there is the potential difficulty that students will not have time to use the reflective strategies they have been taught and hence the standardised test will not show the comprehension gains that may be apparent from other, unspeeded measures of comprehension. This possibility will be investigated in the main experiment.

The Primary Reading Survey is a standardised test devised by the Australian Council for Educational Research (ACER) and normed under Australian conditions. The test consists of seven passages, each made up of several paragraphs. Multiple-choice questions with four alternatives are printed next to each passage.

According to the Teacher's Handbook, 33% of the questions are fact questions, requiring an answer directly from the text. The remainder of the questions require students to go beyond the information given. For example, on Form S of the test in a passage entitled *Poor Dog*, students are asked what the weather was like. They are required to infer from a reference to snowflakes that the weather was wintry. The passage does not state explicitly why a stranger took the lost dog home, but students are asked to infer that he felt sorry for the dog. The test includes 39 questions, and a time of 30 minutes is allowed for its completion. Level C is designed for grade 5 students, and two alternate forms, Form R and Form S exist. The internal reliability (Kuder-Richardson KD-20) is given in the Teacher's Handbook as .92 and .93 for Form R and Form S respectively. No information is given on test-retest reliability for either the same or alternate forms.

4.1.6 Survey of Attitudes to Reading

The survey was designed to assess students' attitudes both towards reading in general, and to their own abilities as readers. The idea was based on a similar survey used by Paris and Oka (1986b). It was included to determine if negative attitudes to reading or a low estimation of competence could be one of the factors limiting the effectiveness of the training, and also if the attitudes of students changed as a result of training. The survey is reproduced in Appendix E.

Ten statements were written for the survey, relating to the interest, enjoyment and importance of reading, as well as opinions about current and future skill levels. To simplify the decision process for grade 5 students, only a four-point rating scale was used. Students were required to rate each statement as either

very true, sort of true, not really true, or not at all true. Two simple practice examples were printed on the survey form to demonstrate the two-stage decision process. Students were directed to first consider whether the statement was true or untrue, and then how true or untrue, before placing a tick in the required box. Only one survey form was designed, and this same form was used for the different times of testing during the main experiment.

4.2 Pilot testing of test materials

Although materials were written on the basis of reference books designed for children, there was no way of knowing *a priori* how difficult grade 5 students would find the passages on various topics. Any floor or ceiling effects on tests in the main experiment could mask comprehension gains as a result of training. Hence the pool of comprehension and transfer tests described above was pilot tested in order to select pieces of an appropriate level of difficulty, and to become aware of any items that were poorly worded or ambiguous.

Contact was made with two adjacent schools which were large enough to have two similar grade 5 classes. They were in the northern suburbs of Adelaide, one in a lower middle class area, the other in a mainly middle class area. The pilot testing was carried out in the week before the main experiment. This timing was not ideal, but since the whole experiment could only be carried out by the experimenter being granted a short-term release from teaching duties, there was no alternative. The short time available for pilot testing meant that the reliability figures reported below had to be calculated later from the performance of children in the main experiment rather than during pilot testing.

Test materials required for the main experiment included nine comprehension tests with associated vocabulary and knowledge tests, three cloze tests, and three sets of error detection tests. To cover these materials and to have the children in each of the four classes spending a similar amount of time on the pilot testing each day, the tests were allocated as follows. One randomly chosen class completed one cloze test and one error detection test on each of five days. The remaining three classes completed a different comprehension test with associated vocabulary

and knowledge tests on each of five days. The fifteen comprehension passages to be pilot testing were chosen from the available pool of twenty-four in such a way that a range of topics from the fields of plants, animals and physical phenomena were covered.

It was explained to students that the work they were about to do was part of on-going research into the reading skills of grade 5 students. In particular, the notion of pilot testing was explained to them via the analogy of testing new aeroplanes. It was made clear to them that the work was not a test of individual ability. Rather, it was to establish whether the material was too easy, too hard, or just right for grade 5 level. To reinforce this idea, students were asked to express their opinion about the level of difficulty at the end of each session. Students were very cooperative, and no-one appeared anxious.

4.2.1 Comprehension tests

Three classes each completed five comprehension tests. Tests chosen for the main experiment on the basis of this pilot testing are reproduced in Appendix D, and tests used for individual practice during training are in Appendix C.

Most of the students in the three classes were of Anglo-Australian background. Class A contained twenty-eight students, with fifteen girls and thirteen boys. Four were of European and one of south-east Asian background. Class B contained twenty-five students, with fourteen girls and eleven boys. Three were of European and three of south-east Asian background. Class C contained twenty-four students, with sixteen girls and eight boys. Five students were of European background. Although the classes contained students of varying abilities, the teachers did not indicate that anyone was receiving remedial teaching or had difficulty

with English.

On each day for one week these classes read one of the comprehension passages and completed the associated questions and vocabulary and prior knowledge exercises. Each session began with the vocabulary and prior knowledge tests. On the first day the procedure was explained via the practice examples on the test. Students had no difficulty completing these examples. Once these tests had been completed, the reading passage was handed out. When students were satisfied that they had read the passage carefully, these sheets were collected and the questions were handed out. There was no time limit on any of this work.

Comprehension questions were scored correct if responses contained all the relevant information from the passage. If only some of this information was given, only half a mark was awarded. Script questions were marked correct if a clear and plausible response was given.

As a result of this pilot testing, three passages were eliminated because they proved to be too difficult. Two dealt with physical phenomena, namely cloud formation and mechanisms of water retention in desert plants. Rather surprisingly, a passage describing the senses used by snakes also proved difficult. Students in general did not understand the workings of heat sensors in pythons.

From the remaining passages, three were chosen from each class. The passages selected from each class were allocated to three different groups in such a way that the groups were as closely matched as possible. The three groups were then randomly allocated to one of three times of testing. Table 4.2 shows the mean scores and standard deviations for these passages, and their random allocation to a particular test schedule for the main experiment.

Three texts on different topics were included at each time of testing to reduce

Table 4.2: Pilot test of comprehension tests (max. score 10)

Test Schedule	Title	Mean	S.D.
Pretest	Why Lions Live in Groups	4.9	1.5
	Spiders as Hunters	5.5	2.4
	Carnivorous Plants	3.1	2.2
Posttest	The Ant Colony	5.1	1.6
	The Moon	4.0	2.3
	Fungi - the Strange Plants	4.3	2.3
Delayed Posttest	The Earth and Other Planets	4.4	2.4
	A for Aardvark	5.3	1.5
	Why Bees Dance	3.9	1.9

effects due to specific passage content or style. As the table shows, some of these topics were more difficult than others. However, the overall averages of the three groups were similar, being 4.47, 4.43 and 4.56 for the pretests, immediate posttests and delayed posttests respectively. These tests are reproduced in Appendix D.

Three passages that proved to be at an appropriate level, but were not selected for the tests in the main experiment, were retained as practice pieces. On the other hand, three passages from those not pilot tested were deleted. The pilot testing showed that grade 5 students have difficulty understanding physical phenomena, so passages dealing with cave formation, the origin of fossil fuels, and space travel were dropped, the latter mainly because it overlapped test topics.

The nine passages to be used as practice exercises during the main experiment are reproduced in Appendix C.

4.2.2 Transfer tests

Class D contained thirty students, with sixteen boys and fourteen girls. Four students were from a non-English speaking background, with two of European and two of south-east Asian background. This class completed one cloze test and a set of three error detection passages daily for one week.

On the first day, the procedure was explained via the examples printed on the test. It was explained that in the cloze test they should read whole sentences before filling in words. If this process did not indicate a suitable word, they should look back to previous information, or forward to subsequent information. No-one had any difficulty completing the practice examples. On the second day, the class was reminded of the three "rules" and worked through the same examples again. Subsequently they were told only to look at the examples to remind themselves of the "rules".

The process of finding errors was also explained via the examples on the test. Both examples contained clear text-based errors, which students easily located. They were instructed to underline only those words or phrases they knew for sure could not be correct. This point was emphasized with reference to the examples in an attempt to focus attention on text-based errors, rather than prior knowledge violations. As with the cloze test, students were told to work through the examples on subsequent tests to remind themselves of the procedure. There was no time limit on either kind of test.

On the cloze test, any words which were both syntactically and semantically

appropriate were accepted as correct. For example, in a sentence intended to read: *Once the eggs are laid, the female's job is finished*, the words *done* and *over* were accepted instead of *finished*.

On the error detection task, a hit was scored if all the words in a problematic phrase were underlined. If fewer words were underlined, only a half mark was awarded, but no marks were deducted if slightly more than the problematic phrase was underlined. For example, in a passage explaining that sloths lived in trees and ate leaves, the whole section in bold italics had to be underlined: *Sloths like to eat grass and juicy berries on the ground.* Any words that were underlined but were not part of the underlining of a problematic phrase were scored as a false alarm.

From the five cloze tests, the three with the most closely matched scores were chosen for inclusion in the main experiment. The passages with the highest score (mean 14.4) and the lowest score (mean 10.1) were deleted. Table 4.3 shows the means and standard deviations for these three tests, and their random allocation to a particular test schedule for the main experiment.

Table 4.3: Pilot test of cloze tests (max. score 20)

Test Schedule	Title	Mean	S.D.
Pretest	The Mallee Fowl	13.1	3.58
Posttest	Shooting Stars	11.6	3.78
Delayed Posttest	Cuckoos - the Lazy Parents	11.6	3.39

From the fifteen error detection passages, those with the highest hit or false alarm rates were eliminated. From the remainder, nine passages were chosen and

allocated to a group of three in such a way that the averages for the three groups were as closely matched as possible. Table 4.4 shows the mean number of hits and false alarms for each of the groups and their random allocation to a particular test-schedule for the main experiment.

Table 4.4: Pilot test of error detection tests (max. hits 6)

Test Schedule	Title	Mean Hits (S.D.)	Mean F.A. (S.D.)
Pretest	Return of the Salmon		
	The Slow Sloth	2.32 (1.31)	2.16 (2.06)
	Seaweed		
Posttest	Sea Mammals		
	Fossils	2.33 (1.20)	2.38 (1.95)
	Pumas		
Delayed Posttest	Starfish		
	Water on the Earth	2.32 (1.28)	2.29 (1.58)
	Sea Snakes		

One cloze test and one group of error detection passages were printed together to constitute one testing session during the main experiment. These tests are reproduced in Appendix E.

4.2.3 Reliability of tests

As explained earlier, due to lack of time the reliability of tests chosen for the main experiment could not be calculated during pilot testing. Therefore the following figures were derived from the analysis of work done by the fifty children in the control and experimental groups of the main experiment. Measures of internal consistency were calculated according to the Kuder-Richardson KD-20 formula for the cloze, knowledge and vocabulary tests, where a response was either correct or incorrect. For the comprehension tests, attitude survey and error detection tests (hits) Cronbach's coefficient alpha was calculated.

Table 4.5: Reliability of comprehension and associated tests

Passage	Comprehension	Knowledge	Vocabulary
Why Lions Live in Groups	.71	.19	.59
Spiders as Hunters	.70	.20	.65
Carnivorous Plants	.68	.25	.51
The Moon	.78	.51	.69
The Ant Colony	.66	.38	.57
Fungi - the Strange Plants	.80	.47	.51
The Earth and Other Planets	.71	.34	.50
A for Aardvark	.60	.29	.61
Why Bees Dance	.72	.03	.47

According to Gronlund (1981), the internal consistency of teacher-devised tests is usually in the range of .60 to .85, which he considers adequate as a basis for routine classroom decisions. He suggests that a higher level of internal

Table 4.6: Reliability of cloze tests

Passage	
The Mallee Fowl	.86
Shooting stars	.78
Cuckoos - the Lazy Parents	.71

Table 4.7: Reliability of error detection tests (hits)

Passages	
The Return of the Salmon group	.70
Sea Mammals group	.66
Starfish group	.47

consistency would be required only for major decisions, such as deciding whether or not a student should be sent to a special school.

As Table 4.5, Table 4.6 and Table 4.7 show, all comprehension and transfer tests fell within this range, with the exception of the final error detection test. However, the internal consistency of knowledge tests was variable and generally low. It may be that the children knew only isolated facts about these topics, rather than being able to draw on a coherent body of knowledge. On the other hand, the tests could have failed to measure any interconnected body of knowledge which did exist. From the data available there is no way of determining which possibility is more accurate.

The coefficient alpha for the Survey of Attitudes to Reading, based on the pretest scores, was .82. Test-retest coefficients were also calculated for the control

group only, since there was no reason to assume that a change in attitudes would take place this class. The correlation between the pretest and the immediate posttest four weeks later was .70, while the correlation between the pretest and the delayed posttest fourteen weeks later was .97.

4.3 The main experiment

4.3.1 Subjects

With generally declining school enrolments, not all schools have two equivalent grade 5 classes. Therefore, size was an important factor influencing the choice of school. Contact was made with a Catholic school which satisfied both the size condition and was in an easily accessible location. Another advantage of this kind of parish school was the socio-economic background of students; it was mixed to some extent, but did not include extremes of wealth or poverty.

Both grade 5 teachers were willing to participate in the experiment. Each class contained 25 students. In class A, there were 13 girls and 12 boys. Their ages ranged from 9 years and 11 months to 11 years, with the average being 10 years and 5 months. In this class, 10 students were of Anglo-Australian background and 15 were of European background, primarily Italian.

In class B, there were 15 girls and 10 boys. Their ages ranged from 9 years and 11 months to 11 years and 5 months, the average age being 10 years and 6 months. In this class, 7 students were of Anglo-Australian background and 18 were of European background, again primarily Italian.

The experimental class was chosen on the basis of pretest results. All the tests showed that there was a different distribution of students in the two classes. Table 4.8 shows the distribution of students according to their percentile rank on the ACER test. One student was absent from each class.

Class B contained a cluster of very weak students on the one hand, and fewer very capable readers on the other. Since one of the aims of the experiment was to determine the effect of the teaching on different ability levels, class A was

Table 4.8: Number of students at different percentile ranks (ACER test)

Class	0 - 9	10-19	20-29	30-39	40-49	50-59	60-69	70-79	80-89	90-99
A	2	3	3	7	1	0	3	0	3	2
B	6	1	4	3	1	3	3	2	0	1

considered preferable as the experimental group.

As mentioned in the previous chapter, Lysynchuk, Pressley, d'Ailly, Smith and Cake (1989) include among their criteria for internal validity of studies that children be randomly assigned to treatment groups, and that all groups read the same materials and believe they are receiving a treatment. In this study children were not randomly assigned, but were left in intact classes, and an untreated control group was used.

If the normal instruction by the control class is to be interrupted, then there should be a good reason for doing so, both from the point of view of the children and their teacher, on whose goodwill access to the class depends. On the basis of studies cited in the previous chapter, there is no good reason to believe that the comprehension of the control class would improve as a result of simply being exposed to the materials used for training and practice.

Several studies show that an experimental class performed significantly better than a control class who had read the same materials and completed similar tasks without instruction (Stevens, 1988; Kurtz and Borkowski, 1987; Borkowski, Weyhing and Carr, 1988; Palincsar, 1987). More significantly, however, some studies have directly compared untreated control groups with control groups who

completed the same independent practice exercises during training as the experimental groups (Adams, Carnine and Gersten, 1982; Palincsar and Brown, 1984; Raphael and Wonnacott). In none of these studies was there a significant difference in posttest performance between the two control groups. Salomon, Globerson and Guterman (1989) compared two control groups who both read the same material as the experimental group, but one group only read the passages, while the other group answered literal and inference questions as well. Again there was no significant posttest difference between the groups. McCormick (1992) found no improvement in the ability to answer inferential questions when students read a passage and answered such questions once a week for twenty weeks. This lack of effect of independent seatwork is not unexpected considering criticisms made of the drill and practice model of instruction in the introduction to chapter three. For this reason, the control group was allowed to continue with normal classroom work to minimise disruption.

Given that the control group was to be untreated, it also made sense to leave the two classes intact rather than randomly assigning children to two groups. In this way the control class teacher could continue with his normal programme for the class. As the table above shows, both classes contained children varying widely in reading ability, although the distribution of scores in class B was somewhat more positively skewed. To take into account pretest differences, pretest scores were used as covariates in the analysis of all posttest comprehension and transfer test scores.

4.3.2 Pretests

Before any tests were conducted, the aim of the experiment was explained to both classes. In particular, it was emphasised that individual students were not being judged. The experimenter explained that, being involved in teacher education, she needed to know whether a particular kind of reading practice would help students improve their reading. This kind of information was important to pass on to trainee teachers, so that they could become good teachers. No student ever appeared anxious or was uncooperative.

Students completed one comprehension test on each of the first three days in the following order: *Why lions live in groups, Spiders as hunters* and *Carnivorous plants.* In each case they began with the knowledge and vocabulary tests. The procedure was explained via the examples printed on each test, and no one had any difficulty with this procedure. Students were encouraged to circle option (d) if they were unsure about the answer.

Students then read the text. They were instructed to read slowly and carefully, so that they would understand and remember the facts. They were warned of the comprehension questions to follow, and that texts would be removed before they answered questions. When all students indicated they had finished reading, the texts were removed. Students then wrote answers to the ten questions on the test sheet from memory.

On the fourth day, both classes completed the Survey of Attitudes to Reading and the transfer tests. It was explained that in an attitude survey there were no right or wrong answers; it was just to see how they personally felt. The experimenter worked through the two practice examples with the class, emphasising the two-stage decision process. In the first stage, students decided on which side

of the double line their opinion fell, True or Not True. In the second stage, they decided which of the remaining two boxes to tick, that is, the strength of their belief. No one had any difficulty with this procedure. The experimenter read each statement to the class, and when all students had made a choice, read the next statement. This procedure was adopted to ensure that responses were not confounded by decoding problems.

For the first of the transfer tests, the cloze test, the experimenter worked through the three practice examples printed on the test, as in the pilot study. The three "rules" of reading whole sentences, looking back and looking forward were again emphasised. Students then read and recorded their responses to the passage *The Mallee Fowl.*

For the second transfer test, the error detection passages, the procedure was explained via the two practice examples printed on the test. Students had no difficulty detecting the relatively simple contradictions in these examples. They were instructed to think carefully about each passage and to underline anything they were sure didn't make sense. It was emphasised that in the examples they knew for sure that the underlined sections could not possibly be correct because they contradicted an earlier statement, and that was the sort of problem they were to look for. The instructions explicitly drew attention to text-based errors, since children of this age may be inclined to overlook contradictions and comment instead on the plausibility of individual statements (Markman, 1979). Following the instructions, children read and marked the *Return of the Salmon* and associated passages.

All the comprehension and transfer tests were untimed. In fact, students were repeatedly encouraged to work slowly and to take as much time as they needed

to do their best work.

On the fifth day, students completed the ACER Primary Reading Survey, Level C, Form R. The experimenter read the standard instructions from the manual to the class. She worked through the three practice examples with students, as specified in the directions. No student had any difficulty with the procedure. A time of exactly 30 minutes was allowed for the test.

The teaching programme with the experimental class began the day after the last test. After this teaching session, the oral reading of students in both classes was tested. The material to be read was randomly chosen from the pretest comprehension passages, and turned out to be the text about carnivorous plants. Each student was seen individually in a quiet room adjoining the classroom. Their reading of the test passage was recorded for later analysis. The time taken to read the 364-word passage was measured from the recording, and the number of words read per minute was calculated. All errors, if uncorrected by the student, were noted. These errors were then categorised as either distorting or non-distorting errors, depending on whether they substantially altered the meaning of the sentence being read. Distorting errors included substitutions such as *usually* for *unusual*, *ferns* for *funnels*, *become* for *because* and *they can get* for *they can't get*. Non-distorting errors often involved adding or deleting an article (*a, the*), or changing number, as in *leaves curl* instead of *leaf curls*.

4.3.3 The teaching programme

On the first day of teaching, the experimenter worked through the introductory section of the work booklet with the class. This section aimed to demonstrate what the three main strategies of Collecting Ideas, Putting Ideas Together and

Checking and Clearing Up were, and why they were useful aids to comprehension and retention of factual information. As noted previously, this kind of information is essential if students are to maintain and transfer the use of the strategies.

As part of the demonstration, students worked through some examples, and then examined how the particular strategy they had used increased their understanding. For example, students were told that the first strategy, Collecting Ideas, involved reading a heading, pausing to recall information already known about that topic, and predicting what the section might be about. They then read the deliberately vague sentence, taken from Bransford and McCarrell (1974): *The haystack was important because the cloth ripped.* Everyone agreed that this sentence was not clear. The heading *The Parachute* was written on the board, and students were invited to say what they knew about about this topic. They then reread the vague sentence, and had no difficulty clarifying the referents and the relationship between the two clauses in the sentence. The same procedure was used with the second example, which stated: *He hid under the bushy tops when he heard the noise.* The heading *The Rabbit in the Carrot patch* was written on the board and discussed.

On the basis of these two examples, students were asked to consider why Collecting Ideas might be a useful strategy. All the students confidently asserted that it helped them to understand the sentences, although even capable students could not be prompted to give more specific reasons, such as helping them to work out exact meanings of words and reasons why things happened.

The same points also applied to the other two strategies, Putting Ideas Together and Checking and Clearing Up; students could apply the strategies successfully to the examples given, but could not give precise reasons why the strategies

had increased their understanding. On reflection, this imprecise articulation of the value of the strategies, although not anticipated, need not pose a problem, as long as the children were convinced that the strategies did indeed improve their understanding and were worth using.

Another problem with the introductory section became apparent; there was too much new information for one session with children of this age. Working through the examples took longer than anticipated. In the course of the experiment, this problem was dealt with by reinforcing points from the introductory section during subsequent teaching. For example, three advantages of generating a summary sentence were considered, namely, as an aid to retention of factual information, as an aid to understanding relationships in a paragraph, and as a test of understanding. So, whenever a rambling summary sentence was produced, students were reminded that a very long sentence would not make the paragraph easier to remember. If students had difficulty generating a summary, they were reminded that this could be a sign that they didn't fully understand the paragraph. Therefore, they should reread it, think about the topic of the paragraph and what was said about it, and try again to summarise.

Instruction continued for the next fourteen school days on a daily basis. The class spent approximately thirty minutes a day working through the teaching texts with the experimenter. Three times a week students independently completed an comprehension practice exercise after the instruction session. On alternate days, and unavoidably also on one of the individual practice days, folders containing the corrected practice exercises and a progress chart were returned to students. The experimenter also provided feedback regarding good responses to comprehension questions. In total, about seven hours were spent on class practice and about

four hours on individual practice during the three weeks of the experiment.

During the first five days, the class worked through the first two texts in the reading booklet, namely *Life on the Great Barrier Reef* and *Natural Disasters.* The experimenter read a paragraph aloud while students followed the text in their booklets. The reasons for this procedure were two-fold. First, it was anticipated that some readers in a mixed ability class would be poor decoders, and this assumption was supported by the results of the oral reading test. Thus, with the experimenter reading each paragraph aloud, even these weaker readers could concentrate on comprehension skills, without being hindered by decoding problems. The second benefit from this procedure became evident during pretesting. Some students worked almost twice as quickly as others. Therefore, the reading by the experimenter also helped to keep the class working together.

During this first week, then, the experimenter read each paragraph aloud, asked the students to choose the best main idea sentence from among the options given, and to mark this option in their work booklets. Several responses were then sought from the children. For the first few paragraphs, the experimenter stated the best option, and explained why the others were less than ideal. Students were reminded of the characteristics of a good main idea sentence, for example, that it should be short and concise enough to remember and should include all parts of the paragraph, not just one aspect. After only a few explanations of this kind, students were able to take over this evaluative function. Every attempt was made to involve a variety of students in stating and justifying their responses.

The same procedure was used for the second question relating to how the paragraph fitted into the overall topic. The class as a whole was then asked whether any words or ideas needed clearing up. If someone detected a problem,

attention was directed to the part of the paragraph where the problematic information occurred. Students were encouraged to use the context to have "a good guess" to find a solution to the problem. Often this strategy was successful, but where there was insufficient information in the text to provide a useful context, the experimenter supplied an answer and reminded the children that if they were reading on their own they could look up such difficult words in a dictionary. If the children did not identify any problematic information in a paragraph, the experimenter sometimes drew attention to difficult words or an incompletely explained idea, and asked whether anyone could clarify the problems. If very few students could do so, it provided an opportunity to reiterate the importance of checking one's comprehension. When an incompletely explained idea was identified it also provided an opportunity to discuss remedial action, for example, by using prior knowledge and the passage context to generate an inference, by reading on to look for further details or, if the information is particularly important, by consulting other books.

Whenever a heading occurred, the class as a whole was asked to state what they knew about the topic and what they thought the section might be about.

Initially students chose from among the multiple-choice alternatives in an apparently haphazard way. It was common to have each of the three options chosen by different students in the class. Gradually, however, they became more skillful and commented on the "obvious" inadequacies of some of the options. By the end of the first week most students, but not all, were not only consistently choosing the correct alternative, but could also state what was wrong with the distractors.

All the texts studied during this training period took longer than one session

to read. Whenever a text was carried over from the previous day, students were first asked to state the main points of the text already studied. This procedure was used throughout the three weeks.

During the second week, students spent four days working through the longer passage entitled *Animal Defence Systems.* For this passage, students were still given the relevant questions to ask in their work booklets, but they now wrote their own responses. The procedure was still the same as in the first week, except that the first response received for the first two questions in the Putting Ideas Together category were written on the board. The class was asked to comment on this response, and amendments were made as necessary. Only occasionally did the experimenter have to supply this evaluation. Several students then read out their responses, usually variations on the one written up. Again evaluation of and amendments to these responses came mainly from students.

After two weeks it became apparent that there would not be enough time to work through the whole reading booklet. Therefore, in the final week the class studied the last passage, *The Age of Dinosaurs*, which could definitely be finished in the time remaining. Once concluded, the class studied a shorter piece, *How Snakes Use Their Senses.* This text had proved difficult during pilot testing, but was manageable given the opportunity for clarification of ideas through class discussion.

In these last five days of training, students were no longer supplied with questions. They had to decide for themselves what to do after reading a heading or a paragraph, and then write down a response. Students had no difficulty generating the required questions. Since they were now very familiar with the strategies, they could answer all four questions before engaging in any class discussion. In

other respects, the procedure was identical to that in the second week.

The time taken to write down responses enabled the experimenter to help some of the weaker students. In weeks two and three, she circulated amongst the students while they were writing. There was no doubt that some of the weaker students still had difficulty generating main idea sentences. In these cases, she reminded students that they might need to read the paragraph again. She then asked the individuals to say what the topic of the paragraph was and what was said about it, and then to combine the two in a sentence. Given this amount of guidance, even weak students could complete the task, but to become skillful at generating their own summaries, these students needed more than three weeks of practice. In hindsight, it would have been wise to begin the instruction with this type of direct instruction, as was done by Kurtz and Borkowski (1987) and Stevens (1988), rather than relying heavily on modelling and inductive generation of this procedure. The modelling of the decision-making process worked successfully for the capable students, but not so successfully for the weaker readers.

Each student was given a personal folder, which contained all their individual work and a progress chart. The pretest comprehension scores were plotted on the chart, as were scores for all the individual practice pieces. Weekly averages were plotted in red. All students eagerly followed their progress on these charts, and most could see their red line going up. A class chart was also prepared and displayed in the room to show the class average for each practice piece, with the weekly averages again marked in red. On the class chart, the red line also steadily went up. The folders were not handed out during pretesting or posttesting, but were examined only after the test period had finished.

The class was extremely enthusiastic about the reading practice, and actively

participated in the work. There were a few quiet students, however, who needed encouragement to volunteer their responses, and it was often difficult to ensure equitable participation in class discussion, especially given the lively nature of most of the class. The experiment was carried out towards the end of winter, at a time when many people were ill with influenza. A few students had extended absences due to illness. Overall, though, the training sessions ran very smoothly. The teacher of the class was present during these sessions, and she was impressed with the programme. She commented not only on the improvement in students' work, but also their obvious enthusiasm and the level of sustained concentration among these ten year-olds. She intended to continue this kind of reading practice with the class, and arranged a seminar for other staff members to find out about the programme.

4.3.4 Immediate posttests

Posttesting was conducted on the five days immediately following the conclusion of training. The procedure was exactly the same as in pretesting. Students completed one comprehension test on each of the first three days in the following order: *The Ant Colony, The Moon* and *Fungi - the Strange Plants.* No mention was made of the trained strategies; students in both classes were simply told to read carefully. On the fourth day, they completed the transfer tests and the Survey of Attitudes to Reading. The cloze test was *Shooting Stars* and the error detection test was *Sea Mammals* and associated passages. On the final day, they completed the ACER Primary Reading Survey, using the alternate form, Form S.

Lysynchuk et al (1989) include among their criteria for internal validity of studies that the time spent on dependent measures by control and experimental

children be recorded and compared, and that some process measures of actual strategy use be included. Neither of these recommendations was followed, for several reasons.

Children sometimes acted as if working quickly was a virtue in itself. It was feared that any overt recording of time taken to complete work would exacerbate this attitude, which was evident in both classes studied. During pretesting children in both classes worked quickly and looked around to see who else had finished. They could not be prevailed upon to check their work during the time when other students were still working. This attitude persisted during posttesting in the control class, but not the experimental class. Discrete, informal noting of time spent completing tests suggested that the experimental class worked longer than the control class on both posttests.

Lysynchuk et al suggest that time spent on dependent measures should be analysed to avoid the possibility that the strategy instruction has an effect on performance 'only' because students spend longer on the task. Considering the impulsive approach of many students in this study, it was no small achievement to have them spend more time reading and answering questions. First, they were motivated to spend more time, and presumably more effort, and second, they knew how to productively spend this time. Similar conclusions were also reached by Adams et al (1982) who noted that their trained students spent more time strategically studying texts rather than reading them through nonstop as they had done previously. In summary, time spent on task was not measured to avoid exacerbating an undesirable and prevalent attitude, and the additional time taken by the experimental group to complete tests was not considered a problem for the reasons given.

No process measure of strategy use during testing was taken, for several reasons. First of all, the posttests were intended as a measure of 'spontaneous' strategy use; children were not reminded to use the strategies. Process measures such as think-aloud procedures or retrospective reporting could potentially interfere with this intention by drawing attention to the strategies during the testing period. As indicated in the first chapter, all such measures have problems, such as interfering with normal processing or requiring high levels of introspection and verbal ability. Finally, several studies have also shown a poor relationship between reported strategy use and the actual strategy use evident in performance (Phifer and Glover, 1982; Kletzien, 1991; Wong and Wong, 1986; Olshavsky, 1976-7; Hare and Borchardt, 1984). For these various reasons, therefore, no process measures were included.

4.3.5 Delayed posttests

Delayed posttesting was carried out ten weeks after the conclusion of instruction. The procedure was the same as in the pretesting. The comprehension tests used were *The Earth and Other Planets*, *A for Aardvark* and *Why Bees Dance.* Again, no mention was made of the trained strategies. The cloze test was *Cuckoos - the Lazy Parents* and the error detection test was the *Starfish* group. Since the ACER Primary Reading Survey does not have three alternate forms, Form S was again used. Thus, performance on Form R in pretesting was compared to performance on Form S in both the immediate and delayed posttesting.

Chapter 5

Results and Initial Comment

5.1 Introduction

In chapter three a number of questions were formulated in a manner which arose directly from the research discussed in that chapter. These questions were as follows:

1. How successfully can a training programme similar to that used by Palincsar be used with children younger than the grade 7 students primarily taught by Palincsar?

2. Can the programme be implemented in a written format so as to still retain its interactive nature and a gradual transfer of responsibility to students, while at the same time providing a more structured guide for teachers to follow?

3. Will any training effects be evident without prompting, on a range of comprehension tests other than those used during training, and will training effects be maintained over time without further intervention?

4. Can the programme be as beneficial to average and above average comprehenders as to the poor comprehenders primarily studied by Palincsar?

5. Can any specific characteristics be identified to distinguish those students who benefit from the programme and those who do not?

Clearly there is some overlap between the first three of these questions. To eliminate this overlap and to provide a framework within which to systematically present the results, the three questions were reworded as follows:

1. Can a training programme similar to that used by Palincsar be implemented in a written format with grade 5 children such that posttest comprehension achievement is improved?

2. Will any training effect be maintained over time without prompting or further intervention?

3. Will any training effect transfer to a range of comprehension measures other than those used during training, and will any such transfer effects be maintained over time?

Results will be analysed in relation to these five questions. Before proceeding with these analyses, however, some preliminary comments need to be made.

The first comment relates to the data to be analysed. It was noted in the previous chapter that some students were absent during part of the testing periods. Analysis of the data actually recorded would therefore result in some loss of information. For example, a student who had completed two of the three comprehension tests at a particular time of testing would not be considered in the analysis of total comprehension scores, involving all three tests. To reduce this

loss of information as much as possible, missing scores for students who had completed at least one of the three comprehension tests at a particular time of testing were replaced with the group mean for the test(s). In this way, all available data was used in the analysis of comprehension tests. This addition was carried out only when at least one score was recorded for the student, and so was not relevant to the ACER Primary Reading Survey, cloze or error detection tests, where only one test was completed at each time of testing.

Results for comprehension tests were analysed using both data files, that is, both the normal file containing only scores actually recorded and the file expanded by the addition of group means for missing values. Since there was on no occasion a difference between the two analyses in terms of whether or not a difference reached the conventional significance level of $p < .05$, the results reported are based on the expanded data file in order to make use of all the available information.

The second preliminary comment relates to the reliability of scoring. The scoring of some tests was unambiguous due to a multiple-choice format; these tests included the ACER Primary Reading Survey, the Survey of Attitudes to Reading, the vocabulary and knowledge tests.

To assess the reliability of scoring comprehension, cloze and error detection tests, all immediate posttests from both the experimental and the control classes were scored independently by both the experimenter and her supervisor. The interrater reliabilities were as follows: total comprehension scores (three tests) $r = .9968$, cloze test $r = .9899$, error detection hits $r = .9968$, and error detection false alarms $r = 1.000$. These high correlations indicate that the tests were scored reliably; any small disagreements were easily resolved through discussion.

The final preliminary comment relates to subgroups within the overall sample of children. There were no significant differences between boys and girls on any of the comprehension and transfer measures at any of the three times of testing, so gender will not be considered further. There were, however, differences due to ethnicity; Anglo students consistently scored higher than students of non-Anglo background. This difference will be considered in more detail once overall group differences have been discussed.

All analyses were conducted using the SPSS-X package (Norusis, 1988). In this chapter, pretest scores for the two classes will initially be compared, followed by the analyses of the overall effectiveness of the intervention, addressing the earlier questions of how well the written format worked for children in grade 5. Their performance on comprehension tests will be examined first, followed by an examination of performance on the transfer tests. Next the effect of the programme on children of above average, average and below average reading ability will be considered, and then an attempt will be made to isolate factors which could explain why some children benefited from the programme while others did not. The final section will examine the relationship between ethnicity and comprehension achievement, as well as the effect of the training programme on attitudes to reading.

5.2 Pretest class comparisons

In the previous chapter it was pointed out that although children were not randomly assigned to the experimental and control groups by the experimenter, the children had in fact been randomly assigned to classes by the school. The intention in this first section is to establish statistically that the two classes can be regarded as random samples drawn from the population of grade 5 students attending parish schools similar to the one studied. The null hypothesis to be tested is that:

There is no difference in experimental and control population means on any pretest measure of comprehension.

Group differences were analysed by means of t-tests. Table 5.1 shows the comparison of pretest comprehension test scores, while Table 5.2 shows the comparison of pretest transfer test scores.

As the two tables indicate, there was no significant difference between the two classes on any pretest measure of comprehension, so the null hypothesis can be accepted and the two classes can be regarded as random samples drawn from the population of grade 5 students.

In the sections to follow, class means for posttests and delayed posttests will be reported and class differences analysed by means of t-tests. However, since the experimental group showed a small but nonsignificant advantage over the control group on most measures, scores were also subjected to analyses of covariance, using the corresponding pretest measure as the covariate, to take into account this initial small difference.

Table 5.1: Class comparison of pretest comprehension test scores

Test	Class	Results					
		Max. score	Mean	S.D.	t	df	$p =$
Why Lions Live in Groups	Expt	10.00	3.62	2.13	0.15	48	.879
	Cont		3.53	1.92			
Spiders as Hunters	Expt	10.00	5.88	1.89	0.38	48	.705
	Cont		5.65	2.46			
Carnivorous Plants	Expt	10.00	4.96	2.29	0.96	48	.342
	Cont		4.40	1.80			
Total comprehension score	Expt	30.00	14.46	5.42	0.58	48	.567
	Cont		13.57	5.42			

Table 5.2: Class comparison of pretest transfer test scores

Test	Class	Results					
		Max. score	Mean	S.D.	t	df	$p =$
Error detection (hits)	Expt	6.00	2.20	1.24	0.60	46	.548
	Cont		2.52	2.14			
Error detection (f.a.)	Expt		2.12	2.32	1.27	46	.212
	Cont		3.26	3.81			
Cloze test	Expt	20.00	11.84	4.42	1.11	46	.272
	Cont		10.22	5.66			
ACER test	Expt	39.00	19.88	8.57	0.95	46	.347
	Cont		17.50	8.76			

5.3 Overall effectiveness of training: comprehension tests

In this section the first two of the questions set out in the introduction will be addressed. Given the training of a class of grade 5 students according to the method described in the previous chapter, it is hypothesised that:

1a) The population mean immediate posttest comprehension score, for each individual test and in total, of trained students is higher than that of untrained students.

2a) The population mean delayed posttest comprehension score, for each individual test and in total, of trained students is higher than that of untrained students.

In statistical terms, hypotheses 1a) and 2a) are alternatives to those actually under test. The hypotheses under test are:

1b) The population immediate posttest comprehension score, for each individual test and in total, of trained students is less than or equal to that of untrained students.

2b) The population delayed posttest comprehension score, for each individual test and in total, of trained students is less than or equal to that of untrained students.

Table 5.3 presents the results of the immediate posttest comprehension tests, as well as the results of one-tailed t-tests, since a directional hypothesis was being tested. The maximum score for each individual test was 10, making the maximum total score 30.

In the previous section it was shown that the experimental and control classes

did not differ significantly in terms of pretest total comprehension scores. However, as Table 5.3 shows, following training the experimental class was significantly superior to the control class on each individual comprehension test and on posttest total comprehension ($p < .001$), allowing rejection of hypothesis 1b) and acceptance of its alternative, hypothesis 1a).

Table 5.3: Class comparison of immediate posttest comprehension scores

Test	Class	Results				
		Mean	S.D.	t	df	$p =$
The Ant Colony	Expt	6.38	1.26	3.91	48	$< .001$
	Cont	4.71	1.73			
The Moon	Expt	6.38	2.09	2.60	48	.006
	Cont	4.85	2.07			
Fungi - the Strange Plants	Expt	5.18	2.54	3.63	48	$< .001$
	Cont	2.71	2.25			
Total comprehension score	Expt	17.94	5.37	3.76	48	$< .001$
	Cont	12.27	5.28			

Table 5.4 presents results of the delayed posttest comprehension tests and results of the one-tailed t-tests. Again, the experimental class was significantly superior to the control class on each individual delayed posttest comprehension test and on delayed posttest total comprehension ($p = .006$), allowing rejection

of hypothesis 2b) and acceptance of its alternative, hypothesis 2a)

Table 5.4: Class comparison of delayed posttest comprehension scores

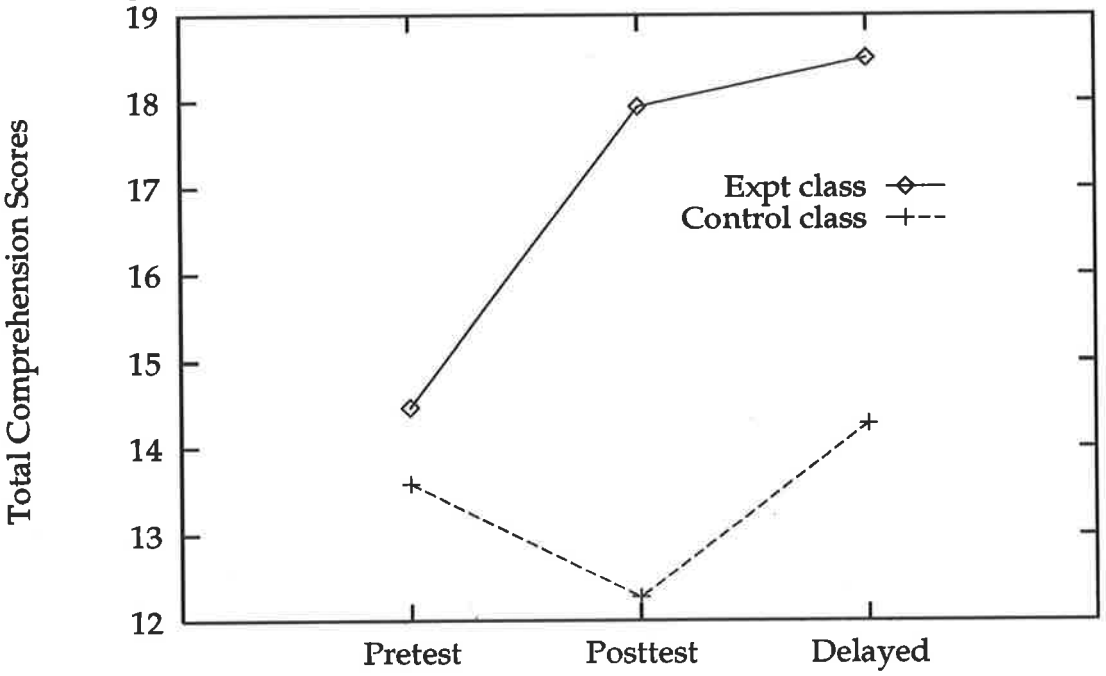
Test	Class	Results				
		Mean	S.D.	t	df	$p =$
The Earth and Other Planets	Expt	6.81	2.02	1.71	46	.047
	Cont	5.75	2.28			
A for Aardvark	Expt	6.06	2.38	1.89	46	.033
	Cont	4.92	1.78			
Why Bees Dance	Expt	5.63	2.44	3.21	46	.001
	Cont	3.60	1.88			
Total comprehension score	Expt	18.50	6.00	2.64	46	.006
	Cont	14.27	5.06			

Hays (1963) recommends that in addition to demonstrating a statistically significant difference between groups as a result of training, an estimate of the strength of association between treatment and achievement should be made. He states: *"A significant result implies that it is safe to say some association exists, but the estimate of* ω^2 *tells how strong that association appears to be. It seems far more sensible to decide to follow up a finding that is both significant and indicates a strong degree of association than to tie this course of action to significance level alone."* (p. 328)

The estimated ω^2 for posttest total comprehension scores was .21, while the estimated ω^2 for delayed posttest total comprehension scores was .11. Thus, training was estimated to account for approximately 21% of variance in immediate posttest scores and 11% of variance in delayed posttest scores, suggesting a degree of association between training and comprehension achievement which has some educational as well as statistical significance.

Total comprehension scores for the two classes at the three times of testing are graphically represented on Figure 5.1.

Figure 5.1: Class comprehension scores at three times of testing



Since the experimental class performed slightly, albeit nonsignificantly better than the control class on comprehension pretests, immediate and delayed posttest total comprehension scores were also analysed using pretest scores as the covariate. Table 5.5 and Table 5.6 show the results of these analyses of covariance.

The tables show that the analyses of covariance confirmed the conclusions

Table 5.5: Analysis of covariance: immediate posttest total comprehension scores

Source of variance	Sum of Squares	DF	Mean Square	F	Sig of F =
Covariate: pretest comp	1038.751	1	1038.751	116.090	$< .001$
Main effects: training	303.614	1	303.614	33.932	$< .001$
Explained	1342.364	2	671.182	75.011	$< .001$
Residual	420.545	47	8.948		
Total	1762.910	49	35.978		

reached on the basis of the t-tests; the experimental class performed significantly better on both the immediate and delayed comprehension posttests, even when pretest differences were accounted for.

Table 5.6: Analysis of covariance: delayed posttest total comprehension scores

Source of variance	Sum of Squares	DF	Mean Square	F	Sig of $F =$
Covariate: pretest comp	949.433	1	949.433	81.701	$< .001$
Main effects: training	158.747	1	158.747	13.660	.001
Explained	1108.180	2	554.090	47.681	$< .001$
Residual	522.940	45	11.621		
Total	1631.120	47	34.705		

In part, the first two questions set out at the beginning of this chapter have been answered by the preceding analyses; posttest comprehension improved as a result of training, and this training effect was maintained for 10 weeks without further intervention. Before addressing the next question, the effectiveness of this programme will be briefly compared to the Palincsar studies, where children were older (grade 7), the groups were smaller, and the programme was conducted orally rather than in writing.

The improvements made by the experimental class represent a substantial gain in terms of being able to learn from text. Table 5.7 shows the mean percentage of questions answered correctly by the two classes for each group of three comprehension tests.

Table 5.7: Percentage of comprehension questions answered correctly

Class	Pretest	Posttest	Delayed
Experimental	48	60	62
Control	45	41	48

Relative to the pretest, the experimental class improved by an average of 12% on the immediate posttest and 14% on the delayed posttest to reach final class averages of 60% and 62% respectively. In comparing these results to results achieved by Palincsar, it must be kept in mind that the relative difficulty of test passages in various studies cannot be assessed. Although in all cases materials were written at a grade-appropriate level, different content can make the passages more or less difficult to understand, depending on how familiar the various topics are to children.

The four grade 7 students taught individually by Brown and Palincsar (1982) achieved scores of 85-90%. The six students taught in groups of two for twenty days by Palincsar and Brown (1984) averaged 80% after training, while the groups of four to seven students taught by classroom teachers for fifteen days averaged 70-80%. In Palincsar (1987), students were taught in groups of seven to sixteen by classroom teachers for twenty days, and 70% of the children reached the criterion of 70% correct on at least four of the five posttests. In Palincsar and Brown (1986), intact classes, of which no details are given, used a written version of the programme in content area (science) lessons for twenty days. Palincsar and Brown note that from a pretest average of 35%, students generally scored 20% higher on their weekly tests during training. In this case, though, the tests tapped information read in class as part of the training, and thus were not independent posttests.

The gains achieved in the present study were not as high as those generally achieved by Palincsar. However, her training period was usually twenty days, rather than the fifteen days used in this study, and her groups were smaller than the twenty-five children taught in this study. As the results above indicate, the magnitude of gains in her studies declined as the number of children taught increased, as would be expected when children can be given less individual guidance. It has been pointed out that the intention in this study was to test a training programme which could potentially be used by teachers with normal, intact classes to prepare them for learning from expository texts. It has to be accepted that comprehension gains in this context are likely to be smaller than when students are taught individually or in small groups. Considering this point, it is fair to conclude that grade 5 children can be taught effectively according to a

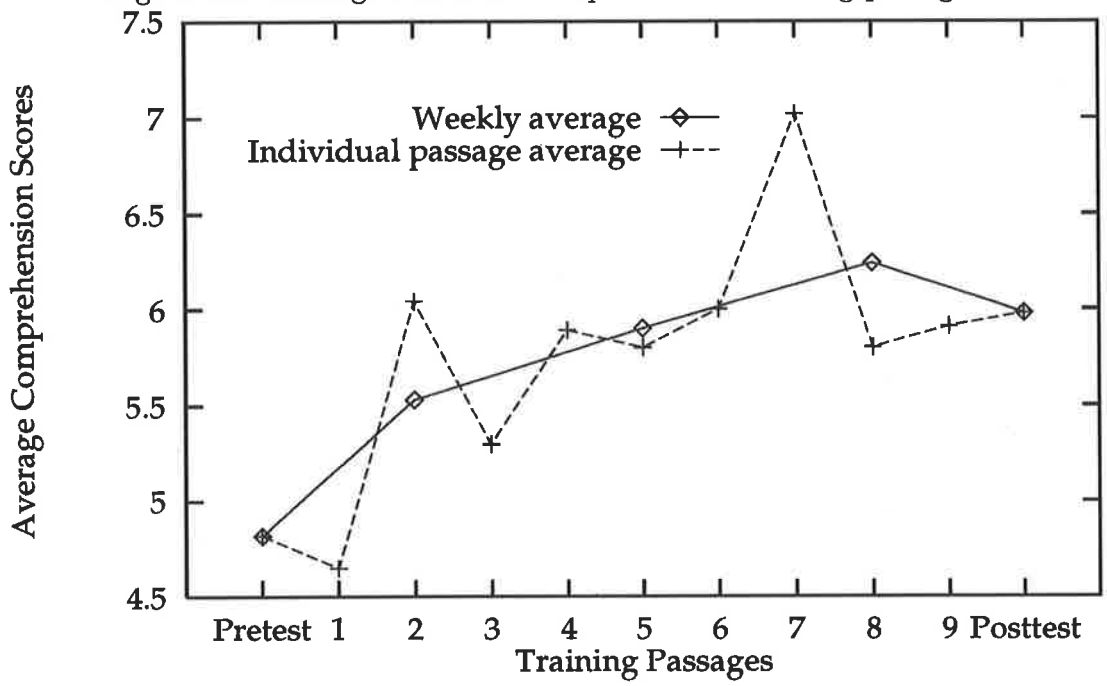
modified and written version of the Palincsar approach, although training effects may be increased by the extension of the training period and some changes to the teaching method, as will be described in more detail in the next chapter.

Before going on to the next section, a brief mention will be made of the children's progress on the practice comprehension passages during training. Three times a week after the class training, the children individually read a passage similar in format to the test passages, and answered ten questions from memory. Although the passages were written at grade 5 level and were of a similar length and based on similar topics to the test passages, the practice pieces had not all been pilot tested. These pieces, therefore, were likely to be more variable in terms of content familiarity than the test passages, which were chosen on the basis of pilot testing scores that were neither very high nor very low.

As Figure 5.2 shows, the individual practice pieces were of varying difficulty, but there was a steady improvement in comprehension during the training period. This improvement is clearly evident in the weekly averages in relation to pretest and posttest averages.

Thus, the children's performance on practice passages during training supports the conclusion of a substantial improvement in comprehension as a result of training. This performance suggests that the children were applying the strategies which they had been taught. However, a detailed analysis of the degree to which the strategies had been mastered was not undertaken. The work of Berkowitz (1986) and Taylor (1982) shows that degree of strategy mastery is an important determinant of comprehension improvement. While comprehension clearly improved in terms of class means, it would be useful to have available information on individual students in order to assess their degree of strategy mastery and the

Figure 5.2: Average scores for comprehension training passages



relationship of this factor to comprehension improvement.

To some extent, information on this point would have been available from students' work booklets, which were left with the class teacher after the delayed posttests. She was very enthusiastic about the programme and asked to keep the materials for future use. When the work booklets were required to check strategy mastery, the teacher had left the school and so this information could not be collected. However, on reflection, the work booklets may not have provided reliable data on strategy mastery, since students could have corrected their responses following the class discussion of each paragraph. It will be suggested in the next chapter that more reliable data could be gathered by including a paragraph summarisation task in the battery of tests.

5.4 Overall effectiveness of the training programme: transfer tests

This section will address the third question at the beginning of this chapter, the question concerning the transfer of training effects to other tests of comprehension. Three tests were used for this purpose, these being an error detection test, a cloze test, and a standardised reading test, the ACER Primary Reading Survey. It was hypothesised that:

3a) The population mean scores for each of three immediate posttest transfer tests are higher for trained grade 5 students than for untrained students.

4a) The population mean scores for each of three delayed posttest transfer tests are higher for trained grade 5 students than for untrained students.

In statistical terms, hypotheses 3a) and 4a) are alternatives to those actually under test. The hypotheses under test are:

3b) The population mean scores for each of three immediate posttest transfer tests for trained grade 5 students are less than or equal to those of untrained students.

4b) The population mean scores for each of three delayed posttest transfer tests for trained grade 5 students are less than or equal to those of untrained students.

Class differences on each of the tests will first be assessed by means of t-tests, and then by analysis of covariance with the corresponding pretest score as the covariate. In each case the t-tests are two-tailed for the pretest comparison and one-tailed for posttest comparisons, since in the latter directional hypotheses are being tested.

5.4.1 Error detection tests

The number of hits (detection of inconsistencies) and false alarms (underlining of unproblematic information) were analysed separately. Table 5.8 presents the average number of inconsistencies detected (hits) by the two classes at the three times of testing, and the significance of class differences according to t-tests. The maximum score for each test was 6.

Table 5.8: Class comparison of error detection (hits) scores

		Results				
Test	Class	Mean	S.D.	t	df	$p =$
Pretest	Expt	2.20	1.24	0.60	46	.548
	Cont	2.52	2.14			
Posttest	Expt	3.17	1.79	1.23	44	.112
	Cont	2.55	1.61			
Delayed	Expt	2.92	1.43	1.50	46	.181
	Cont	2.30	1.42			

The results indicate that the experimental class was not significantly different to the control class at any time of testing, thereby not allowing the rejection of the hypotheses under test (hypotheses 3b) and 4b)) as these relate to error detection (hits). However, the score for the experimental group was lower than that for the control group on the pretest. To take this difference into account, analyses of covariance were conducted, using pretest scores as the covariate. Results for the

immediate and delayed posttests are presented in Table 5.9 and Table 5.10.

Table 5.9: Covariance analysis of immediate posttest error detection: hits

Source of variance	Sum of Squares	DF	Mean Square	F	Sig of F =
Covariate: pretest ED Hits	24.241	1	24.241	10.380	.002
Main effect: training	6.812	1	6.812	2.917	.095
Explained	31.054	2	15.527	6.648	.003
Residual	98.091	42	2.335		
Total	129.144	44	2.935		

These covariance analyses reveal that following training the experimental group moved in the predicted direction relative to the control group on both the immediate and delayed posttests, although the differences did not quite reach a conventional level of significance, being $p < .10$ and $p < .07$ respectively. Fig-

Table 5.10: Covariance analysis of delayed posttest error detection: hits

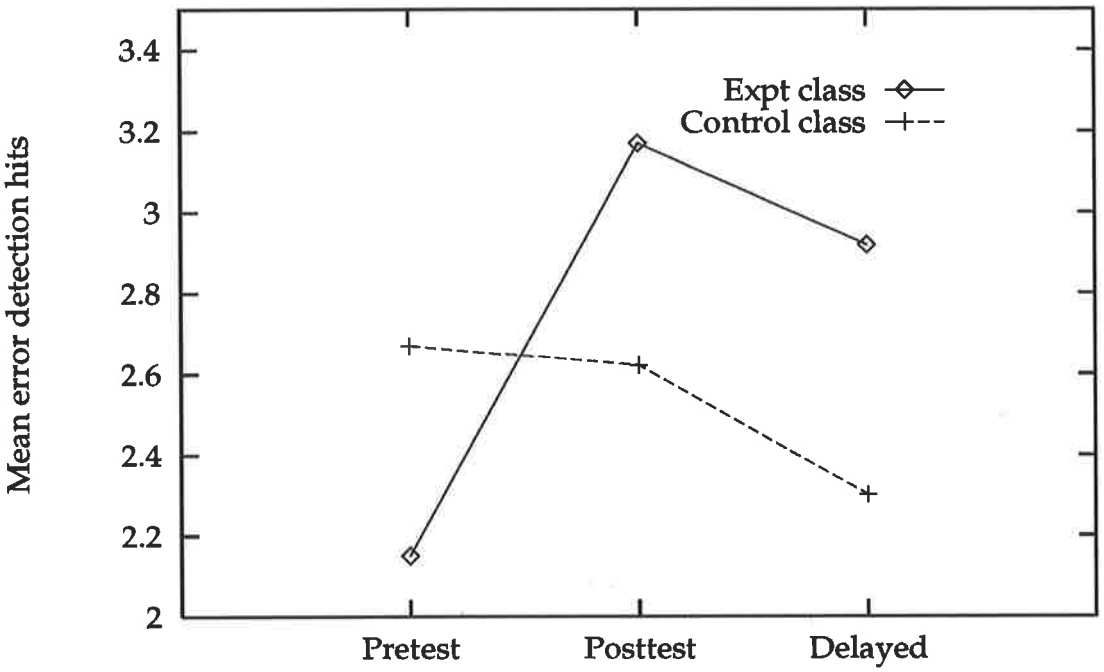
Source of variance	Sum of Squares	DF	Mean Square	F	Sig of $F =$
Covariate: pretest ED Hits	14.191	1	14.191	8.247	.006
Main effect: training	6.126	1	6.126	3.560	.066
Explained	20.317	2	10.158	5.904	.005
Residual	77.433	45	1.721		
Total	97.750	47	2.080		

ure 5.3 graphically represents the average number of hits by the two classes at the three times of testing.

Turning now to false alarm scores. Table 5.11 shows average scores for false alarms in the two classes at three times of testing, and the significance of the differences according to the appropriate t-tests.

From the table, the two groups were not significantly different on the pretest, but the experimental group produced significantly fewer false alarms on both the immediate and the delayed posttest with $p = .003$ and $p < .03$ respectively, allowing rejection of the hypotheses under test and the acceptance of their alternatives (hypotheses 3a) and 4a)) as these relate to error detection (false alarms).

Figure 5.3: Class hit scores at three times of testing



However, since the score for the experimental group was also lower, albeit non-significantly, on the pretest, these results were checked by means of covariate analyses, using the pretest score as the covariate. The results of these analyses are shown in Table 5.12 and Table 5.13.

These covariance analyses reveal that the experimental group produced significantly fewer false alarms on the immediate posttest ($p < .007$). Results on the delayed posttest were in the same direction, but did not quite reach a conventional level of significance, being $p < .08$. Figure 5.4 graphically represents the average number of false alarms produced by the two classes at three times of testing.

To conclude this section, the results do not allow the rejection of hypotheses 3b) and 4b), except in the case of false alarms, where the alternative hypothesis 3a) relating to immediate posttest results can be accepted. In the case of im-

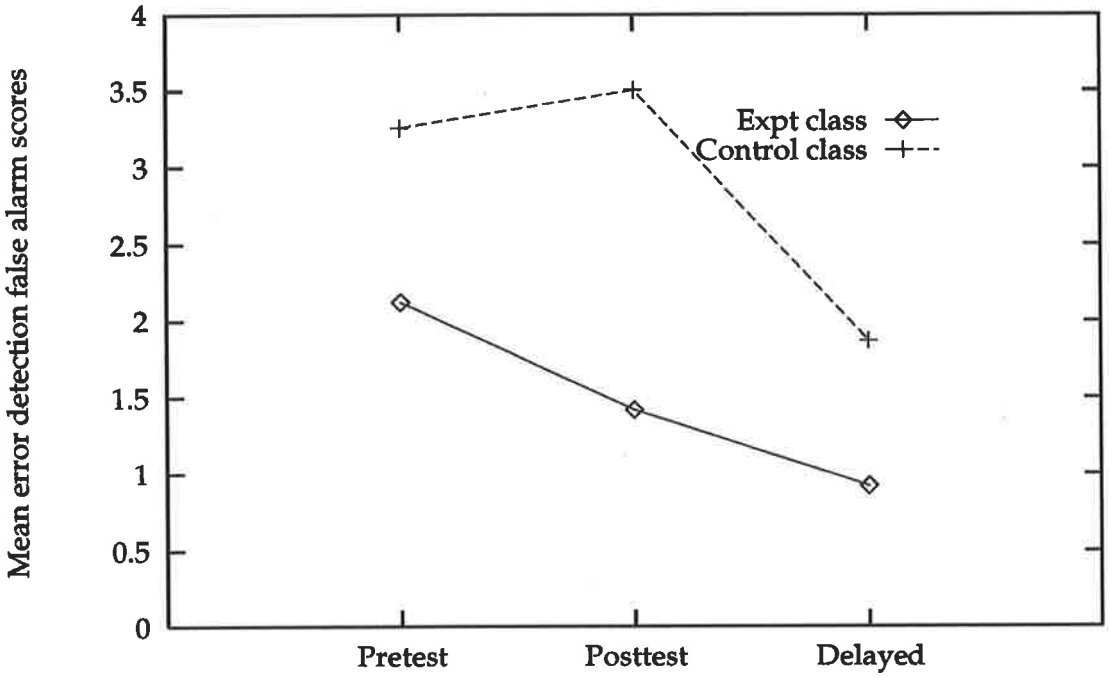
Table 5.11: Class comparison of error detection (false alarm) scores

Test	Class	Results				
		Mean	S.D.	t	df	$p =$
Pretest	Expt	2.12	2.32	1.27	46	.212
	Cont	3.26	3.81			
Posttest	Expt	1.42	1.59	2.91	44	.003
	Cont	3.50	3.10			
Delayed	Expt	0.92	1.15	2.07	46	.023
	Cont	1.87	1.96			

mediate and delayed posttest hits and delayed posttest false alarms, the results did not allow rejection of hypotheses 3b) and 4b) at the conventional levels of significance, but indicated gains in the predicted direction.

In chapter three it was pointed out that an error detection test involving internal inconsistencies was chosen as a particularly stringent test of constructive processing and comprehension monitoring. Although the differences between the two classes reached the conventional level of significance of $p < .05$ on only one of the four comparisons, differences on both the immediate and delayed posttests were in the predicted direction and approached this level of significance when pretest scores were taken into account. Thus, it would be fair to conclude that the experimental group tended to detect more internal inconsistencies and tended to underline less unproblematic information. Gains that were made by the ex-

Figure 5.4: Class false alarms at three times of testing



perimental class were maintained on the delayed test.

Table 5.12: Covariance analysis of immediate posttest error detection: false alarms

Source of variance	Sum of Squares	DF	Mean Square	F	Sig of F =
Covariate: pretest ED false alarms	40.540	1	40.540	7.671	.008
Main effect: training	44.610	1	44.610	8.441	.006
Explained	85.150	2	42.575	8.056	.001
Residual	221.961	42	5.285		
Total	307.111	44	6.980		

Table 5.13: Covariance analysis of delayed posttest error detection: false alarms

Source of variance	Sum of Squares	DF	Mean Square	F	Sig of $F =$
Covariate: pretest ED false alarms	5.622	1	5.622	2.235	.142
Main effect: training	8.413	1	8.413	3.344	.074
Explained	14.035	2	7.018	2.789	.072
Residual	113.215	45	2.516		
Total	127.250	47	2.707		

5.4.2 Cloze tests

As described in the previous chapter, cloze tests were constructed by leaving the first sentence of each passage intact and deleting every fifth word thereafter. A cloze test was included as a measure of intersentence integration. Table 5.14 shows the average scores for the two classes at three times of testing and the significance of class differences according to the appropriate t-tests. The maximum score for each test was 20.

Table 5.14: Class comparison of cloze test scores

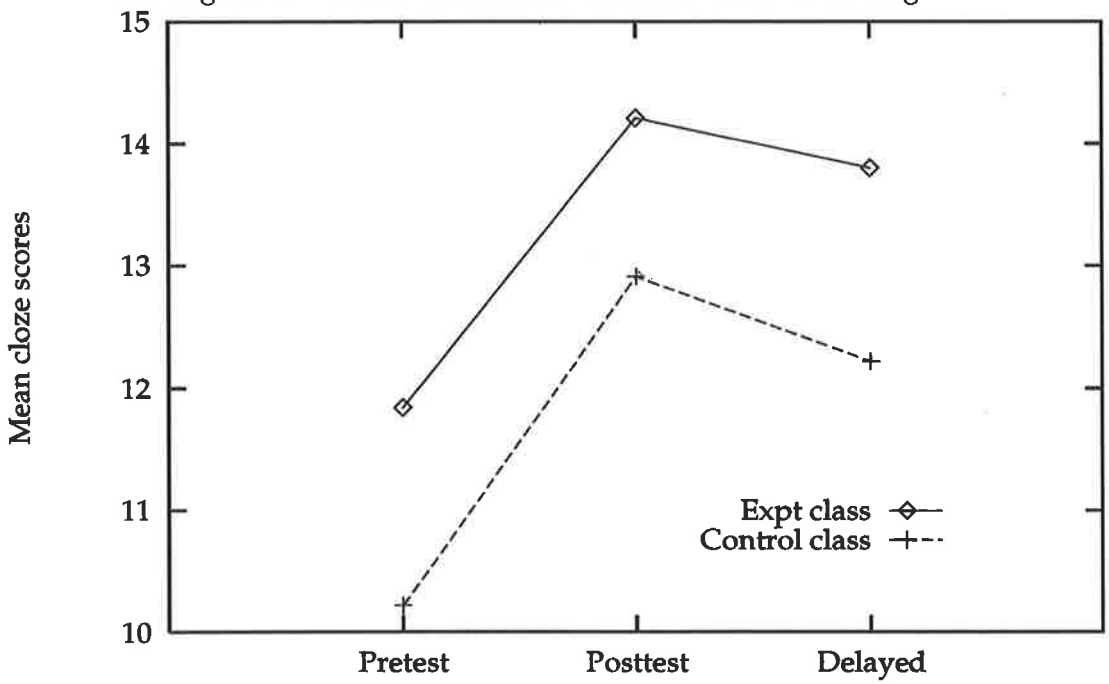
		Results				
Test	Class	Mean	S.D.	t	df	$p =$
Pretest	Expt	11.84	4.42	1.11	46	.272
	Cont	10.22	5.66			
Posttest	Expt	14.21	3.89	1.16	44	.126
	Cont	12.91	3.66			
Delayed	Expt	13.80	3.34	1.58	46	.061
	Cont	12.22	3.68			

As the table indicates, the two classes were not significantly different at any time of testing. This result was confirmed by covariance analyses using pretest scores as the covariate for both the immediate posttest $F(1,43) < 1$, $p = .38$ and the delayed posttest $F(1,47) = 1.2$, $p = .28$. Once again, the hypotheses that the population means for cloze tests of trained students are less than or equal to

the population means for untrained students at either time of testing cannot be rejected, and so the alternative hypotheses 3a) and 4a) stated at the beginning of this overall section cannot be accepted with regard to cloze tests.

Figure 5.5 shows the average scores for the two classes at the three times of testing.

Figure 5.5: Class cloze scores at three times of testing



The cloze test was intended as a measure of intersentence integration, and it was hypothesised that the training programme would improve this ability. The hypothesised improvement did not occur. However, in chapter three problems with cloze tests with every nth word deleted were discussed, particularly in relation to comments made by Shanahan, Kamil and Tobin (1982). The filling of blanks created by such systematic deletion does not necessarily require intersentence integration, and so the failure to find no significant class differences on these tests does not necessarily indicate that no improvement in such integration oc-

curred. Given the fact that the experimental group showed some improvement in the detection of internal inconsistencies on the error detection test, which required the integration of nonadjacent sentences, it is possible that this group would also have shown improvement relative to the control class on a cloze test with content words selectively deleted to require intersentence integration. Before concluding that a training programme such as that used in this study has no effect on the integration of successive sentences, it is suggested that tests designed specifically to measure this ability be used. In hindsight, the use of a cloze test with every nth word deleted was not a wise choice, given that the intention was to measure intersentence integration.

5.4.3 ACER Primary Reading Survey

The ACER Primary Reading Survey is an Australian standardised reading test, where a time limit of thirty minutes is set and answers to multiple-choice questions are recorded on a computer scoring sheet. Table 5.15 shows the average scores for the two classes at three times of testing, as well as the significance of class differences according to the appropriate t-tests. The maximum score for each test was 39.

As the table indicates, the difference between the two classes on the immediate posttest was significant ($p < .04$), but there was no significant difference on the delayed posttest. Since there was a small pretest difference between the classes, immediate and delayed posttests were also subjected to covariance analyses, using pretest scores as the covariate. Table 5.16 shows the analysis of immediate posttest scores.

As the table indicates, the difference between the two classes was in the pre-

Table 5.15: Class comparison of ACER test scores

Test	Class	Results				
		Mean	S.D.	t	df	$p =$
Pretest	Expt	19.88	8.57	0.95	46	.347
	Cont	17.50	8.76			
Posttest	Expt	25.52	7.60	1.81	44	.038
	Cont	21.22	8.47			
Delayed	Expt	26.88	6.77	0.92	46	.181
	Cont	24.96	7.60			

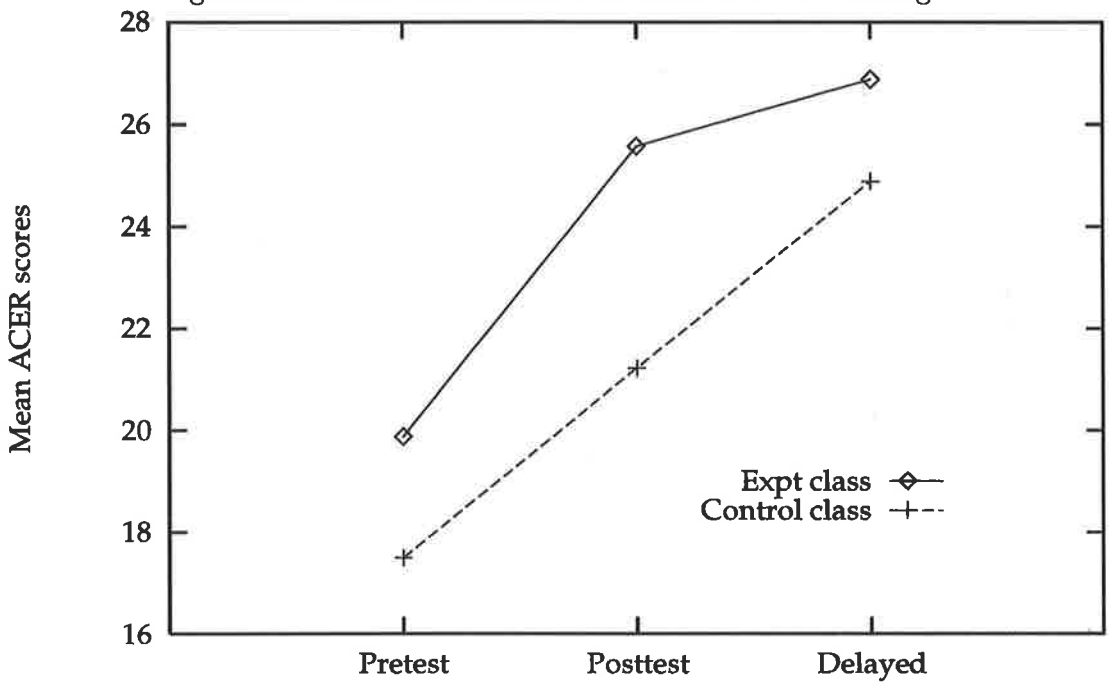
dicted direction, but did not reach the conventional level of significance, with $p = .13$. For the delayed posttest, the analysis of covariance confirmed the result already obtained from the t-test. There was no significant difference between the classes, with $F(1, 46) < 1$, $p = .62$. Thus, the results do not allow the rejection of the hypotheses that the population means for the ACER test of trained students are less than or equal to the population means of untrained students at either time of testing, hence the alternative hypotheses 3a) and 4a) put forward at the beginning of this overall section cannot be accepted with regard to the ACER test.

Table 5.16: Covariance analysis of immediate posttest ACER scores

Source of variance	Sum of Squares	DF	Mean Square	F	Sig of $F =$
Covariate: Pretest ACER scores	1879.176	1	1879.176	94.627	$< .001$
Main effect: training	47.522	1	47.522	2.393	.130
Explained	1926.698	2	963.349	48.510	$< .001$
Residual	814.211	41	19.859		
Total	2740.909	43	63.742		

Figure 5.6 graphically represents average scores for the two classes at the three times of testing.

Figure 5.6: Class ACER scores at three times of testing



It is apparent from the graph that both classes improved over the three times of testing. This improvement is most likely to be due to a practice effect with the unfamiliar format of a standardised test. These children had not previously completed a speeded test where answers were marked on a computer scoring sheet.

The ACER test is a stringent test of comprehension since it requires efficient processing not only in terms of accuracy of constructed meanings, but also in terms of speed. Results from the comprehension tests reported in the first section of this chapter suggest that the experimental group did improve in its ability to construct meaning from texts relative to the control group, but used additional reading time to do so. This improvement would, therefore, be difficult to demon-

strate on a speeded test. However, the class difference approaching significance on the ACER immediate posttest suggests that the experimental class was beginning to develop the ability to apply its skills under time pressure immediately following training, but the development of this efficiency did not continue over the ten weeks without instruction to the delayed posttest.

It is possible that with further practice following the initial intensive training, for example, by having two reading sessions per week to increase efficiency in applying the trained skills, the children may be in a better position to demonstrate improvements in comprehension on a speeded test.

To conclude this overall section on transfer tests, the results did not allow rejection of the hypotheses that the population means of trained students were less than or equal to the population means of untrained students for the three tests at the two times of testing, with one exception. The alternative hypotheses 3a) and 4a) stated at the beginning of this section could not be accepted, except in the case of immediate posttest false alarms. However, in several instances results were in the predicted direction and approached a conventional level of significance. In particular, children in the experimental class showed signs of improvement in error detection on both the immediate and the delayed posttests and the ACER immediate posttest.

5.5 Treatment-ability interactions

As explained previously, children in the experimental class had widely varying levels of reading ability. In order to examine whether the training programme benefited all ability groups to a similar extent, three groups were formed post hoc on the basis of pretest ACER scores. The scores themselves suggested the division into groups of approximately equal size. Table 5.17 shows the number of children and the percentile range in each group.

Table 5.17: Ability groups and their percentile range

Group	N	Percentile range
Poor	8	8 - 24
Average	8	30 - 42
Good	9	61 - 95

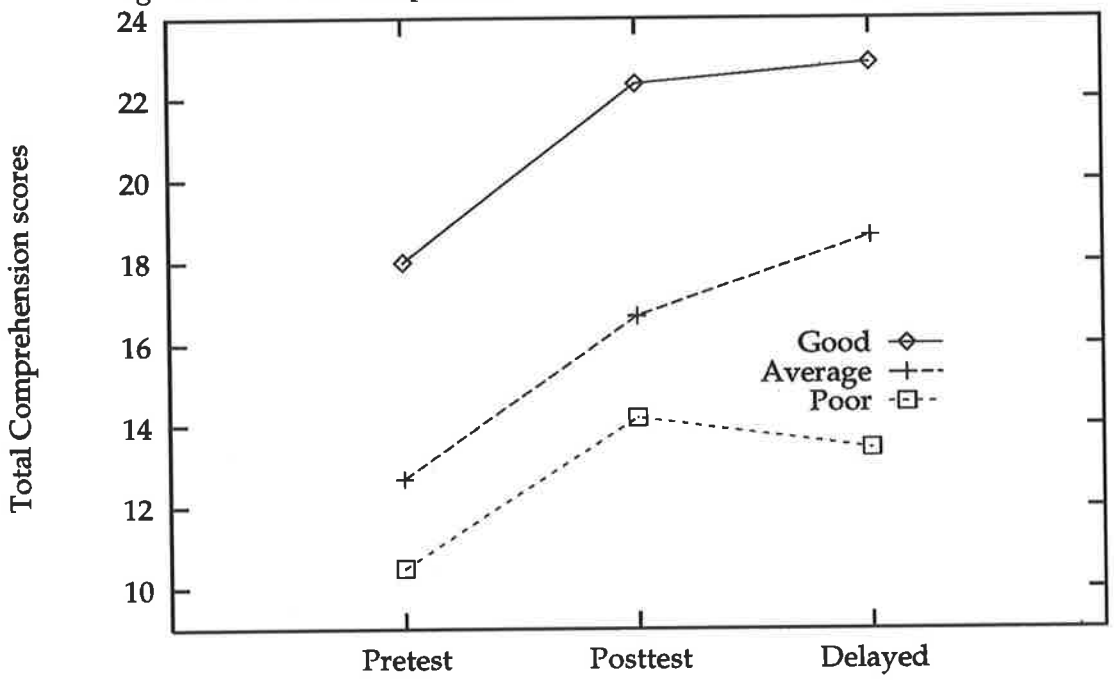
It should be noted that the middle group was average only in relation to this sample; in relation to national Australian norms the group tended to be below average, although obviously less so than the poor group.

In order to assess whether each of these groups improved on comprehension tests to a similar degree from the pretest to the posttest and the pretest to the delayed posttest, post-training scores were subjected to analysis of covariance, using pretest comprehension scores as the covariate. It was hypothesised that the training would be equally beneficial for all ability groups. More specifically, it was hypothesised that:

5) There will be no difference in the mean comprehension gain scores of populations of good, average and poor readers in response to training.

On neither the comprehension posttest nor the delayed posttest was there a significant difference between the three ability groups once pretest scores were accounted for, with $F(2,24) < 1$, $p = .48$ for the posttest and $F(2,23) = 2.50$, $p = .11$ for the delayed posttest. These results indicate that the null hypothesis 5) stated above can be accepted. As a point of interest, though, group differences on the delayed posttest did approach significance, with the adjusted means indicating that the average readers tended to improve most and the poor readers least between the pretest and the delayed posttest, as can be seen on Figure 5.7, which graphically represents the total comprehension scores for the three groups at the three times of testing.

Figure 5.7: Total comprehension scores for three ability groups



When the ACER immediate and delayed posttests were subjected to analysis of covariance with the pretest ACER scores as the covariate, the F-values in both analyses were below 1, indicating that there was no significant difference between

the ability groups once pretest scores were accounted for. Hence the training affected all ability groups in a similar way in terms of their performance on the standardised ACER test.

A very similar pattern emerged from the analyses of immediate and delayed posttest error detection hits and false alarms and the cloze tests. In only two of these six analyses did the F-value exceed 2, these being posttest false alarms $F(2, 23) = 2.05$, $p = .16$ and the posttest cloze $F(2, 23) = 2.08$, $p = .15$.

Thus, from these analyses of comprehension tests and three transfer tests at two times of testing there was no significant evidence of an interaction between training and pretest reading ability, supporting hypothesis 5) that there would be no difference in the magnitude of comprehension gains achieved by good, average and poor readers as a result of training. In response to the fourth question set out at the beginning of this chapter, the training programme can be beneficial to above average and average readers, as well as to below average readers.

Before going on to the next section, a brief mention will be made of the decoding adequacy of the ability groups. Not surprisingly, the group of poor readers included the weakest decoders in terms of both reading speed and distorting errors. Table 5.18 shows the reading speed (w.p.m.), the number of distorting errors (D.E.), nondistorting errors (N.D.E.) and total errors, with total errors per minute in parentheses, for the eight poor readers when reading a 364-word passage (Carnivorous plants).

The students classified as adequate decoders in the Palincsar studies read at least 80 w.p.m. with less than two errors per minute. Since no statement to the contrary is made, this criterion presumably included errors of all kinds, making no distinction between distorting and nondistorting errors. According

Table 5.18: Reading speed and accuracy for the group of poor readers

Student	ACER %tile	w.p.m.	D.E.	N.D.E.	Total
17	8	92	3	10	13 (3.29)
08	8	117	2	9	11 (3.54)
24	14	109	2	2	4 (1.20)
07	14	116	5	10	15 (4.78)
21	18	108	0	7	7 (2.08)
25	21	121	0	1	1 (0.33)
19	24	119	3	0	3 (0.98)
03	24	124	5	6	11 (3.75)
Class mean		143	1.0	4.6	5.3 (2.11)

to this criterion, less than half of the group of poor readers in this study would be classified as adequate decoders when total errors were considered; for the remainder, reading speed was adequate but they were too inaccurate to meet the criterion. However, the total error score included nondistorting errors which, while indicating less than perfect visual analysis, were not considered detrimental to comprehension. As explained in the previous chapter, nondistorting errors included substitutions such as *a sweet nectar* for *sweet nectar* or *leaves* for *leaf.*

Of greater concern were the distorting errors; twenty-six such errors were made by the entire experimental class, and of these, twenty were contributed by the poor readers. As Table 5.18 shows, students 03 and 07 made the highest number of such errors. The mistakes made by student 03 suggest that when uncertain she guessed words on the basis of their beginning and general shape. Her distorting

substitutions included *usual* for *unusual*, *funless* for *funnel*, *vitamin* for *victim*, *nature* for *nectar*, and *attack* for *attract*. While there was clearly a relationship between the visual appearance of the words and her pronunciation, these errors would make it difficult if not impossible for student 03 to comprehend the passage as intended by its author. Student 07 appeared not to know the word *prey* and read the nonsense word *perie* whenever it occurred. Among her other distorting errors she also rather surprisingly substituted *small plants* for *small animals*.

To conclude, most of the students in the experimental class were adequate decoders. All had adequate reading speed according to the criterion used by Palincsar, although two individuals in particular were inaccurate readers, to the point where the inaccuracy of their visual analysis of print potentially interfered with comprehension.

5.6 Variables which predict comprehension gains

A further question to be addressed in this study is why some students fail to benefit from instruction while others make substantial gains. In Palincsar and Brown (1984) one of the six students in the first experiment failed to improve, although her entry level in terms of comprehension and Gates-MacGinitie scores and decoding ability was similar to that of students who did improve. In Palincsar, Brown and Martin (1987), 20% of the students tutored by their peers failed to make substantial gains.

In the previous section it was shown that treatment did not interact with reading ability, as measured by the ACER pretest. Therefore, the ACER test, on its own, cannot be used to discriminate between students who benefit from training and those who do not. It seems unlikely that such a discrimination can be made on the basis of any other single variable. Discriminant analysis is a multivariate statistical technique which allows analysis of variables together rather than one at a time (Norusis, 1988). Therefore, discriminant analysis was used to evaluate the hypothesis that:

6) There is no difference in population mean comprehension improvement scores of groups formed from any combination of potentially discriminating variables.

A discriminant analysis, when used to discriminate between two groups, constructs a single discriminant function, which is a linear combination of variables weighted to maximise within group similarity and between group difference. The weighting of each variable is provided by the discriminant coefficient for the vari-

able in question and allows calculation of a discriminant score for each student. Students are then classified into groups according to whether their discriminant score is above or below a certain cut-off point.

In the discriminant analysis the Wilk's method was used, which selects variables in step-wise fashion according to their importance in discriminating between groups (Norusis, 1988). The criterion is the ability of each variable to minimise the overall Wilk's lambda statistic. All eleven pretest measures were included in the analysis, these being reading speed and oral errors (distorting and non-distorting), vocabulary, subject knowledge, attitudes to reading, comprehension scores, ACER scores, error detection scores (hits and false alarms) and cloze test scores.

Two groups of children were formed post hoc on the basis of their degree of improvement. The criterion chosen for classification as an 'improver' was at least a 10% increase from the pretest total comprehension score on *both* the immediate and delayed posttests. Thus, a student who scored 18 of the possible 30 marks on the pretest would need to score at least 21 on both posttests to meet the criterion. The criterion was made more stringent by the fact that any posttest gain also had to be maintained on the delayed test. Group 1 were the improvers and group 2 the nonimprovers.

Only 16 of the 25 children in the experimental class were included in the discriminant analysis. Of the remainder, six had incomplete comprehension test scores for at least one time of testing, two (students 01 and 18) had complete comprehension scores but were missing data for other variables, and one student (student 15) was deliberately excluded. This student scored 85% on the pretest and 93% on the delayed posttest, and so was performing close to ceiling level. It

was thought that his inclusion in the group which did not meet the criterion of improvement may have been confusing, since it is likely that the reason for his small improvement was different to that of other students.

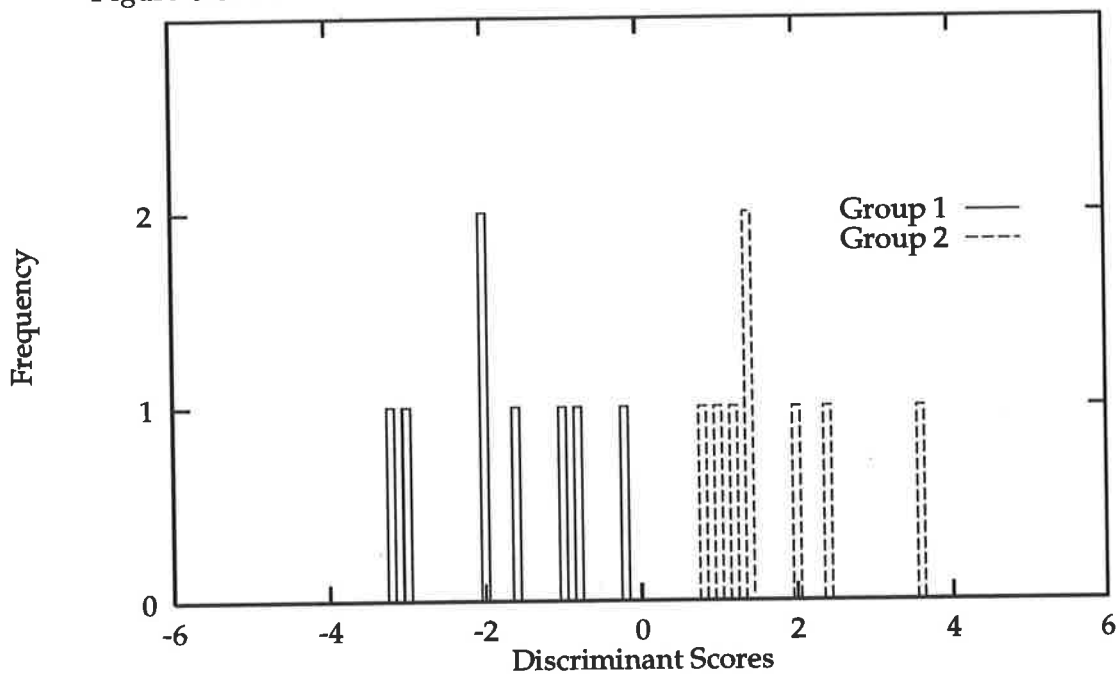
A significant discriminant function was found, with Wilk's lambda = 0.22, equivalent $F(6,9) = 5.17$, $p < .02$, allowing rejection of the hypothesis under test (hypothesis 6). A discriminant model based on six variables discriminated between students who met the criterion of improvement (group 1) and those who did not (group 2). Table 5.19 shows the order in which these six variables were entered into the equation according to the criterion of minimising Wilk's lambda.

Table 5.19: Variables in the discriminant model

Variable	Wilk's lambda	sig.
1. Reading speed	.81	.09
2. Pretest ACER	.58	.03
3. Pretest knowledge	.45	.02
4. Pretest cloze	.38	.02
5. Pretest Attitude	.30	.02
6. Non-distorting oral errors	.22	.01

The discriminant model based on these six variables successfully classified 100% of the 16 students with complete data; 8 students were correctly classified into group 1 and 8 were correctly classified into group 2. By chance 50% of students could be correctly classified into the two groups, whereas the discriminant model correctly classified 100% of students. Figure 5.8 shows the distribution of discriminant scores for these 16 students.

Figure 5.8: Distribution of discriminant scores for groups 1 and 2



As indicated previously, the discriminant score for each student is calculated by multiplying his or her score for each variable in the model by the discriminant score for that variable and then summing the products. Raw scores are multiplied by the relevant unstandardised canonical discriminant function coefficient, however Table 5.20 also lists the standardised coefficients, since they indicate the relative contribution of each variable to the discriminant scores once the scale of measurement has been equalised.

Table 5.21 and Table 5.22 show the discriminant scores of students classified into groups 1 and 2 respectively, and the probability of each student being classified in this way given their discriminant score.

As Table 5.21 and Table 5.22 indicate, a negative discriminant score placed students in group 1, and a positive score placed them in group 2. Two variables making an important contribution to these discriminant scores were reading speed

Table 5.20: Canonical discriminant function coefficients

Variable	Canonical discriminant function coefficients	
	Standardised	Unstandardised
Reading speed	-2.91	-0.10
ACER test	2.17	0.24
Knowledge	0.92	0.35
Cloze test	1.69	0.33
Attitude	-0.87	-0.25
Non-distorting errors	0.98	0.31
(Constant)		4.53

and ACER scores, as can be seen by examining previous tables. Table 5.19 shows that the discriminant function became significant after the entry of these two variables, although the remaining four variables continued to reduce the Wilk's lambda and to improve the ability of the discriminant model to correctly classify students. Table 5.20 indicates that these two variables had the highest standardised canonical discriminant function coefficients; reading speed made a relatively large negative contribution to a student's discriminant score, whereas the ACER test made a relatively large positive contribution. In other words, students who had a high reading speed *relative to* their ACER score were likely to receive a negative discriminant score and hence to be classified into group 1. On the other hand, students who had a high ACER score *relative to* their reading speed were likely to be classified into group 2. It should be noted that it is the relative rather than the absolute value of these scores which influenced

Table 5.21: Discriminant scores and probabilities of inclusion in group 1

Student ID	Discrim. score	Probability
04	-3.28	1.000
08	-0.23	0.691
11	-1.64	0.997
13	-1.92	0.999
14	-2.94	1.000
16	-1.03	0.973
22	-2.09	0.999
24	-0.76	0.934
Mean	-1.74	

a student's classification; as Table 5.23 shows, both groups included students with varying levels of reading speed and ACER achievement. With regard to the remaining four variables, the discriminant coefficients listed in Table 5.20 indicate that students were likely to have negative discriminant scores, and hence to be classified into group 1, if they had good decoding ability (high reading speed, low non-distorting errors) and high attitude scores *relative to* their ACER, cloze and knowledge scores.

Table 5.23 and Table 5.24 show scores for the 16 students in the analysis for the six variables in the discriminant model, as well as their percentage improvement in comprehension from the pretest to the posttest and the pretest to the delayed posttest, and their classification as good, average or poor readers according to the ACER pretest (see the previous section on treatment by ability interactions

Table 5.22: Discriminant scores and probabilities of inclusion in group 2

Student ID	Discrim. score	Probability
02	0.72	0.925
06	3.66	1.000
07	1.94	0.999
17	1.05	0.975
19	1.20	0.985
20	1.43	0.993
21	2.47	1.000
23	1.41	0.993
Mean	1.74	

for details of this classification).

Table 5.23: Values for discriminating variables in group 1

Stud	ACER	Pretest	Posttest	Delayed	Read	Pretest	Pretest	Pretest	Pretest	Read
I.D.	class.	Comp.	% change	% change	Speed	ACER	Know.	Cloze	Attitude	N.D.E.
04	good	18.5	11.7	16.7	190	25	12	19	18	2
08	poor	8.5	20.0	36.7	117	9	9	6	9	2
11	av.	7.5	10.0	23.3	142	16	12	13	13	0
13	av.	9.0	28.3	25.0	157	16	14	14	18	5
14	good	17.0	25.0	30.0	182	34	10	8	15	5
16	good	19.5	10.0	11.7	168	32	10	11	13	3
22	av.	13.0	21.7	15.0	137	18	9	9	18	7
24	poor	10.5	35.0	16.7	109	11	11	10	16	2
Mean		12.9	20.2	21.9	150.3	20.1	10.9	11.3	15.0	3.3
18[a]	av.	17.0	20.0	25.0	149	17	(6)	17	14	7

[a]student not included in the original analysis

Table 5.24: Values for discriminating variables in group 2

Stud	ACER	Pretest	Posttest	Delayed	Read	Pretest	Pretest	Pretest	Pretest	Read
I.D.	class.	Comp.	% change	% change	Speed	ACER	Know.	Cloze	Attitude	N.D.E.
02	av.	19.5	-8.3	15.0	140	20	14	12	16	5
06	good	20.5	8.3	23.3	168	35	17	19	19	4
07	poor	8.5	-3.3	-11.7	116	11	14	6	11	10
17	poor	9.5	0.0	3.3	92	9	14	0	15	10
19	poor	10.0	18.3	1.7	119	14	12	14	11	0
20	good	14.0	21.7	3.3	146	25	8	13	7	4
21	poor	17.0	10.0	3.3	108	12	15	9	15	7
23	av.	9.5	5.0	21.7	116	17	8	15	15	4
Mean		13.6	6.5	7.5	125.6	17.9	12.8	11.0	13.6	5.5
01[a]	good	15.5	6.7	1.7	168	25	(12)	13	18	3
15[a]	good	25.5	6.7	8.3	172	32	15	16	19	3

[a]student not included in the original analysis

Students marked with subscripts (^{a}) were not included in the analysis for reasons stated previously, although they had complete comprehension scores and so could be classified as improvers or non-improvers. Student 15 had been deliberately excluded because of his high comprehension scores, and students 01 and 18 were missing one of three knowledge test scores. In order to be able to test the discriminant model on these three students, class mean values were substituted for the one knowledge test missed by students 01 and 18. The following discriminant scores were calculated:

Table 5.25: Discriminant scores for three additional students

Stud. ID	Discrim. score	Classification
01	-3.03	group 1
15	0.00	unclassified
18	-1.40	group 1

Student 18 was correctly classified according to the discriminant model as an improver in group 1. However, student 01 was incorrectly classified as an improver, and the score of student 15 fell between the two groups, although he should have been classified into group 2 on the basis of his small comprehension improvement. Thus, only one of the three students was correctly classified. It should also be noted that although all of the 16 students in the original analysis were correctly classified, some of the discriminant scores did not reflect the students' actual comprehension gains. Examination of Table 5.21, Table 5.22 and Table 5.23 shows that student 08 was the most marginal member of group 1 according to her discriminant score, although she showed very large comprehension gains. On the other hand, the high discriminant score of student 06 clearly placed

her in group 2, whereas her comprehension gains only just fell short of the criterion for improvement. These points suggest limitations in the generalisability of the model.

Norusis (1988) notes that perfect classification in a discriminant analysis is unusual, and that a model will normally fit the sample from which it is derived to a greater degree than another sample from the same population. The predictive power of the present model could not be tested, since no data were available from another sample. However, the classification of the three students not included in the original analysis is an example of the point made by Norusis. It should also be noted that limitations in the model would be expected, since it was derived from a small number of cases.

Nevertheless, bearing in mind these provisos, the discriminant model indicated that students who met the criterion of improvement and clearly benefited from training were those whose comprehension skills (ACER, cloze) were not commensurate with their decoding ability (reading speed, non-distorting errors), and who had relatively positive attitudes to reading but less subject knowledge. This pattern applied to all reading ability levels within this sample. Those students who did not meet the criterion of improvement had high comprehension skills relative to their decoding ability. The educational implications of this finding will be discussed further in the next chapter.

It is interesting to note that the students trained so successfully in, for example, Brown and Palincsar (1982) and Palincsar and Brown (1984) were adequate decoders but poor comprehenders. In other words, they were students likely to be predicted as comprehension improvers by the discriminant model derived from data in the present study.

5.7 Additional comments

This final section will examine two further aspects of the study, namely, differences between students of Anglo and non-Anglo background, and attitudes to reading within the experimental class following training.

5.7.1 Ethnic differences

As was pointed out in the previous chapter, fifteen students in the experimental class and eighteen students in the control class were of non-Anglo background, primarily Italian. However, all but one of these children was Australian-born. Pretest comparisons indicated a significant difference between students of Anglo and non-Anglo background on many of the variables that were measured. For the purposes of this analysis, scores from the two classes were combined, since no training had yet taken place. Table 5.26 shows average scores for the two groups of children on all pretest measures, as well as standard deviations, t-values and their significance, and estimated ω^2 values. The t-tests here are two-tailed, since no difference was originally predicted in population means of Anglo and non-Anglo students.

Table 5.26: Pretest differences between Anglo and non-Anglo groups

Test	Group	N	Mean	S.D.	t	df	p	ω^2
Comprehension total	Anglo	17	17.12	4.83	3.91	48	.003	.23
	non-Anglo	33	12.42	4.99				
ACER	Anglo	16	25.81	8.14	4.93	46	< .001	.33
	non-Anglo	32	15.13	6.50				
Cloze test	Anglo	17	13.35	4.82	2.44	46	.019	.10
	non-Anglo	31	9.81	4.81				
Error detection: hits	Anglo	17	3.26	1.85	2.90	46	.006	.14
	non-Anglo	31	1.87	1.44				
Error detection: false alarms	Anglo	17	1.53	1.74	1.91	46	.062	.05
	non-Anglo	31	3.29	3.56				
Reading speed	Anglo	17	152.65	28.94	3.27	47	.002	.17
	non-Anglo	32	126.78	24.86				
Distorting errors	Anglo	17	0.76	1.05	1.64	47	.108	.03
	non-Anglo	32	2.31	3.96				
Non-distorting errors	Anglo	17	5.59	4.68	0.41	47	.687	.00
	non-Anglo	32	6.13	4.27				
Attitudes to reading	Anglo	17	16.76	1.92	3.07	46	.004	.15
	non-Anglo	31	12.90	4.97				
Total vocabulary	Anglo	17	25.29	2.62	2.87	48	.006	.13
	non-Anglo	33	21.61	4.94				
Total knowledge	Anglo	17	12.35	3.10	1.68	48	.099	.04
	non-Anglo	33	10.58	3.74				

As the table shows, significant differences between the groups were found on the comprehension test, the ACER test, the cloze test, error detection, attitudes to reading, oral reading speed, and vocabulary test. Differences approached significance for error detection false alarms, oral reading distorting errors and subject knowledge. As a group, then, the children of non-Anglo background were less skilled readers and had less knowledge relevant to reading than their peers of Anglo background. Where there was a significant difference between the groups, the estimated ω^2 value indicates that ethnicity accounted for variance in scores ranging from 10% for the cloze test to 33% for the ACER test. Thus, for the comprehension and ACER tests in particular, there was a moderately strong association between ethnicity and pretest achievement.

However, of greater interest than their entry level was whether children of Anglo and non-Anglo background benefited to a similar extent from the training programme. Considering that there was no interaction between treatment and reading ability, as demonstrated previously, it was anticipated that the two groups would make similar comprehension gains. Hence the hypothesis to be tested is as follows:

7) There is no difference in the mean comprehension gain scores of populations of children of Anglo and non-Anglo descent in response to training.

Scores for the immediate and delayed comprehension and transfer tests for the experimental class were subjected to analysis of covariance, with the corresponding pretest score as the covariate. In each analysis the F-value was below 1 and hence far short of significance, indicating that the performance of the two groups was not significantly different once pretest scores were accounted for. Therefore, hypothesis 7) predicting no difference can be accepted.

This finding is particularly important from the point of view of being able to generalise the findings of this study to a population of grade 5 students wider than that in the catchment area of this school. The percentage of students of Italian background was high within the samples studied, however, since both Anglo and non-Anglo students benefited equally from the training the findings can be generalised to grade 5 classes with varying ethnic compositions. It was pointed out, though, that all but one of the non-Anglo children were Australian-born, so the findings may not apply to more recently arrived migrant children.

5.7.2 Attitudes to reading

The Survey of Attitudes to Reading was administered at each time of testing in order to examine whether negative attitudes could limit the effectiveness of training, and also whether attitudes would change as a result of the training. Table 5.27 shows the average attitude scores for the two classes at the three times of testing. Standard deviations are shown in parentheses. The possible range of scores was 20 to -20.

Table 5.27: Average attitude scores for the two classes

Class	Pretest	Posttest	Delayed
Experimental	14.68 (3.35)	15.25 (4.44)	15.35 (2.91)
Control	13.83 (5.58)	14.50 (5.27)	14.38 (6.26)

The two classes were not significantly different at any time of testing, nor was there a significant difference within classes between times of testing. Thus, the training programme did not alter the attitudes of the experimental class towards

reading. Paris and Oka (1986b) suggest that such attitudes are not easily altered by a training programme of limited duration. It certainly was the case that individual students were fairly consistent in their ratings from one time of testing to the next. This lack of change was not a problem considering the positive attitudes expressed by most students, as can be seen from the class averages in Table 5.27.

There were, however, some students in the experimental class who expressed less positive attitudes. Student 17 and student 07 were among the weakest readers in the class and did not improve as a result of training. Their attitude scores, respectively, dropped from 15 to 6 and from 11 to 7 over the course of training. In the course of the daily reading practice they may have become more aware of their below average achievement and their lack of improvement, even though every effort was made to encourage all students.

Student 08 retained less positive attitudes than her peers, even though she made substantial gains during training. Her pretest comprehension average was a low 28%, but she improved by 20% on the immediate posttest and 37% on the delayed posttest to raise her above the class average. Nevertheless, her attitude scores were 9, 11 and 8 on the pretest, posttest and delayed posttest respectively. On the other hand, the attitude of student 20 improved substantially. This boy frequently announced that he was a poor reader; he had four older sisters who told him so. However, he was above the 60th percentile on the ACER pretest, and made substantial gains on both the ACER test and the comprehension tests following training. His attitude score moved from 7 to 16, and he conceded that he was no longer a poor reader.

After the completion of the posttests, a few students spontaneously wrote

thank you notes and their comments about the training programme. Prompted by this action, the class teacher asked all students, once the experimenter had left the school, whether they would care to write comments about the programme and what they thought they had gained from it. The comments were on the whole very favourable, although students also expressed some reservations and suggestions.

Most students remarked that they had learned more about reading from the programme and felt that they were now reading with more understanding. Some comments of this nature included:

> *"I learned to read slowly through the book, not rushing through it like a competition."*
>
> *"The lessons helped me to find out about stories instead of reading them and not learning anything."*
>
> *"The lessons helped me to understand what I am reading and to shorten things down in my own words."*
>
> *"I learned how to break down paragraphs and understand what I read more. The lessons helped me to read more, and to read more carefully."*

Several students commented specifically that the lessons had helped them in gathering information for projects and in writing projects.

> *"I think the lessons helped me in my project work to pick out important information because I found the project we did after the lessons easier than projects I'd done before."*
>
> *"Before when I did projects I used to copy writing out of the book, but now I write projects in my own words."*

"The lessons helped me a lot to improve my reading and the way I write information for projects."

The class teacher also formed the impression that students' project work had improved. Aspects she particularly noted were less verbatim copying out of books, greater organisation of information, and more use of headings and subheadings.

Although most students expressed favourable opinions on the effectiveness of the programme and most, if not all, students found the topics studied interesting, there were also some reservations and suggestions. As one student commented: *"I think the lessons were interesting but went from easy to hard too quickly."* A number of other students also commented that the lessons became quite difficult towards the end, when students were working more independently. Some students also commented that they were getting bored towards the end. This comment was supported by a greater restlessness observed in the class in the last few days of training. One boy made the sensible suggestion that it might be better to have lessons three times a week instead of every day.

The comments written by the children were on the whole quite perceptive, and will be considered further when modifications to the programme are suggested in the next chapter.

Chapter 6

Discussion and Conclusions

6.1 Introduction

The intention in this study was to test a training programme designed to improve comprehension of expository texts. It was not intended as a remedial intervention for below average readers, but rather aimed to improve the ability of all students to learn from expository texts. Many researchers have drawn attention to the need to develop this skill in the upper primary grades. For example, Armbruster et al (1991) note that American high school students have difficulty understanding expository texts, which these authors link, in part, to limited exposure to the genre in the primary years. B. Taylor (1986) observed that grade 6 students produced almost no mental gist after reading a few pages of a content area textbook. Berkowitz (1986, p. 162) concludes that Taylor's observation is a common one when she states:

> *"A common observation among both classroom teachers and reading researchers is that middle-grade students have difficulty reading*

and remembering what they have read in their content area textbooks. Because these textbooks become an increasingly more important mode for learning in the higher grade levels, it is imperative that middle-grade students learn how to read and study content area material."

In Australia, primary teachers do not use basal reading series as appears to be the case in the United States. Teachers to a greater extent design their own lessons, and it would be useful for them to have available effective materials for developing the ability to comprehend expository texts. Research cited in chapter three clearly shows that this ability can be enhanced in primary school children, but as Paris, Wasik and van der Westhuizen (1987, p. 38) note:

"Until interventions are more available in a format that is easily used by educators, experimental research is unlikely to be adopted in classroom settings."

In particular, Paris et al suggest that teachers should be able to use a programme without intensive in-service training, and that it should be able to be used with existing groups of students rather than requiring small group teaching. As set out in chapter three, the reciprocal teaching programme implemented by Palincsar has proved to be very effective in enhancing comprehension of expository texts, incorporating strategies which encourage macrostructure formation, activation of prior knowledge, and more active goal-setting and comprehension monitoring. However, in Palincsar (1987) half the teachers in the study found the reciprocal teaching dialogue difficult even after in-service training. In this study the largest group of children taught by the Palincsar method was 16, still considerably smaller than most normal classes. Thus, the present study attempted

to teach similar strategies to the ones used by Palincsar in a format that could potentially be easier for teachers to use and that was suitable for use with intact classes of normal size. The written format was intended to provide a framework within which a teacher could work, and written responses were intended to ensure maximum engagement by all students. The study also investigated the effectiveness of the programme with a mixed ability class in grade 5, in contrast to the below average readers in grade 7 taught in most of the Palincsar studies.

6.2 Comprehension tests

The experimental class showed a significant improvement in comprehension following three weeks of training. This improvement was maintained on the delayed test administered ten weeks after training, albeit with a lower percentage of explained variance, and occurred without students being prompted to use trained strategies during testing. Performance on practice passages during training also showed a steady improvement, indicating that the demonstrations of enhanced comprehension were robust and durable. Although students in the class did not have a high level of knowledge about the topics of the test passages prior to reading, scoring 39%, 41% and 42% on the pretest, posttest and delayed posttest knowledge tests respectively, they were able to answer an increasing number of comprehension questions correctly, indicating that they had improved their ability to learn new information from texts.

Macrostructure formation was a central component of the training programme; students were instructed to devote additional processing to headings and subheadings, to invent topic sentences for paragraphs, and to relate the topic

sentences to section headings. The improvements in comprehension resulting from this training provide some support for the view of van Dijk and Kintsch (1983) regarding the importance of macrostructure formation. According to these authors, the macrostructure provides a basis for logically connecting sequences of text, a means of organising information in memory, retrieval cues for the recall of this information, as well as a construction defining the most important elements of a text. Loman and Mayer (1983) show that when macrostructure formation is facilitated through signalling, students recall more conceptual information and are able to make better use of that information for creative problem solving. Winograd (1984, p. 420) concludes that *"the ability to identify important elements in a text is a strategic skill that underlies both comprehension and summarization."*

Van Dijk and Kintsch (1983) postulate that readers use many sources of information when forming a macrostructure, including titles, thematic statements and explicitly signalled text structure, as well as subject knowledge. Among adults, titles influence the construction of what a text is about and its recall (Bransford and Johnson, 1972; Schallert, 1976), and readers spend longer processing topic sentences than non-topic sentences (Lorch et al, 1985).

Among children, though, Winograd (1984) notes that poor readers have difficulty using textual cues to identify important information. Williams et al (1981) show that grade 5 children have not yet learned to identify and attend to topic sentences, and Garner and Gillingham (1987) similarly conclude that grade 5 students are not well equipped for tasks which require them to attend to structural properties of texts. Brown and Day (1983) observed that grade 5 children rarely invented topic sentences in their passage summaries, and did not even make good use of explicit topic sentences.

Given the importance of macrostructure formation and the poorly developed skills grade 5 children generally have in this area, it is likely that focusing on explicit structural cues and inventing topic sentences during training played an important role in the improved comprehension which resulted. Such a conclusion is consistent with other studies which have improved comprehension as a result of focusing, at least in part, on macrostructure formation, including the Palincsar studies, Salomon et al (1989), Stevens (1988), Kurtz and Borkowski (1987), Borkowski et al (1988) and Adams et al (1982).

However, it would be desirable to have data on the degree to which the trained strategies, in particular, paragraph summarisation, were mastered, in order to be able to analyse the association between strategy mastery and comprehension achievement more precisely. In Palincsar (1987), improvements in summarisation performance predicted improvements in comprehension, but this relationship did not occur in Palincsar, Brown and Martin (1987). As was noted previously, students' work booklets were not available for inspection, but in any case, it would be preferable to include a paragraph summarisation task in the battery of tests to ensure that students were working independently. This point will be referred to again in the final section of this chapter.

Comprehension improvements in this study were achieved without prompting students to use the trained strategies and were maintained for ten weeks after training, a delay longer than that investigated in most studies involving macrostrategies; only Brown and Palincsar (1982) and Palincsar and Brown (1984) investigated the maintenance of strategies over a similar or longer period. As indicated previously, Paris and Oka (1986a) consider that students will only use strategies without prompting if they consider them to be worthwhile and important. The

utility and rationale for the trained strategies was explained to students, most of whom could also observe for themselves improvements in comprehension on their individual progress charts during the course of training. It is a plausible inference, given the comprehension gains and their maintenance over ten weeks, that students were convinced of the value of the strategies, and therefore continued to use them, unprompted, over time. However, since such beliefs were not directly measured, this point remains conjectural.

Similarly speculative is the possibility that as a result of training students developed more sophisticated goals for reading expository texts. As Brown (1980, p. 454) notes: *"The reader's purpose determines how he or she sets about reading."* McKoon and Ratcliff (1992) concluded that global inferences were only made and used to guide local processing when readers had particular goals. Graesser et al (1980) observed that adult readers devoted more attention to macrostructure formation when they were reading in preparation for an essay rather than multiple-choice questions, thus supporting Brown's contention regarding the importance of goals. Lorch et al (1993) found that college students recognised different purposes for reading and adjusted their processing accordingly. Skilled adult readers can set different goals according to the task they are engaged in.

However, poor readers, whether children (Garner and Kraus, 1981-2) or adults (Gambrell and Heathington, 1981) may set goals which are too heavily focused on word-level processes, and may see the good reader as one who can read quickly and fluently. Even if goals are not limited to this extent, it is possible that children may focus on understanding individual sentences. The training provided a model of a different purpose to aim for, one focused on understanding main

ideas. In the comments from students quoted in the previous chapter, one boy noted that he no longer tried to read fast like in a competition, and another said that he now read to find out what a story was about rather than reading without learning anything. It may be that such comments reflect more sophisticated goals for reading. The possibility that reading goals evolve in the course of training could be tested by the use of questionnaires or interviews such as those used by Garner and Kraus (1981-2) and Gambrell and Heathington (1981). This point will be referred to again in the final section of the chapter.

6.3 Transfer tests

Results on the transfer tests were not as positive as on the comprehension tests. Nevertheless, error detection scores did move in the predicted direction and approached significance. On the pretest, the experimental class detected 36% of the errors, with this score increasing to 53% and 49% on the immediate and delayed posttests respectively. In contrast, the control class declined from an initial 45% to 44% and 38% on the immediate and delayed posttests respectively. Class differences approached significance when pretest scores were accounted for ($p < .10$ and $p < .07$ for immediate and delayed posttests respectively). The experimental class also tended to underline less unproblematic information ($p < .007$, $p < .08$ for immediate and delayed posttests respectively). Studies in which training effects transferred to error detection tests include Palincsar and Brown (1984), Paris and Jacobs (1984), Paris, Cross and Lipson (1984) and Paris and Oka (1986b). However, in the Paris studies the errors were nonsense words and scrambled clauses which, as will be discussed further below, are easier to detect

than internal inconsistencies. Also Paris and Oka note that improvement came about largely due to a reduction in false alarms rather than an increase in hits.

Several studies indicate that rates of detecting explicit internal inconsistencies are often low. For example, Garner and Kraus (1981-2) observed that detection of intersentence contradictions was low even among good readers in grade 7, with poor readers detecting none at all. In Reis and Spekman (1983), poor readers in grades 6 and 7 scored an average of 17% for detection of such errors. In Markman (1979), 43% of the children in grades 3, 5 and 6 missed all or all but one of the three explicit contradictions. Even the performance of adults leaves much room for improvement in this area (Grabe et al, 1991; Baker and Anderson, 1982; Baker, 1985b; Glenberg et al, 1982).

One reason why the detection of internal inconsistencies is difficult is because it requires constructive processing before the error can become apparent (Markman, 1979; Markman and Gorin, 1981; Beal, 1990). For example, Baker (1985b, p. 311) states that *"detection of internal inconsistencies involves the ability to retrieve and integrate nonadjacent propositions and draw logical inferences."* Skilled readers are normally unaware of the processing they are carrying out when all operations are proceeding smoothly. However, their efforts become conscious when integration operations fail (Capelli and Markman, 1982; Baker, 1985a). As Markman (1981, p. 75) notes: *"... much information about one's comprehension is a by-product of active attempts to understand and not just of attempts to monitor."*

Markman (1985) goes on to point out that the likelihood of detecting a particular error should vary with the amount of processing required. Considerable processing is required to detect internal inconsistencies, so it is not surprising that they are difficult errors to detect. Consistent with this point, Scardamalia

and Bereiter (1984) observed that it was easier to detect errors which could be found by questioning the plausibility of individual statements rather than requiring the integration of successive statements. Similarly, Garner and Kraus (1981-2) found that intrasentence contradictions were easier to detect than intersentence contradictions. Not surprisingly, Zabrucky and Moore (1989) concluded that the ability to detect internal inconsistencies developed later than the ability to detect nonsense words.

Considering that the detection of internal inconsistencies relies heavily on integration operations and awareness that they are not proceeding smoothly, two possible reasons can be suggested as to why children in particular find this type of error difficult to detect. Firstly, they may not carry out the necessary constructive processing or, secondly, even if they do, they may not devote sufficient attention to processing difficulties.

Baker (1984b) concludes that younger and poorer readers may see the goal of reading as understanding the words, rather than constructing larger units of meaning. Garner (1981) refers to 'piecemeal processing' carried out by poorer readers, in which they focus on word-level operations rather than integration. Consistent with the hypothesis of inadequate integration, Markman (1979) noted that children were inclined to question the plausibility of individual statements, rather than integrating successive statements to find contradictions. A similar observation was made by Tunmer et al (1983, p. 107), who comment that children such as those in their study *"question the truth of individual sentences rather than integrate the story as a whole and examine its overall logical structure."*

Inadequate integration could be one reason why readers fail to detect internal inconsistencies. However, even when children show signs of the integration dif-

ficulties which in adults often give rise to awareness of problems, they may still fail to report an error. For example, both Harris et al (1981) and Zabrucky and Ratner (1986) found similar increases in reading times in the region of an error for grade 3 and grade 6 children, but grade 3 children reported fewer errors, leading Zabrucky and Ratner to conclude that *"beginning readers process information in a constructive fashion and show rudimentary skills at applying both internal and external standards of evaluation, but ... their processing may not result in conscious awareness of non-comprehension or ... they may not know how to interpret such awareness"* (p. 1414).

The training programme to some extent addressed both these problems. The strategy of subsuming the individual sentences in a paragraph into one summary statement should encourage students to form larger units of meaning wherein contradictions would potentially be more apparent. Students were also encouraged to explicitly ask themselves whether there were any words or ideas that were unclear. Although Markman (1985) notes that adults do not usually engage in such self-questioning, this strategy was included to attempt to make students attend more closely to any processing difficulties they might have experienced, rather than overlooking such problems as in Harris et al (1981) and Zabrucky and Ratner (1986). While the training programme included strategies useful for error detection, it should be noted that students did not engage in any error detection tasks during training, so that this test involved transfer of skills to a novel task.

Epstein et al (1984, p. 355) emphasise the significance of error detection when they state that *"ability to detect a contradiction is very close to what we normally mean by comprehension. ... If a reader cannot detect a contradiction of a main idea (inserted into the text by an experimenter), then it is unlikely that the reader*

has understood the text as the author intended."

The improved error detection shown by the experimental class, even though it did not quite reach significance, at least suggests improved comprehension in the manner indicated by Epstein et al. If false alarms are an example of questioning the plausibility of individual statements rather than comparing successive statements, then the tendency to produce fewer false alarms and more hits suggests that these students were beginning to engage in more of the constructive processing necessary for contradictions to become apparent. It will be proposed in the final section of this chapter that the training period needs to be extended beyond three weeks. It is possible that class differences in error detection would reach significance if students had more time to consolidate the gains they were beginning to show in this study.

Turning now to the ACER test, the hypothesis of improved performance as a result of training was not supported. On the immediate posttest results were in the predicted direction once pretest scores were taken into account (p=.13), but the delayed posttest showed no difference between the classes.

Paris and his colleagues have not found improvements in standardised test performance as a result of their ISL programme, and have questioned the appropriateness of such instruments to measure the effectiveness of classroom interventions. For example, Paris, Cross and Lipson (1984, p. 1249-50) conclude:

> *"The GATES and TORC measure student achievement that is primarily determined by general intelligence and background experiences. They are good tests for discriminating students according to ability, but the instruments seem ill-suited as measures of variations in classroom curricula. One might argue that successful performance*

on standardized tests depends on comprehension strategies, but these seem dissimilar to many ISL strategies and, indeed, unlike many other ordinary reading activities of elementary school children. For example, rereading and paraphrasing in a deliberate fashion may be too time consuming while taking a standardized test."

Working under time pressure does make it more difficult to apply reflective comprehension strategies. Although students' reading time for comprehension tests was not formally measured in this study for reasons explained previously, informal observations indicated that following training the experimental class spent considerably longer reading test passages than the control class. This more reflective approach would be difficult to apply to the ACER test, when there are 39 questions to be answered in 30 minutes. Nevertheless, there was a movement in the predicted direction on the immediate posttest. Therefore, it is suggested that another possibility should be investigated before standardised tests such as the ACER are dismissed as inappropriate instruments for the purpose.

The longer time spent reading comprehension test passages suggests that students were still slow in the application of strategies, which were not yet sufficiently well practised to be applied efficiently. However, improved comprehension scores and a movement in this direction on the ACER immediate posttest suggests that at least the rudiments of the strategies had been learned. Thus, if the training period is extended, as recommended in the final section of this chapter, it is possible that strategies could become sufficiently well mastered to be applied under the time pressure of the ACER test.

Performance on the cloze tests showed no signs of improvement as a result of training. If the cloze tests measured intersentence integration as intended,

then this result is rather surprising, considering that there were signs of improvement on the error detection test, requiring the integration of nonadjacent sentences. It was suggested, following Shanahan, Kamil and Tobin (1982), that cloze tests with every nth word deleted are not a good measure of intersentence integration. If such tests do, to a considerable extent, measure *intra*sentence integration, then there would be no reason to assume that the present study, focused on macrostructure formation, would improve cloze performance.

In order to assess whether students will transfer their new skills to a test of intersentence integration, it is suggested that cloze tests with selectively deleted content words should be used, as in the studies by Carr et al (1983) and Dewitz et al (1987), where such tests were used as training to improve inferential comprehension.

6.4 Discriminating students who improve from those who do not

Eleven pretest variables were entered into a discriminant analysis in order to determine which combination of variables, if any, would distinguish those students who improved their comprehension as a result of training from those who did not. The criterion for improvement was set at a 10% gain from the pretest comprehension test to *both* the immediate and delayed posttests. A significant discriminant function was found, based on six variables. All 16 students who had complete data were correctly classified as either improvers or non-improvers according to this model.

An important component of the discriminant model was the relationship be-

tween decoding skill and comprehension measures; the improvers had high decoding skill (high reading speed, low non-distorting errors) *relative to* their performance of cloze and ACER tests. Put another way, it appeared that these students were not fully utilising the processing capacity which should have been available for comprehension tasks, given their level of decoding skill. In this regard these students resembled the students taught by Brown and Palincsar (1982) and Palincsar and Brown (1984), who were adequate decoders but below average comprehenders, except that in the present study the model applied to students of varying levels of comprehension ability. The non-improvers, on the other hand, had lower decoding skill relative to their cloze and ACER performance. It should be kept in mind, though, that the relationship between decoding skill and comprehension measures in the model was a relative one; the group of improvers did not show uniformly high decoding skill or low levels of comprehension in absolute terms, nor did the group of non-improvers show uniformly low decoding skill or high comprehension.

The implication of the discriminant model will be considered in relation to the view of working memory put forward by Daneman (1984). However, it must be noted that these comments are speculative, since the model was derived from a small sample of students.

Daneman (1984, p. 368-9) describes the relationship between processing efficiency, working memory capacity and comprehension performance as follows:

> *"The major tenet of the theory is that individual reader differences influence the integration processes of reading comprehension and that one crucial source of reader difference is processing efficiency in working memory. ... Slower and less efficient processes may consume more*

of the available capacity and so lessen the amount of additional capacity for storing and maintaining information in working memory. If working memory has less information subsequent processing will not be optimal."

The Reading Span test (Daneman and Carpenter, 1980) was designed to measure the relationship between storage and processing functions; those readers who processed sentences efficiently as they read aloud should have more capacity to store the final words for later recall. In support of Daneman's theory, large span readers processed difficult syntactic constructions more efficiently (Carpenter and Just, 1989). Large span readers were better able to recover from inconsistent interpretations (Daneman and Carpenter, 1983). Large span readers also showed more accurate pronoun assignment, fact retrieval and title generation, which Daneman and Carpenter (1980) attribute to their greater capacity to form rich and coherent chunks in memory, from which inferences and generalisations can more effectively be made.

In contrast, Walczyk (1990) found no relationship between subcomponent processing and the detection of explicit and implicit inconsistencies, leading him to state that: *"This implies a threshold beyond which efficient encoding, semantic access and logical proposition integration afford no facilitating effects for strategic reading activities, at least as measured by error detection"* (p. 494).

In the present study, overall processing efficiency was not measured; reading speed and accuracy are only aspects thereof. Also, the aim here was not to explain the level of comprehension achievement at a particular time, but rather to explain the capacity to improve comprehension, at whatever level, as a result of training. In spite of these differences to the situation described by Daneman (for

example, in Daneman and Carpenter, 1980), it is suggested that the Daneman model, rather than that of Walczyk, has relevance to the present study.

Assuming that a particular level of decoding skill, as one component of processing efficiency, leaves a certain capacity for comprehension processes, then the reader who is not fully utilising that capacity is the one who is likely to benefit from instruction. The reader who is fully utilising the available capacity, whether it be high or low, may have difficulty integrating new strategies into his or her repertoire and applying them to comprehension tasks.

As indicated, these comments are speculative, but they do have testable implications. If the comments have any validity, then increasing decoding speed and accuracy should increase the probability that a student's comprehension will improve as a result of training. In the final section of this chapter, it will be suggested that the last phase of the training programme could involve small group work where students read aloud to each other. This increased opportunity to practise existing decoding skills may have the effect of improving processing efficiency and creating increased capacity to master new comprehension strategies. Another possibility is to use a vocabulary expansion programme such as that developed by Beck et al (1982), in conjunction with strategy training. As explained in chapter one, this programme taught groups of thematically related words in such a way that an interconnected memory representation was formed and words could be rapidly accessed.

Within the context of the decoding-comprehension relationship described above, another variable which contributed to the discriminant model was attitude to reading. The improvers expressed relatively more positive attitudes than the non-improvers regarding the importance of reading, the enjoyment they de-

rived from it, and their ability as readers. It is not surprising to find that readers who believe that reading is a skill sufficiently important to be worth devoting additional effort to, and who have sufficient confidence in their ability to believe that this effort will pay off, are more likely to improve their comprehension. As Paris, Lipson and Wixson (1983, p. 204) state: *"... agents' choices of actions and goals are influenced by individual styles such as risk taking, achievement aspirations, self-concept, or fear of failure."*

As indicated in the previous chapter, the discriminant model derived from the sample of children in this study still needs to be tested on further samples from the population. Almost undoubtedly, the model will not classify such children as effectively as was the case here. In such an event, the inclusion of additional variables in a discriminant analysis may increase predictive power. For example, although attitude to reading was a variable already included in the model, it may be that a finer-grained measure of students' beliefs would enhance predictive power. Of the relationship between attitudes and performance, Carr, Borkowski and Maxwell (1991, p. 117) note: *"If the impetus for achievement is external to children (e.g., a belief in luck or the necessity of help from others), it is unlikely that they will develop feelings of self-esteem and a repertoire of high-level metacognitive skills (especially executive skills) necessary for good performance."*

Carr et al (1991) used an attribution questionnaire where students were asked to explain why they thought success and failure outcomes in specific examples had occurred. Butkowsky and Willows (1980) linked their investigation of attributions to actual performances by the children on solvable and unsolvable anagrams. This approach has the advantage of being able to investigate problem-solving persistence in the face of difficulty at the same time as attributions. Students who

lack persistence due to self-defeating attributions are unlikely to invest the effort necessary to successfully apply trained strategies, or to persist if new strategies are initially difficult to master.

Another variable which could be relevant is categorisation ability, such as being able to classify concepts into superordinate categories, but perhaps more importantly, being able to identify relationships between sentences and grouping them under a superordinate topic sentence. Williams (1984) considers that main idea identification rests on such classification skills, since detail sentences are exemplars of the category stated in the topic sentence, at least in simple enumeration paragraphs. While main idea comprehension in general is likely to involve more than such simple classification, having this skill is at least a foundation on which to build. This skill could be assessed by a task such as that used by Danner (1976), where he presented children with the individual sentences and topic sentences of a simple enumeration passage and asked them to place the sentences into coherent groups. Children lacking this ability may well find it more difficult to master the trained strategies.

6.5 Suggested changes and further research

In this final section, a number of changes to the teaching programme will be suggested, followed by a range of questions which could be addressed in future research, and an evaluation of the overall effectiveness of the programme.

A number of changes to the actual teaching programme could enhance its effectiveness. The aim of these changes would be to enable more students to achieve and maintain comprehension gains and to strengthen the signs of im-

provement that were evident on two of the three transfer tests. While a longer training period is desirable to achieve these ends, additional practice has to be organised in such a way as to avoid the boredom which was beginning to emerge at the end of three weeks of intensive training. Another problem to address is the perceived difficulty of applying the strategies independently in the last stage of training. Ideally training would also involve more oral reading practice since a few students were weak in this area, and it was speculated that more efficient decoding may allow more readers to benefit from training.

The first suggested change would be to simplify the introduction to the programme, both in terms of explaining the nature and value of the strategies and explaining the process of paragraph summarisation. In the current programme, learning of this process was heavily based on induction, that is, learning from the modelling of the teacher and working with exemplars of topic sentences provided in the work booklets. Observations of classroom performance indicated that some students found this inductive method more difficult than others. It is suggested that more direct explanation would be desirable, as included in the studies by Stevens (1988) and Kurtz and Borkowski (1987). Students could be taught a three-step procedure for generating summary sentences, namely, identify the topic of a paragraph as the concept or action referred to repeatedly in the sentences of the paragraph, identify what is said about the concept or action, and combine the two into a succinct summary sentence. As in the Stevens study, students could also be taught to check their own summary statement against each sentence in the paragraph to evaluate whether the summary is sufficiently general and inclusive, thus strengthening the self-control component of the instruction (Brown, Campione and Day, 1981; Brown and Palincsar, 1982).

The introduction to the programme was intended to inform students of the value of the strategies they were being taught. However, in hindsight this section was overly ambitious, being too long and detailed for grade 5 students. If the revised programme begins with single paragraph summarisation only, as outlined above, then this introduction could be broken up accordingly. In other words, the value of summarisation would be explained during this practice, demonstrating through concrete examples how the process enhances recall of important information and provides one means of monitoring the adequacy of comprehension. The remainder of the introduction, now shortened by the removal of the above, would then be linked to the introduction of the full programme, retaining only the most important points of how the strategies assist understanding and recall.

The simplification of the beginning of the programme, in particular the more direct instruction of paragraph summarisation, may enable more students to master the trained skills and may lessen the perceived difficulty of the last phase of instruction when students are applying the strategies more independently.

Three weeks of intensive instruction was probably not the most productive timetable for these grade 5 students. On the one hand, at least some students needed additional practice to fully master the strategies and to apply them efficiently, but on the other hand, the class was beginning to become restless by the end of the three weeks. It is suggested, therefore, that after an intensive period of training to familiarise students with the strategies, instruction take place less frequently and change from class work to small group work.

The end of phase two would appear to be a suitable point to make this change, since at this point students have been introduced to the strategies, have practised identifying correct options from a multiple-choice format in response to the key

self-questions, and have generated their own responses to the questions. They are now ready to both ask and answer the questions. This third phase could begin as a class activity, perhaps two or three times a week. Once the teacher is satisfied that students understand the procedure, further practice can take place in mixed ability groups of three or four students. These students take turns in reading aloud a paragraph to the group and applying the strategies to that paragraph. The group provides feedback on the accuracy of oral reading and the adequacy of strategy application. Having received about three weeks of class instruction, students should be in a position to work more independently of the teacher, at least in an environment where they are supported by peers. Being in a small group where students know what to do should still ensure a high proportion of academically engaged time.

The same procedure could be used for spaced review as part of content area learning. In other words, whenever a topic is introduced in science or social studies, for example, some reading could be included regardless of how the topic is otherwise to be taught. Students would then divide into their groups and apply the trained strategies to this reading.

These suggested changes are intended to have several effects. Firstly, it was noted that some students in the experimental class would benefit from further oral reading practice. In the small groups, each student could spend more time reading aloud than would be possible in a whole class. Stevens, Madden, Slavin and Farnish (1987) report that oral reading improved when students spent increased time reading to and receiving feedback from a partner. As speculated in the previous section, it could be that improved decoding skills could increase the likelihood of making comprehension gains as a result of training, even for students

who are not poor decoders.

Secondly, the extended practice in applying the strategies orally rather than in writing may make it more likely that the strategies will become part of students' normal reading. This point was suggested in relation to the study by Berkowitz (1986), where trained students showed improved recall only when they were allowed to construct written maps. Ideally the strategies should be used by students even when no written product is required.

Considering these suggested changes and points raised in previous sections, there are several questions which could be the subject of future research:

- If the suggested changes were made, would comprehension gains be larger, and would more students reach the criterion for improvement? Would significant gains also be evident on the error detection and ACER tests?
- To what extent is improvement in paragraph summarisation linked to improvement in comprehension tests?
- Do students develop more sophisticated reading goals for expository texts as a result of the training?
- Does training result in improved intersentence integration, as measured by a cloze test with selectively deleted content words?
- How well does the discriminant model predict which students will benefit from training in further samples from the grade 5 population? Could a more powerful model be developed by the inclusion in the analysis of additional variables, such as attributions for success and failure and categorisation ability?

- If reading speed and accuracy are increased, for example, through increased oral reading practice and/or a vocabulary expansion programme, will more students benefit from training?

Improvements could be made to the training programme, but in its present form it can be considered a qualified success. It produced durable comprehension gains in an intact, mixed ability grade 5 class. Even though these gains were not shown by all children, and the transfer to other tasks was not as great as predicted, the results suggest that this approach is feasible with normal classes of this age to improve understanding and recall of expository texts.

One limitation of the study is clearly that only one class was taught. The particular Catholic parish school involved had a high percentage of children of Italian descent. Although a range of reading abilities was represented, overall the class was slightly below average according to national norms. However, there was no interaction between treatment and reading ability or treatment and ethnicity, making it more likely that the findings of this study can be generalised to the population of grade 5 children in Australia. Nevertheless, the findings may not apply to children who are very different to those in the sample, including more recently arrived migrants, Aboriginal children in isolated schools, or children well outside the ability range represented in this sample.

Finally, the written formal worked well, both from the point of view of producing a high level of student engagement, since even in a large class every individual is required to respond to each paragraph, and from the point of view of providing a framework to assist the instructor in organising the lessons. This approach provides a basis for the easily used materials advocated by Paris et al (1987). However, it is important to emphasise that the written material is a framework

only, and would lose much of its value if used mechanically as a series of teacher-directed exercises. It is essential that the teacher encourage the students to use the framework to direct their own processing and discussion thereof. By the end of the programme the teacher should need to make only a minimal input to the application and discussion of strategies. If the teacher continues to be directive, it is unlikely that students will develop the executive control necessary for maintenance and transfer of strategies.

While the original aim of this study was, in part at least, achieved, two further questions remain to be investigated to extend the educational application of the findings, namely:

- Is there an age at which this type of training produces optimum comprehension gains, or could the programme, with texts written at the appropriate grade level, be implemented with equal effectiveness in any class in the upper primary school?

- Can classroom teachers easily be trained to use the programme, and what kinds of comprehension gains are achieved when classroom teachers conduct the training?

Appendix A:

Teaching Texts

Life on the Great Barrier Reef

How a Reef is Built

1 The Great Barrier Reef is the largest coral reef in the world. Coral can only grow in some places. The water must be warm all year round. The water must be shallow because in deep water the coral would not get enough light. The water must also be calm because rough surf would destroy the coral. The coast of Queensland is perfect, and that's why the Great Barrier Reef has been able to grow so well.

2 There are many types of hard coral on the reef. Some are round and look like human brains, while others are long and thin and look like deer antlers. Some grow into big, flat plates, while others look like the underside of a mushroom. There are also some soft corals, which don't feel rocky, but are leathery and slimy to touch.

3 The hard corals look like white rocks, but these rocks are made by small animals called polyps. Each polyp makes a hard shell around itself. As thousands of polyps each make a shell, the white rock starts to grow. If you look closely at a coral, you'll see a lot of minute holes. That's where the polyps live. During the day they stay inside their rocky shell, but during the night they stretch out to feed.

4 Polyps have bright colours and look like flowers, with petals around a centre, but they are really carnivorous animals. The ''petals'' are tentacles for catching prey and the ''centre'' is a greedy mouth. Each tentacle has thousands of stinging needles, which are fired into small animals. The victim is held and poisoned by the needles. The tentacles then pull it into the mouth.

5 Polyps are very small, but they have some larger relatives on the reef. All the members of this family have stinging tentacles, and include sea anemones and jellyfish. Sea anemones live on the sea floor with their tentacles pointing up to catch fish as they swim past. Jellyfish float in the sea and trail their long tentacles in the water below them. Our skin is too thick for the stinging needles of some of these animals, but some jellyfish can give us a nasty sting.

Reef Animals with Shells

6 Many animals with shells live on the reef. Seasnails are one group of animals with a shell. They are a bit like garden snails, but they live in water and can be much bigger. The largest seasnails are called Giant Tritons. Their pointed shells may be half a metre long, so they really are giants! They are large enough to eat starfish. Their shells have beautiful shapes and markings, but beauty can bring danger, and many Giant Tritons have been killed for their shells.

7 Snail shells are in one piece, but some other shells have two parts which can open and close. These shells are called bivalves. They work a bit like the valve on a tyre, which controls how much air goes in and out of the tyre. Bivalves control how much water goes in and out of the shell. They get their food by filtering small particles out of the water, so they have to let water in. But they have to close to protect themselves. Otherwise, at low tide, they would dry out. They also close if there is danger.

8 The largest bivalve is the Giant Clam. Large clams may be over one hundred years old, over one metre long and weigh over two hundred kilograms. Their flesh is a pretty blue-green colour. Like other bivalves, they eat small particles from the water. Some movies show people being caught when Giant Clams close. It is true that the shells are strong, but they close very slowly, so it is most unlikely that a person could get caught.

9 Animals like crabs have a different type of shell. They have a hard skeleton around the outside of their bodies. These animals are called crustaceans because they have a hard ''crust'' around them. They all have legs, antennae and claws or nippers to hold their food. Some of these animals are too small for us to see. Lobsters, crabs and prawns are some of the larger members of this family. They thrive on the reef.

Fish on the Barrier Reef

10 There are hundreds of different types of fish on the reef. Huge numbers of fish can live there because there are good supplies of food and a lot of places to shelter in the coral. Many reef fish are very bright. Some are small, but some, like the sharks and stingrays, are quite large. They have a wide range of feeding habits.

11 Stonefish get their name because they look like stones. They are a dull brown colour and have lumpy skin. They sit very still on the sea floor, often half buried in sand. They are hard to see because of their perfect camouflage. When a fish swims past, the stonefish quickly sucks it into its mouth. Stonefish have poisonous spines which can kill people.

12 Clownfish are very brightly coloured small fish which live among the tentacles of sea anemones. How do they manage, when sea anemones have poisonous stinging needles? The clownfish are immune to this poison but bigger fish, which might eat the clownfish, are not immune. So the little clownfish is safe in the sea anemone. In return, the clownfish keeps the anemone clean by eating algae which grows on it.

13 Cleanerfish do just that, clean other fish. These small animals wait on a flat piece of rock. Larger fish swim into this ''cleaning station'' like we might go to a carwash. The cleanerfish dart around the face and gills of their customers, picking off parasites, algae and dirt. They even swim into the mouths of large fish. They are in no danger because no-one ever eats a cleanerfish. They are much too useful.

14 Sharks are some of the largest fish on the reef. Most of the dangerous ones, like the tiger and hammerhead sharks, don't like the shallow water around the coral and stay further out to sea. The four types that live on the reef are harmless. They are fairly small, being less than two metres long. With so many fish on the reef, sharks have no trouble finding food.

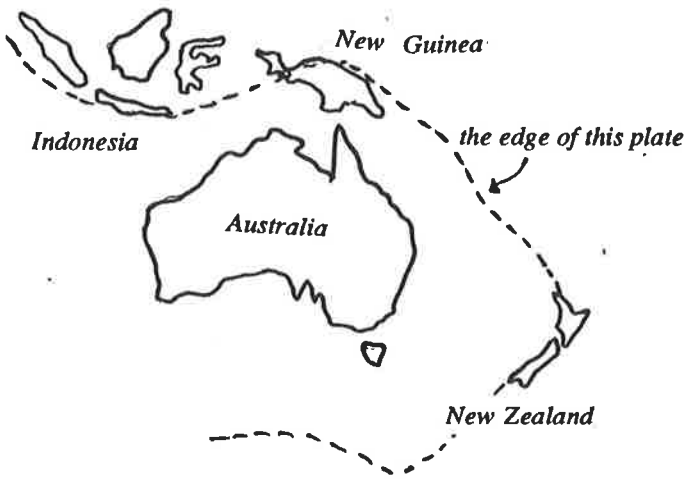
Natural Disasters

If we want to know why we get earthquakes and volcanoes, we first have to look at how the Earth was formed many, many years ago.

How the Earth was Formed

1 When the Earth was first formed, it was so hot that all the rocks were still molten. Over a very long time the surface cooled down and slowly became solid. This solid crust is very thick, but under the crust the rock is still hot and liquid, a bit like toffee.

2 The crust of the Earth is not just one piece, but is made up of a few big pieces called plates. These plates carry the land and the seas. The plates are like rafts that float on the liquid rock inside the Earth. Because they float, these huge plates can move about, but they move very slowly.



The diagram shows our part of the Earth's crust. Australia is in the middle of a plate. New Zealand, New Guinea and Indonesia are on the edge of this plate.

Earthquakes

3 When there is an earthquake, the ground shakes and there may be large cracks in the ground. Buildings collapse and people may be killed. Earthquakes are caused by the Earth's plates bumping into each other. When two plates are near each other but don't move the same way, they put a lot of pressure on the rocks deep in the ground. Over a long time the rocks split and shift and the ground shakes.

4 The force of an earthquake is greatest where the rocks start to move. This point is called the epicentre. From there shock waves spread out, just like ripples spread out when you drop a stone into water. As the shock waves spread out, they slowly weaken. A long way from the epicentre, people might feel only a slight tremor in the ground.

5 The places most likely to have earthquakes are those on the edge of plates. In this country we are fairly safe. We might feel earth tremors from time to time, but there is not much chance that we will have big earthquakes. Japan and some parts of the United States, though, are in much greater danger because they are on the edge of plates. Many years ago, the city of San Francisco came close to being destroyed by a huge quake.

6 These days experts can predict earthquakes with the help of instruments that can pick up even slight temors in the ground. People can also be warned of earthquakes by animals. Some animals act strangely before a quake. It may be that they can feel tremors too small for us to notice. A long time ago, the Chinese used a clever method. They made a pot with dragons' heads all around it. They put a metal ball between the teeth of each dragon. Even slight tremors would make the balls fall, so the Chinese had an early warning of earthquakes.

Volcanoes

7 A volcano is a place where hot gas and liquid rock escape from the inside of the Earth. If this mixture can come through the ground freely, there is no explosive eruption. The hot rock, called lava, just pours through the Earth's crust. In places where this happens, you may also find pools of boiling mud, and gas with a foul smell. There may be jets of hot water, called geysers, that shoot into the air.

8 But if the volcano is blocked by old, hard lava, pressure builds up inside it. Over a long time, this pressure becomes so great that there is a huge eruption. Smoke and lava shoot high into the air. The volcano cone can even explode into small pieces. It is a bit like shaking a bottle of fizzy drink. Pressure builds up and may even break the bottle. If you take the top off the bottle, the pressure inside will make the drink shoot out.

9 As with earthquakes, you are most likely to find volcanoes on the edge of plates. Liquid rock comes through the Earth's crust where two plates meet. New Zealand and the United States still have active volcanoes. Australia doesn't, because we are in the centre of a plate.

10 There was a huge eruption on the island of Krakatoa in Indonesia about one hundred years ago. The volcano had been dormant for nearly two hundred years when it started to rumble again. There was a series of mighty explosions. They could be heard even in places a long way from the island. Most of the island exploded into small pieces. So much ash and dust shot into the air that the sun's light was blocked out for three days.

11 The explosion was so great that it caused huge waves around the island. The waves spread out to many other islands. A wall of water swamped houses all around the coasts. Forty thousand people were drowned. These waves killed more people than the explosion itself.

12 Another famous eruption took place nearly two thousand years ago in Italy. The people there thought that the volcano Vesuvius was extinct. They built fine towns around it and grew crops on the rich volcanic soil. But Vesuvius was only dormant, and well plugged with solid rock. It rumbled for sixteen years and in the end shook the plug loose. It erupted with great force.

13 Hot ash and lava poured down on the town of Pompeii. Many people choked under the cover of ash. It rained at the same time, and the rain mixed with the ash to make a mixture like cement. The whole town was covered. For hundreds of years the town lay hidden and no-one gave it much thought.

14 In more recent times, Pompeii has been uncovered. The buildings were well preserved under the ash. You can now walk through the town and see what Roman buildings looked like. You can also see models of what the people and animals looked like before they died. As ash hardened around their bodies, it left an outline of them. Scientists have poured plaster into these hollows to make models of the victims.

Animal Defence Systems

1 No animal can live on its own, and each one plays its part in the complex web of life. Some animals eat only plants. They in turn are eaten by other animals, who may themselves be eaten by larger animals. But the hunted have many ways to protect themselves from their hunters.

Appearance - Hiding from a Hunter

2 The simplest way to hide and avoid being caught is to have the same shape and colour as the surroundings. The hunter just can't see its prey. This sort of look is called camouflage, and is used by many animals.

3 The ptarmigan lives in the far north near the Arctic. It is a plump bird which can't fly well and spends most of its time on the ground. Foxes and birds of prey like to eat the ptarmigan. Its only defence is to hide. In summer, its feathers are mottly brown to blend into the rocks and low grasses. In winter, it is pure white to blend into the snow.

4 Many fish also use camouflage. The flounder is a weak swimmer that lives near the seafloor. Its body is flat so that it can stay close to the sandy bottom, and its mottly brown colour blends into the sand. The flounder can't escape being caught by using speed, so it tries not to be seen.

5 Ground-nesting birds could be caught while they sit on their eggs, so some of them also use camouflage. The bittern is a water bird with a long neck. It makes its nest in tall, brown reeds. It has brown stripes running up and down its neck. If the bittern feels in danger while sitting on its eggs, it stretches up its long neck. The stripes make it look just like the reeds.

6 Some animals use their markings not to hide, but to trick hunters. Many butterflies have dark spots, like eyes, on the edge of their wings. When birds try to catch them, they mostly peck at these spots. The butterfly loses the edge of its wing, but it can still fly away. It would most likely be killed if birds pecked the body instead.

7 Some insects copy both the colour and the shape of the places where they live. Some caterpillars have long, thin bodies. They are brown and have small bumps on them to make them look just like the twigs they live on. Leaf insects have wide, flat bodies like the shape of leaves. Even their legs have wavy edges to make them look like leaves with the edges eaten off. Birds and frogs looking for food can only see them if they move.

8 Some animals pretend to be someone else to protect themselves. The cleanerfish is a small fish that is not eaten by larger fish. The reason is that it does a very useful job by cleaning other fish. It picks dirt and bits of food from the mouth and skin of its customers. But another fish looks just like the cleanerfish, and so is not eaten. When big fish swim up to this imposter to be cleaned, it attacks them. It tears off and eats small bits of their fins and flesh. It makes use of the good reputation of the cleanerfish to avoid being eaten.

9 Some animals don't have any sort of camouflage, but are bright and easy to see. Why aren't they caught by hunters? The reason is that bright colours are often a warning that an animal is poisonous or just tastes awful. Hairy caterpillars are often brightly coloured. If the hairs on them are touched, they can cause a nasty rash. Some colourful frogs have poison in their skins. It is safer for a hunter to avoid animals with bright colours.

Behaviour - Bluffing a Hunter

10 Many animals are not able to cause any real harm to their preditors. Some of them, though, pretend that they are more fierce than they really are. Others behave in a way which helps them to escape from their hunters.

11 One of the main ways of ''acting tough'' is for an animal to make itself look bigger than it is. The frill-necked lizard is only small. But when it erects its frill and opens its mouth wide, it looks much bigger and quite fierce. The pufferfish can fill its stomach with air or water when it is in danger. It puffs up into a ball and makes the spines on its skin stand up. This prickly ball is enough to scare off most fish who would like to eat it.

12 Some ground-nesting birds use bluff to keep preditors away from their eggs and chicks. Seagulls attack the nests of terns, but terns can't harm the larger seagull. When a lot of terns fly at the seagull and make a lot of noise, though, they seem quite fierce. Mostly this display is enough to scare the seagull away from the nests.

13 Some birds use a different kind of bluff to lure preditors away from their nests. The adult pretends to be injured. It limps away from its nest with one of its wings drooping. The hunter follows the ''injured'' bird, ready to get an easy meal. When the adult is far enough away from its nest, it takes off and flies back to its young.

14 Many animals have parts of their body which help them to escape. Some lizards can drop their tails when they are grabbed by this part of the body. The tail keeps wiggling and distracts the preditor while the lizard runs away. Some starfish can drop one of their arms when they are touched. These arms, like the lizards' tails, grow back after a while.

15 The octopus and all its relatives use a smoke screen to help them escape from danger. When they are frightened, they squirt a jet of dark ink into the water. The enemy can't see in the murky water, so the octopus has time to get away.

16 Some animals have well protected backs, but soft bellies that could be attacked. The echidna has spines all over its back. The armadillo has armour plates over its back. Both have only fur on their belly. To protect this part, the echidna digs into the ground as soon as it feels in danger. It leaves only a pile of spines above the ground. The armadillo rolls into a hard ball, with all its soft parts protected on the inside.

17 The most straight-forward way to escape a preditor is to run. Antelopes are good runners. A healthy adult can run much faster than its hunters, such as lions. Springboks also leap into the air as they run. This bouncy leap with four stiff legs is called pronking. It warns other animals of danger. It may also be confusing to the hunter if animals are bouncing out of the tall grass in all directions.

Behaviour - Attacking a Hunter

Some animals can attack their enemies in a number of ways.

18 Many animals have horns or tusks. Musk oxen make good use of their curved horns. These large, hairy cattle live in the cold of the Arctic. The big adults are safe, but calves could be attacked by wolves. When in danger, the large bulls form a circle around the young animals. They use their horns to butt the wolves, and trample them with their hooves. The wolves can't get to the weaker animals inside the circle.

19 One of the best known animals to use a chemical defence is the skunk. If it is in danger, it turns its back to its enemy and lifts its tail. It may even stand on its front legs. If the hunter comes closer, the skunk squirts out two jets of liquid. The smell of this liquid is so repulsive that most animals change their mind about trying to eat a skunk.

20 An insect that makes use of chemicals is called the bombadier beetle. It has stores of chemicals in its body. They are mixed together when the beetle is threatened. It turns its back to its enemy and squirts out a jet of this liquid. Not only is the smell awful, but the liquid is also hot. Preditors like frogs take a while to recover when they have been hit by a bombadier beetle.

The Age of Dinosaurs

Dinosaurs lived a long time ago, but life on earth began a long time before then.

The Beginning of Life

1 A very long time ago, there was heavy rain most of the time and the earth was flooded with water. Life began in this ocean. The first living things were minute plants and animals, no more than specks of jelly.

2 Over many years these tiny specks changed into all sorts of plants and animals. All of them still lived in the sea. As new ones appeared, some old ones died out. After a very long time, creatures a bit like our modern fish first appeared. These fish were the ancestors of the dinosaurs and other land animals. But why did some of these fish leave the water?

3 A very long time ago there were huge movements in the earth's crust. Some land appeared above the water, although this land was still wet and swampy. There were a lot of benefits for the fish that could live in this swamp. They could escape bigger fish in the water and reach food other fish could not get. But if their waterhole dried out, they had to be able to crawl over the mud to find a new pool.

4 To be able to come ashore, some fish developed lungs to be able to breathe in the air. Their front fins grew stronger to pull them across the mud. These fish slowly changed into amphibians, a bit like the frogs we have today. They could breathe, feed and move on land, but they had soft skins. They had to keep returning to water to keep their skins moist. They also had very soft eggs, which had to be laid in water. They could never move far from the swamps.

5 Over a very long time, some of these frog-like creatures evolved into reptiles. Reptiles were much better suited to life on the land. They had far stronger legs and could run across the earth and rock. They had tough skins which didn't dry out and laid eggs with hard shells. The reptiles no longer had to stay close to the swamps.

When Dinosaurs Ruled

6 The first reptiles were still small animals, but during the many years that reptiles ruled the world, some grew very big indeed. The earth was still warm and wet. Huge jungles grew in this climate, so even large animals had plenty of plants to eat. With so many plant-eating reptiles, the hunters also had plenty to eat.

7 The largest dinosaurs were gentle giants. Their bodies were huge and could weigh as much as seven elephants. They carried this great weight on four sturdy legs. These reptiles had long necks to reach leaves on tree tops. Diplodocus was the largest land animal that has ever lived. It measured thirty metres from its tiny head to the end of its long tail.

8 Not all the giant reptiles ate plants. The largest meat-eating dinosaur was Tyrannosaurus rex. This name means ''king of the tyrant lizards'', and it really was a tyrant. It was as tall as a giraffe and ran upright on two very strong hind legs. It had sharp claws to grab its victims and rows of razor-sharp teeth to tear their flesh. It measured fifteen metres from its huge head to its tail. It is the largest carnivore that has ever lived on land.

9 But even Tyrannosaurus rex did not always have things its own way. Some plant-eating dinosaurs had a thick armour. Triceratops was the size of an elephant, with three long, pointed horns on its head and a frilled, bony plate on top of its skull. When in danger, it charged at its attacker. Even the ''tyrant lizard'' may have retreated from the sharp horns.

10 Not all dinosaurs lived on the land. Some reptiles took to the air. A tough skin stretched between their long front legs and their short back legs. They could glide down from cliff tops to catch fish. Other reptiles lived in the sea, and their legs changed into flippers to help them swim.

11 Not all dinosaurs were large. Even while the giants roamed the earth, there were many much smaller reptiles. Some were only the size of a chicken or an ostrich. These animals were mostly insectivores or ate the eggs of other dinosaurs.

The End of the Dinosaurs

12 Most scientists now think that a change in the earth's climate caused the giant reptiles to die out. There are two reasons why a cold climate would not suit dinosaurs. The huge jungles which could feed animals the size of Diplodocus needed hot, wet weather to grow. In cold weather the jungles would die, and so would the huge plant-eaters. With fewer animals to catch as prey, the large hunters also would starve.

13 The second reason is that reptiles can't make their own body heat. They can only warm themselves by basking in the sun. In a cold climate, they also become cold and freeze to death. The smaller reptiles could find places to shelter, but the others were too large to be able to escape cold winds and rains.

14 Some of the smaller animals didn't die out. The crocodile is the largest reptile that has survived from this time. There were also small, furry creatures, a bit like rats, and the first creatures with feathers. With the giants out of the way, these smaller animals took over the land. Over thousands of years they changed into all the animals that we see today.

How Snakes Use Their Senses (additional text)

All snakes have to be able to catch prey to survive. They can't move very fast, except to strike, so they depend on their senses. But in some ways these senses aren't very good because snakes are deaf, their daytime sight is not as good as ours, and they don't have noses to smell with like we do.

Snakes have sensitive scales on their belly that let them feel even slight movements in the ground. So, even though they can't hear their prey coming closer, they can feel it. This may be the first signal to the snake that a meal is close by.

Most people have seen snakes flicking their tongues in and out. What the snake is doing is "tasting" the air. The tongue picks up small particles in the air. When the tongue is pulled in, these small bits are flicked onto the roof of the snake's mouth. From there, the snake can tell what the taste is, for example, another snake, a frog or a mouse.

The snake's tongue makes it a very good tracker, better than a dog. This is an important sense for the snake. Often when it bites its prey, the victim will run away and die in a hidden place. The snake needs to use its tongue to find the dead animal.

All snakes use their belly scales and their tongues, but pythons use even more senses. Pythons are the largest snakes in Australia. They hunt at night and choke rather than poison their prey. Their special eyes let them use even bad light to get an outline of their prey. They can see in the dark, rather like cats.

But there is more to this story. Pythons have small pits around their jaw. These pits have heat sensors in them that find anything with a higher body temperature than the snake. Any warm blooded animal, like a mouse, is not safe even metres away from the snake. Cold blooded prey, like lizards and frogs, though, are safe because they don't give off any heat for the snake to feel.

The special senses of the python make possible the tricky job of catching bats in pitch dark caves.

Appendix B:

Student Work Booklet

BEING A BETTER READER

There are a few simple things that everyone can do to help them understand what they read.

Collecting Ideas

Whenever you see a heading, stop for a moment and ask yourself these questions:

- What do I know about this topic?
- What might this story be about?

Lets look at some examples to see how collecting ideas helps you to understand what you are reading.

The haystack was important because the cloth ripped.

Does the sentence make sense? Lets collect some ideas before we look at it again. **(Write heading on board)**

How did collecting ideas help you to understand this sentence?

_____

_____

Lets look at another example and then collect some ideas. **(Write heading on board)**

He hid under the bushy tops when he heard the noise.

How did collecting ideas help you to understand this sentence?

_____

_____

In a general way, how would you say that collecting ideas helps you to understand?

_____

_____

Putting Ideas Together

When you come to the end of a paragraph, stop for a moment and ask yourself these questions:

- What is the main idea of this paragraph?
- How does this paragraph fit into what I've already read?

Main Ideas

The main idea of a paragraph is one sentence that sums up what that paragraph is about. For example:

He went into the fruit shop and bought some bananas, apples and oranges. Then he bought some potatoes, beans and carrots for dinner.

We could sum up the main idea as something like: *He went into the fruit shop and bought some fruit and vegetables.* Notice that in a main idea sentence we leave out little details, like what sort of fruit and vegetables he bought. Lets look at another example.

He took a pair of trousers and three shirts from the wardrobe. He collected some socks and underwear from a drawer and put them with the other clothes. He went to the bathroom to get his toothbrush. Then he got out a suitcase and packed everything into it.

How could we sum up the main idea of this paragraph?

Which do you think would be easier to remember, the paragraph or the main idea sentence?

Lets look at another example.

..... No, nothing there would do. ..... No, they were too hard. ..... What could she do with that? ..... Now there was something .....

Can we put these ideas together? Could we sum up the main idea in one sentence? Does a paragraph make sense if we don't (or can't) put the ideas together? Lets look at the whole paragraph now and think again about the main idea.

She opened the fridge and had a good look inside. No, nothing there would do. She looked through the cupboard next to the fridge. Some dry biscuits? No, they were too hard. A tin of beetroot. What could she do with that? She got a chair to reach the top cupboard. Now there was something. A tin of soup would do nicely!

What is the main idea of this paragraph?

How does putting the ideas in a paragraph together into a main idea sentence help you to understand the paragraph?

Lets look at one last example. See if you can make up a main idea sentence.

She skipped down the path. The greengrocer packed the tomatoes. The car skidded across the road and hit a pole. The dogs played with the ball.

What do you think it means if you can't make up a main idea sentence? What should you do when this happens?

_____

_____

To sum up, in what three ways does putting ideas together into a main idea sentence help you when you are reading?

1.
_____

2.
_____

3.
_____

Fitting Paragraphs Together

Think about the paragraph in the last section which described a person looking for food. This is the next paragraph.

She heated up some soup and put it into the plate on her tray. She carefully carried the tray upstairs. Grandma was awake now. Poor Grandma. She was so weak, and she hardly had any teeth left. At least now she would be able to eat some soup.

How does this paragraph fit in with the one we read before?

_____

_____

How does fitting the paragraphs together help us to understand more about what we've read?

_____

_____

Checking and Clearing up

When you finish reading a paragraph, you should make sure that all the words and ideas make sense before you go on to the next paragraph. Lets look at an example.

He felt embarrassed by his dishevelled appearance as he staggered out of the forest into the town. He hadn't seen a bathroom or changed his clothes for three days. He made his way home as unobtrusively as possible. It worked, because no-one paid any attention to him.

What do the words *dishevelled* and *unobtrusive* mean? How well do you understand the paragraph without knowing these words? Can you work out from the paragraph what the words mean? How could you do that?

Why do you think it is a good idea to check for words you don't know and to try to clear up their meaning?

Sometimes the ideas in a story don't make sense to us. There can be many reasons for this. Maybe the ideas are ones we don't know much about. Maybe we weren't concentrating very well. In this case it is a good idea to read the paragraph again before we go on. But sometimes stories don't give us every single bit of information. If they did, we might never finish a book because it would be so long! To really understand what we are reading we sometimes have to fill in these missing bits. Lets look at an example.

He threw a cigarette butt out of the window as they drove through the forest. She said he shouldn't do that, but he paid no attention. She was right, though, because the fire destroyed many acres of the forest.

Does this paragraph make sense? Is there anything that needs clearing up?

How do you think it helps us to check how well we understand ideas and to clear up anything that doesn't make sense?

Life on the Great Barrier Reef

Collecting Ideas: Main Heading

What do I know about life on the Great Barrier Reef? What might this text be about?

_____

_____

Collecting Ideas: Heading

What do I know about how a reef is built? What might this section be about?

_____

_____

Putting Ideas Together: (Paragraph 1)

What is the main idea in this paragraph?

a The water has to be calm because rough surf would destroy the coral.

b We have the largest coral reef in the world because conditions in Queensland are just right.

c Coral needs warm water, calm water and shallow water.

How does this paragraph fit in with what I've already read? (the heading)

a It says that Queensland is perfect.

b It says that the Great Barrier Reef is the largest in the world.

c It says where a coral reef can be built.

Checking and Clearing up:

Are there any words I don't understand? What can I do about it?

destroy shallow reef perfect coast _____

Do all the ideas make sense? Is there anything that isn't explained?

- Why does coral need light?
- Why does the water have to be warm?
- Anything else? _____

Putting Ideas Together: (Paragraph 2)

What is the main idea in this paragraph?

a Soft corals are leathery and slimy.

b Hard corals can have many different shapes.

c Corals can look like brains, antlers, plates or mushrooms.

How does this paragraph fit in with what I've already read?

a It says that some corals look like brains.

b It says that some corals are soft.

c It says what sorts of corals can be made.

Checking and Clearing up:

Are there any words I don't understand? What can I do about it?

antlers leathery underside _____

Do all the ideas make sense? Is there anything that isn't explained?

- What is the difference between soft and hard coral?
- How do corals grow into different shapes?
- Anything else? _____

Putting Ideas Together: (Paragraph 3)

What is the main idea in this paragraph?

a Hard coral is made by little animals making a shell around themselves.

b Polyps are small animals that come out to feed at night.

c Hard corals look like white rocks.

How does this paragraph fit in with what I've already read?

a It says what hard corals look like.

b It says how a hard coral grows.

c It says what polyps are.

Checking and Clearing up:

Are there any words I don't understand? What can I do about it?

minute polyp _____

Do all the ideas make sense? Is there anything that isn't explained?

- Why can't you see polyps when you look at a piece of coral?
- Why do polyps make a hard shell?
- Why does it takes thousands of polyps to make a coral?
- Anything else? _____

Putting Ideas Together: (Paragraph 4)

What is the main idea in this paragraph?

a The pretty polyps poison small animals with stinging tentacles.

b Polyps have tentacles which make them look like pretty flowers.

c Tentacles have stinging needles which are fired into small animals, and when the animal is poisoned the tentacles pull it into the polyp's mouth.

How does this paragraph fit in with what I've already read?

a It says that polyps have bright colours and look like pretty flowers.

b It says what the polyps that make coral reefs are like.

c It says that polyps have stinging needles.

Checking and Clearing up:

Are there any words I don't understand? What can I do about it?

carnivorous tentacles prey victim poisoned _____

Do all the ideas make sense? Is there anything that isn't explained?

- How does a tentacle know when to fire the stinging needles?
- What sort of small animals do polyps eat?
- Anything else? _____

Putting Ideas Together: (Paragraph 5)

What is the main idea in this paragraph?

a Jellyfish can give us a nasty sting.

b Sea anemones have their tentacles pointing up and jellyfish trail their tentacles in the water below them.

c Sea anemones and jellyfish are related to polyps and all have stinging tentacles.

How does this paragraph fit in with what I've already read?

a It says what the relatives of the polyps are like.

b It says that polyps are smaller than sea anemones and jellyfish.

c It says that sea anemones and jellyfish sting.

Checking and Clearing up:

Are there any words I don't understand? What can I do about it?

relatives trail anemones members nasty _____

Do all the ideas make sense? Is there anything that isn't explained?

- What do sea anemones look like?
- Why do jellyfish trail their tentacles below them?
- Anything else? _____

Collecting Ideas: Heading

What do I know about reef animals with shells? What might this section be about?

_____

_____

Putting Ideas Together: (Paragraph 6)

What is the main idea in this paragraph?

a Giant Tritons are the largest sea snails and have beautiful shells.

b Giant Tritons are killed for their beautiful shells.

c Giant Tritons can be half a metre long and they eat starfish.

How does this paragraph fit in with what I've already read?

a It says that sea snails are like garden snails.

b It says what one reef animal with a shell is like.

c It says what Tritons eat.

Checking and Clearing up:

Are there any words I don't understand? What can I do about it?

markings danger beauty _____

Do all the ideas make sense? Is there anything that isn't explained?

- Who kills Tritons?
- How can sea snails live in water when garden snails can't?
- Anything else? _____

Putting Ideas Together: (Paragraph 7)

What is the main idea in this paragraph?

a Bivalve shells close at low tide and when there is danger.

b Bivalve shells let water in so that they can filter out bits of food.

c Bivalve shells have two parts that open and close to control the amount of water in the shell.

How does this paragraph fit in with what I've already read?

a It says what bivalves eat.

b It says that the bivalve shell is like a valve on a tyre.

c It says what another kind of reef animal with a shell is like.

Checking and Clearing up:

Are there any words I don't understand? What can I do about it?

valve filter particle protect tide _____

Do all the ideas make sense? Is there anything that isn't explained?

- What sort of danger is there for bivalves?
- What are the particles in the water?
- Why do bivalves have to stop themselves drying out?
- What sort of animals live in bivalve shells?
- Anything else? _____

Putting Ideas Together: (Paragraph 8)

What is the main idea in this paragraph?

a People couldn't get caught in Giant Clams like they do in the movies.

b The largest bivalve is the Giant Clam, which can be very old and big.

c Giant Clams can be 100 years old, one metre long and weigh 200 kilograms.

How does this paragraph fit in with what I've already read?

a It says what colour Giant Clams are and what sort of food they eat.

b It says that Giant Clams aren't dangerous.

c It says what one kind of bivalve is like.

Checking and Clearing up:

Are there any words that I don't understand? What can I do about it?

unlikely flesh _____

Do all the ideas make sense? Is there anything that isn't explained?

- Why wouldn't a person get caught in a Giant Clam?
- How do people know how old clams are?
- Anything else? _____

Putting Ideas Together: (Paragraph 9)

What is the main idea in this paragraph?

a Crustaceans have a hard skeleton around them instead of a sea shell.

b Some animals are called crustaceans.

c Lobsters, crabs and prawns have antennae and nippers.

How does this paragraph fit in with what I've already read?

a It says that lobsters, crabs and prawns have claws to hold their food.

b It says what animals with a different kind of shell are like.

c It says that lobsters, crabs and prawns are crustaceans.

Checking and Clearing up:

Are there any words I don't understand? What can I do about it?

skeleton antennae nippers thrive crustaceans _____

Do all the ideas make sense? Is there anything that isn't explained?

- In what way is the hard skeleton of crustaceans different to other shells?
- Why do crustaceans thrive on the reef?
- Anything else? _____

Collecting Ideas: Heading

What do I know about fish on the reef? What might this section be about?

_____

_____

Putting Ideas Together: (Paragraph 10)

What is the main idea in this paragraph?

a There are large sharks and stingrays on the reef.

b Many different kinds of fish can live on the reef.

c A lot of fish on the reef are brightly coloured.

How does this paragraph fit in with what I've already read?

a It says that there a lot of fish with different sizes and habits.

b It says that reef fish are bright.

c It says there is a lot of food and shelter on the reef.

Checking and Clearing up:

Are there any words I don't understand? What can I do about it?

type huge supplies shelter range habits _____

Do all the ideas make sense? Is there anything that isn't explained?

- What are the feeding habits of reef fish?
- Why is there a lot of food and shelter on the reef?
- Anything else? _____

Putting Ideas Together: (Paragraph 11)

What is the main idea in this paragraph?

a Stonefish have poisonous spines which can kill people.

b Stonefish look like stones.

c The perfect camouflage of stonefish helps them catch fish.

How does this paragraph fit in with what I've already read?

a It says that stonefish are dangerous.

b It describes what one unusual fish looks like.

c It describes how the stonefish sucks fish into its mouth.

Checking and Clearing up:

Are there any words I don't understand? What can I do about it?

camouflage spine poisonous _____

Do all the ideas make sense? Is there anything that isn't explained?

- Why do stonefish have perfect camouflage?
- How do people get poisoned by its spines?
- Why do fish swim past the stonefish when they should know it will eat them?
- Why do stonefish sit still and bury themselves in sand?
- Anything else? _____

Putting Ideas Together: (Paragraph 12)

What is the main idea in this paragraph?

a Clownfish are safe in sea anemones because they are immune to the poison.

b Clownfish are brightly coloured, small fish.

c Clownfish keep sea anemones clean by eating the algae that grows on them.

How does this paragraph fit in with what I've already read?

a It says that clownfish clean the sea anemone.

b It says that most fish are not immune to sea anemone poison.

c It says where one kind of fish lives.

Checking and Clearing up:

Are there any words I don't understand? What can I do about it?

manage immune in return algae _____

Do all the ideas make sense? Is there anything that isn't explained?

- Why are the clownfish immune to sea anemone poison?
- How does the clownfish get its name?
- Anything else? _____

Putting Ideas Together: (Paragraph 13)

What is the main idea in this paragraph?

a Cleanerfish pick small bits off other fish to clean them.

b Cleanerfish set up "cleaning stations" on flat pieces of rock.

c No-one eats cleanerfish because they are useful.

How does this paragraph fit in with what I've alrady read?

a It describes a "cleaning station".

b It describes the unusual habits of one kind of fish.

c It says that cleanerfish swim into the mouths of big fish.

Checking and Clearing up:

Are there any words I don't understand? What can I do about it?

dart gill parasite customer _____

Do all the ideas make sense? Is there anything that isn't explained?

- What does the cleanerfish get out of cleaning other fish?
- How do other fish know not to eat cleanerfish?
- How do other fish find "cleaning stations"?
- Why do fish want to be cleaned?
- Anything else? _____

Putting Ideas Together: (Paragraph 14)

What is the main idea in this paragraph?

a Reef sharks are harmless and fairly small, being less than two metres long.

b Tiger and hammerhead sharks stay further out to sea.

c The sharks on the reef are the smaller and less dangerous types.

How does this paragraph fit into what I've already read?

a It describes different kinds of sharks.

b It describes some of the largest fish on the reef.

c It describes the sharks that live out to sea.

Checking and Clearing up:

Are there any words I don't understand? What can I do about it?

harmless dangerous _____

Do all the ideas make sense? Is there anything that isn't explained?

- Why are the sharks on the reef harmless?
- What sort of food do sharks look for?
- Anything else? _____

Natural Disasters

Collecting Ideas: Main Heading

What do I know about natural disasters? What might this text be about?

_____

_____

Collecting Ideas: Heading

What do I know about how the Earth was formed? What might this section be about?

_____

_____

Putting Ideas Together: (Paragraph 1)

What is the main idea in this paragraph?

a When the Earth was very hot, all the rocks were molten.

b The outside of the Earth is solid, but the inside is hot and liquid.

c The inside of the Earth is hot and runny, like toffee.

How does this paragraph fit in with what I've already read? (the heading)

a It says how the crust of the Earth was formed.

b It says what the rocks were like when the Earth was hot.

c It says what the inside of the Earth is like.

Checking and Clearing up:

Are there any words I don't understand? What can I do about it?

formed molten surface liquid _____

Do all the ideas make sense? Is there anything that isn't explained?

- Why are the rocks in the centre of the Earth still liquid?
- How do people know what the centre of the Earth is like?
- Why did the surface become solid when it cooled down?
- Anything else? _____

Putting Ideas Together: (Paragraph 2)

What is the main idea in this paragraph?

a The pieces of the Earth's crust are called plates.

b The plates carry the land and the sea.

c The Earth's crust is made up of a few big plates that can move about.

How does this paragraph fit in with what I've already read?

a It says what the pieces of the Earth's crust are called.

b It says that the pieces of the Earth's crust are like rafts.

c It says what the crust of the Earth is like.

Checking and Clearing up:

Are there any words I don't understand? What can I do about it?

raft crust plate _____

Do all the ideas make sense? Is there anything that isn't explained?

- Why do things that float move about?
- How can the plates carry something as big as the land and the sea?
- If the Earth's crust is made up of pieces, why don't they fall apart?
- Anything else? _____

Collecting Ideas: Heading

What do I know about earthquakes? What might this section be about?

_____

_____

Putting Ideas Together: (Paragraph 3)

What is the main idea in this paragraph?

a In an earthquake the ground shakes and there may be large cracks in the ground.

b In earthquakes buildings collapse and people may be killed.

c When two plates grind against each other, the ground shakes and can cause a lot of damage.

How does this paragraph fit in with what I've already read?

a It says that the ground shakes in an earthquake.

b It says why earthquakes happen.

c It says that earthquakes can kill people.

Checking and Clearing up:

Are there any words I don't understand? What can I do about it?

collapse shift pressure _____

Do all the ideas make sense? Is there anything that isn't explained?

- How do the plates put pressure on rocks underground?
- How do rocks make the ground shake when they split and shift?
- How do earthquakes kill people?
- Anything else? _____

Putting Ideas Together: (Paragraph 4)

What is the main idea in this paragraph?

a An earthquake is strongest at the epicentre and gets weaker further away.

b Shock waves from an earthquake spread out like ripples in water.

c Where the rocks start to move in an earthquake is called the epicentre.

How does this paragraph fit in with what I've already read?

a It says that shock waves are like ripples in water.

b It says what an epicentre is.

c It says where the force of an earthquake is greatest.

Checking and Clearing up:

Are there any words I don't understand? What can I do about it?

shock waves ripples weaken tremor epicentre force

_____

Do all the ideas make sense? Is there anything that isn't explained?

- Why do shock waves get weaker as they spread out?
- Why is the force of an earthquake greatest at the epicentre?
- What are shock waves?
- Anything else? _____

Putting Ideas Together: (Paragraph 5)

What is the main idea in this paragraph?

a Australia is unlikely to have big earthquakes.

b Japan and parts of the United States are on the edge of plates.

c Countries that have big earthquakes are on the edge of plates.

How does this paragraph fit in with what I've already read?

a It says that Australia doesn't have big earthquakes.

b It says where earthquakes are most likely to happen.

c It says that the big city of San Francisco was almost destroyed by an earthquake.

Checking and Clearing up:

Are there any words I don't understand? What can I do about it?

chance danger destroyed _____

Do all the ideas make sense? Is there anything that isn't explained?

- Where is San Francisco?
- Why is Australia safe from big earthquakes?
- Anything else? _____

Putting Ideas Together: (Paragraph 6)

What is the main idea in this paragraph?

a Over time people have used different methods to predict earthquakes.

b The Chinese used balls between dragon's teeth to predict earthquakes.

c Scientists have special instruments that pick up even slight tremors.

How does this paragraph fit in with what I've already read?

a It says how experts predict earthquakes.

b It says what method the Chinese used to predict earthquakes.

c It says how people know when earthquakes are going to happen.

Checking and Clearing up:

Are there any words I don't understand? What can I do about it?

experts instruments method notice dragon _____

Do all the ideas make sense? Is there anything that isn't explained?

- How do slight tremors warn people that there might be an earthquake?
- Why do people want to get an early warning of earthquakes?
- Anything else? _____

Collecting Ideas: Heading

What do I know about volcanoes? What might this section be about?

_____

_____

Putting Ideas Together: (Paragraph 7)

What is the main idea in this paragraph?

a Volcanoes don't erupt explosively when hot gas and rock come freely through the crust.

b Around volcanoes you may find pools of boiling mud and foul smelling gas.

c Hot rock called lava pours through the Earth's crust

How does this paragraph fit in with what I've already read?

a It says what geysers are like.

b It says why volcanoes don't always erupt explosively.

c It says where you get pools of boiling mud and foul smelling gas.

Checking and Clearing up:

Are there any words I don't understand? What can I do about it?

mixture eruption geysers foul explosively _____

Do all the ideas make sense? Is there anything that isn't explained?

- How can hot gas and rock come through the Earth's crust?
- Why does mud boil?
- Why do geysers shoot up?
- Anything else? _____

Putting Ideas Together: (Paragraph 8)

What is the main idea in this paragraph?

a When a volcano erupts, smoke and lava shoot into the air.

b Volcano cones can explode into small pieces when there is an eruption.

c Volcanoes erupt when they are blocked and pressure builds up inside.

How does this paragraph fit in with what I've already read?

a It says why volcanoes erupt explosively.

b It says what volcanoes look like.

c It says what volcanoes can do.

Checking and Clearing up:

Are there any words that I don't understand? What can I do about it?

blocked cone build up explode _____

Do all the ideas make sense? Is there anything that isn't explained?

- How can a volcano erupt when it is blocked?
- Why do smoke and lava shoot into the air?
- In what way is a volcano like a bottle of fizzy drink?
- Anything else? _____

Putting Ideas Together: (Paragraph 9)

What is the main idea in this paragraph?

a Volcanoes are found where liquid rock comes up between two plates.

b Places like New Zealand and the United States have active volcanoes.

c Australia doesn't have active volcanoes.

How does this paragraph fit in with what I've already read?

a It says that New Zealand and the United States have active volcanoes.

b It says that Australia doesn't have active volcanoes.

c It says where you find volcanoes.

Checking and Clearing up:

Are there any words I don't understand? What can I do about it?

_____

Do all the ideas make sense? Is there anything that isn't explained?

- What is an active volcano?
- Why do New Zealand and the United States have active volcanoes?
- Anything else? _____

Putting Ideas Together: (Paragraph 10)

What is the main idea in this paragraph?

a A mighty volcanic eruption exploded the island of Krakatoa.

b Ash and dust from the volcano blocked out the sun's light for three days.

c The explosion was one hundred years ago and the volcano had been dormant for two hundred years.

How does this paragraph fit in with what I've already read?

a It says how long ago the explosion on Krakatoa was.

b It says how a volcano on Krakatoa exploded.

c It says why you couldn't see the sun's light.

Checking and Clearing up:

Are there any words I don't understand? What can I do about it?

dormant rumble mighty explosion _____

Do all the ideas make sense? Is there anything that isn't explained?

- Why did the volcano start to rumble after two hundred years?
- Where is Indonesia?
- How could the island explode into small pieces?
- Anything else? _____

Putting Ideas Together: (Paragraph 11)

What is the main idea in this paragraph?

a The waves from Krakatoa spread out to many other islands.

b A wall of water killed forty thousand people.

c The explosion caused huge waves that killed many people.

How does this paragraph fit in with what I've already read?

a It says that houses all around the coast were swamped.

b It says that the explosion caused huge waves.

c It says how many people were killed.

Checking and Clearing up:

Are there any words I don't understand? What can I do about it?

swamped caused _____

Do all the ideas make sense? Is there anything that isn't explained?

- How does an explosion cause waves?
- Why did the waves kill more people than the explosion?
- Why were so many houses swamped?
- Anything else? _____

Putting Ideas Together: (Paragraph 12)

What is the main idea in this paragraph?

a The people around Vesuvius thought that the volcano was extinct.

b Vesuvius erupted violently after being dormant for many years.

c People built fine towns and grew crops around Vesuvius.

How does this paragraph fit in with what I've already read?

a It says how another famous volcano erupted.

b It says what the people thought about the volcano.

c It says what the people did around the volcano.

Checking and Clearing up:

Are there any words I don't understand? What can I do about it?

took place famous extinct _____

Do all the ideas make sense? Is there anything that isn't explained?

- Why did people think the volcano was extinct when it was only dormant?
- Why did Vesuvius take sixteen years to erupt?
- Why did the eruption have such great force?
- Anything else? _____

Putting Ideas Together: (Paragraph 13)

What is the main idea in this paragraph?

a Many people choked under the ash from Vesuvius.

b Ash mixed with rain made a mixture like cement.

c Ash and lava from Vesuvius covered the town of Pompeii.

How does this paragraph fit in with what I've already read?

a It says that Pompeii was buried for hundreds of years.

b It says what damage the volcano caused in Pompeii.

c It says that the ash choked many people.

Checking and Clearing up:

Are there any words I don't understand? What can I do about it?

choked cement _____

Do all the ideas make sense? Is there anything that isn't explained?

- Could people run into their houses to escape the volcano?
- Did the rain have anything to do with the volcano erupting?
- Anything else? _____

Putting Ideas Together: (Paragraph 14)

What is the main idea in this paragraph?

a Pompeii has been uncovered and you can see the buildings and models of the victims.

b In Pompeii you can see what Roman buildings looked like.

c Scientists have made models of the people and animals who died in Pompeii.

How does this paragraph fit into what I've already read?

a It says what Pompeii looks like now.

b It says that Pompeii was well preserved under the ash.

c It says how scientists made models of the victims.

Checking and Clearing up:

Are there any words I don't understand? What can I do about it?

recent preserved outline plaster hollow victim _____

Do all the ideas make sense? Is there anything that isn't explained?

- Why were the buildings well preserved?
- How did the ash leave an outline of the bodies?
- How can you make a model from these outlines?
- Anything else? _____

Animal Defence Systems

Collecting Ideas:

What do I know about animal defence systems? What might this text be about?

_____

_____

Putting Ideas Together: (Paragraph 1)

What is the main idea in this paragraph?

_____

How does this paragraph fit in with what I've already read? (the heading)

_____

Checking and Clearing Up:

Are there any words I don't understand? What can I do about it?

_____

Do all the ideas make sense? Is there anything that isn't explained?

_____

_____

Collecting Ideas: Heading

What do I know about animals hiding from a hunter? What might this section be about?

_____

_____

Putting Ideas Together: (Paragraph 2)

What is the main idea in this paragraph?

_____

How does this paragraph fit in with what I've already read?

_____

Checking and Clearing Up:

Are there any words I don't understand? What can I do about it?

_____

Do all the ideas make sense? Is there anything that isn't explained?

_____

_____

Putting Ideas Together: (Paragraph 3)

What is the main idea in this paragraph?

_____

How does this paragraph fit in with what I've already read?

_____

Checking and Clearing Up:

Are there any words I don't understand? What can I do about it?

_____

Do all the ideas make sense? Is there anything that isn't explained?

_____

_____

Putting Ideas Together: (Paragraph 4)

What is the main idea in this paragraph?

_____

How does this paragraph fit in with what I've already read?

_____

Checking and Clearing Up:

Are there any words I don't understand? What can I do about it?

_____

Do all the ideas make sense? Is there anything that isn't explained?

_____

_____

Putting Ideas Together: (Paragraph 5)

What is the main idea in this paragraph?

_____

How does this paragraph fit in with what I've already read?

Checking and Clearing Up:

Are there any words I don't understand? What can I do about it?

Do all the ideas make sense? Is there anything that isn't explained?

Putting Ideas Together: (Paragraph 6)

What is the main idea in this paragraph?

How does this paragraph fit in with what I've already read?

Checking and Clearing Up:

Are there any words I don't understand? What can I do about it?

Do all the ideas make sense? Is there anything that isn't explained?

Putting Ideas Together: (Paragraph 7)

What is the main idea in this paragraph?

How does this paragraph fit in with what I've already read?

Checking and Clearing Up:

Are there any words I don't understand? What can I do about it?

Do all the ideas make sense? Is there anything that isn't explained?

Putting Ideas Together: (Paragraph 8)

What is the main idea in this paragraph?

_____

How does this paragraph fit in with what I've already read?

_____

Checking and Clearing Up:

Are there any words I don't understand? What can I do about it?

_____

Do all the ideas make sense? Is there anything that isn't explained?

_____

_____

Putting Ideas Together: (Paragraph 9)

What is the main idea in this paragraph?

_____

How does this paragraph fit in with what I've already read?

_____

Checking and Clearing Up:

Are there any words I don't understand? What can I do about it?

_____

Do all the ideas make sense? Is there anything that isn't explained?

_____

_____

Collecting Ideas: Heading

What do I know about animals who bluff a hunter? What might this section be about?

_____

_____

Putting Ideas Together: (Paragraph 10)

What is the main idea in this paragraph?

_____

How does this paragraph fit in with what I've already read?

_____

Checking and Clearing Up:

Are there any words I don't understand? What can I do about it?

_____

Do all the ideas make sense? Is there anything that isn't explained?

_____

_____

Putting Ideas Together: (Paragraph 11)

What is the main idea in this paragraph?

_____

How does this paragraph fit in with what I've already read?

_____

Checking and Clearing Up:

Are there any words I don't understand? What can I do about it?

_____

Do all the ideas make sense? Is there anything that isn't explained?

_____

_____

Putting Ideas Together: (Paragraph 12)

What is the main idea in this paragraph?

_____

How does this paragraph fit in with what I've already read?

_____

Checking and Clearing Up:

Are there any words I don't understand? What can I do about it?

Do all the ideas make sense? Is there anything that isn't explained?

Putting Ideas Together: (Paragraph 13)

What is the main idea in this paragraph?

How does this paragraph fit in with what I've already read?

Checking and Clearing Up:

Are there any words I don't understand? What can I do about it?

Do all the ideas make sense? Is there anything that isn't explained?

Putting Ideas Together: (Paragraph 14)

What is the main idea in this paragraph?

How does this paragraph fit in with what I've already read?

Checking and Clearing Up:

Are there any words I don't understand? What can I do about it?

Do all the ideas make sense? Is there anything that isn't explained?

Putting Ideas Together: (Paragraph 15)

What is the main idea in this paragraph?

_____

How does this paragraph fit in with what I've already read?

_____

Checking and Clearing Up:

Are there any words I don't understand? What can I do about it?

_____

Do all the ideas make sense? Is there anything that isn't explained?

_____

_____

Putting Ideas Together: (Paragraph 16)

What is the main idea in this paragraph?

_____

How does this paragraph fit in with what I've already read?

_____

Checking and Clearing Up:

Are there any words I don't understand? What can I do about it?

_____

Do all the ideas make sense? Is there anything that isn't explained?

_____

_____

Putting Ideas Together: (Paragraph 17)

What is the main idea in this paragraph?

_____

How does this paragraph fit in with what I've already read?

_____

Checking and Clearing Up:

Are there any words I don't understand? What can I do about it?

_____

Do all the ideas make sense? Is there anything that isn't explained?

_____

_____

Collecting Ideas: Heading

What do I know about animals who attack a hunter? What might this section be about?

_____

_____

Putting Ideas Together: (Paragraph 18)

What is the main idea in this paragraph?

_____

How does this paragraph fit in with what I've already read?

_____

Checking and Clearing Up:

Are there any words I don't understand? What can I do about it?

_____

Do all the ideas make sense? Is there anything that isn't explained?

_____

_____

Putting Ideas Together: (Paragraph 19)

What is the main idea in this paragraph?

_____

How does this paragraph fit in with what I've already read?

_____

Checking and Clearing Up:

Are there any words I don't understand? What can I do about it?

Do all the ideas make sense? Is there anything that isn't explained?

Putting Ideas Together: (Paragraph 20)

What is the main idea in this paragraph?

How does this paragraph fit in with what I've already read?

Checking and Clearing Up:

Are there any words I don't understand? What can I do about it?

Do all the ideas make sense? Is there anything that isn't explained?

The Age of Dinosaurs

Main Heading

Heading

Paragraph 1

Paragraph 2

Paragraph 3

Paragraph 4

Paragraph 5

Heading

Paragraph 6

Paragraph 7

Paragraph 8

Paragraph 9

Paragraph 10

Paragraph 11

Heading

Paragraph 12

Paragraph 13

Paragraph 14

Appendix C:

Comprehension

Practice Passages

Living in the Cold

Winter is a hard time for most animals, but in some places it can be deadly. The land is covered with ice and snow, so a lot of animals find it hard to keep warm and find enough food.

Animals that live in the Arctic Ocean can still find food among the icebergs. Seals catch fish, and killer whales catch seals and fish. To keep warm, all these animals have a thick layer of fat, called blubber, around their bodies. Seals also have a thick, oily fur which keeps the cold water off their bodies.

The caribou is a type of reindeer that lives in the far north of America. They solve the problem of finding enough food by migrating. When winter comes, huge herds of caribou move south to find shelter in forests. Here they can eat bark, shoots and moss off the trees till the spring sun melts the snow and lets the grass grow again. A caribou that is too weak to move south may well die.

Big bears like the grizzly solve the problem of winter feeding by hibernating. In summer and autumn bears spend most of their time eating. They eat berries, roots, honey, fish, and whatever else they can find. By the time winter comes they are fat enough to curl up in their dens and doze until spring comes.

Some animals, though, can stay active in the freezing winter. Voles are small animals, a bit like a guinea pig. They make burrows under the snow to protect them from the cold winds. Voles cut grass in summer and make "haystacks" in sheltered places. They can use this food when the ground is covered with snow. But they can also tunnel through the snow to find buried plants.

The fox is another animal that can face the winter cold. It has a thick fur to protect it. It even has hairy feet, a bit like warm slippers, to stop its toes freezing. It can curl up with its bushy tail folded over it and survive a snow storm. The fox has to eat anything it can find. Its main food is the vole; it can smell voles and dig through the snow to find them.

Living in the Cold

Name: _____

1 Which animals can find food among the icebergs of the Arctic Ocean?

_____

2 What do all the animals in the Arctic Ocean have to keep warm?

_____

3 What exactly do caribou find when they migrate south to shelter in forests?

_____

4 Why do bears like the grizzly spend most of their time eating in summer and autumn?

_____

5 What sorts of food do bears like the grizzly eat?

_____

6 What kind of animal is a vole?

_____

7 Why do voles cut grass to make "haystacks"?

_____

8 How can foxes find voles?

_____

9 Why do you think the little voles can survive in the snow, but the caribou have to move south?

_____

10 Why do you think animals in the Arctic Ocean can find food all year round, but some animals on land can't?

_____

Beavers - the Dam Builders

Beavers are the master builders of the animal world. They cut down small trees and use the pieces to make big piles of branches in a river. They stick mud around these pieces and when the mud goes hard, it is like cement. Inside the pile they make a home for themselves, called a lodge. The only way to get into the lodge is through an underwater tunnel.

The beaver is well suited for its lifestyle. It has large, strong front teeth, which let it chew through the trunks of young trees and cut them into pieces for building.

The beaver is also a very good swimmer. It has a flat, scaly tail that it uses to steer. It can push the tail down through the water to get a rapid burst of speed. It can slap the water with its tail to warn other beavers of danger.

The beaver has strong front claws for digging in the mud. Its back feet are webbed to help it swim. On land the beaver is slow and clumsy. When it is in danger, it dives into the water as quickly as possible and tries to get to its lodge. Hunting animals can't swim as well as the beaver, so they can't get through the tunnel into the lodge.

Beavers like to eat wood and bark. In cold places they collect young branches and leave them in the water. When they can't get out onto the land because the surface of the water is frozen, they can pull these branches into their home for food from under the water.

The beaver's lodge is also a dam wall that stops the river flowing and makes a lake. They cleverly control the amount of water in this lake. If the water goes down, the beavers are in danger because the tunnel into the lodge will be above the water and let other animals in. The beavers build the lodge higher. This stops the river flowing past and keeps more water around the lodge. If the water rises, the lodge could be flooded, so they construct another wall further up the river to stop too much water flowing down into their lake.

Beavers - the Dam Builders

Name: _____

1 How do beavers stick together pieces of wood?

_____

2 What is the only way to get into a beaver's lodge?

_____

3 Why do beavers need strong front teeth?

_____

4 What do you think makes the beaver slow and clumsy on land, but fast in the water?

_____

5 Why does a beaver try to get to its lodge as quickly as possible to escape danger?

_____

6 What do beavers eat?

_____

7 In *cold* places, why do beavers need to collect young branches and leave them in the water?

_____

8 What is one reason why beavers have to control the amount of water in their lake?

_____

9 What do beavers do when the water in their lake goes down?

_____

10 What do you think could be one of the hunting animals beavers have to escape from?

_____

How Birds Fly

Birds are well designed for flying. Their bones are hollow to make them light, but they have a fine criss-cross structure on the inside to give them extra strength. The muscles on the chest are large and powerful to flap the wings.

Birds have two kinds of feathers. Some are soft and fluffy to keep birds warm. The feathers on the wings are bigger and stronger to push birds through the air. Feathers are made of the same material as our hair and nails.

When a bird's wings flap down, they push air down and the bird rises. The wings can also push air backwards to move the bird forward. This is rather like what we do when we swim; we push water backwards to make ourselves move forward.

The wings wouldn't work very well if the feathers had bits missing or were covered in dirt, so birds have to look after their feathers. Many birds take baths in water and then shake out their feathers to dry them. They use their beaks like a comb to clean out dirt from their feathers. They spread wax from a gland near the tail over the feathers to make them waterproof.

Not all birds fly by flapping their wings in the usual way. Some birds have very long wings which they keep stretched out. They use air currents to glide and use almost no energy at all. The albatross is a large bird that lives far out to sea. It uses its long wings to glide over the ocean. Vultures glide in circles looking for dead animals.

Some birds can hover, which means that they stay in the same place. The kestral is a hunting bird. It faces into the wind and flaps its wings only just enough to match the speed of the wind. So, it stays in the same position and can watch for small animals on the ground.

The hummingbird is the only bird which can turn its wings like we can turn our wrist. It beats its wings very quickly and turns them just the right way to stay in the same position so that it can use its long beak to eat nectar from flowers.

How Birds fly

Name: _____

1 What makes birds' bones light?

_____

2 What are the two kinds of feathers for?

_____

3 What are feathers made of?

_____

4 In what way is flying like swimming?

_____

5 Why do birds have to look after their feathers?

_____

6 Why do you think it is important that feathers are waterproof?

_____

7 What do birds do with their wings when they use air currents to glide?

_____

8 In what way do you think it helps birds to glide rather than flapping their wings in the usual way?

_____

9 What does the kestral do when it *hovers* to stay in the same place?

_____

10 Why does the hummingbird want to stay in the same position?

_____

How Animals Communicate

Animals can't talk or write like we do, but they use sight, sound, smell and touch to send signals to others of their kind.

One important signal is to warn others of danger. In small fish called minnows, an injured fish gives off a chemical signal that others can smell. The whole school of minnows takes off in fright.

Deer often use sight signals as warnings. In herds of white-tail deer the leader flashes the white underside of her tail to warn others. When an African springbok senses danger, it leaps high into the air with four stiff legs. This jump, called pronking, can be seen by springboks all around and tells them to beware.

Many birds and mammals use special sounds to signal danger. The rabbit uses its strong back legs to thump the ground. Beavers slap the water with their flat tail.

Animals also have to tell others that they are ready to mate. Some male birds offer the female a "present". Some penguins put down a stone in front of the female. Birds like the tern walk up and down the beach with a fish in their mouth, and the roadrunner goes up to the female with a lizard. If the females pick up these "presents", it means they are ready to mate.

If a male spider goes straight up to a female, he may get eaten. He has to tap her web and wait for her to tap back. If she shakes the web, she doesn't want to mate and he had better run away!

Many animals have signals to help parents look after their young. The female crocodile buries her eggs in the ground. She waits near the nest until she hears the peeping sound which tells her that the young are ready to hatch. She gently digs them out and helps them get to the water.

Animals that live in groups have ways of telling others who is "top dog" and must be obeyed. The leader of a baboon troup walks with his tail in the air. He is allowed to stare at others, but no-one is allowed to stare at him. If they do, this is a threat and may lead to a fight.

How Animals Communicate

Name: _____

1 Why do you think minnows take off in fright when they smell an injured fish?

_____

2 What is unusual about the way the African springbok "pronks" to warn others of danger?

_____

3 What special sound does the rabbit use to signal danger?

_____

4 What do birds like the tern do on the beach when they are looking for a mate?

_____

5 How do some kinds of female birds show they are ready to mate?

_____

6 What would happen if a male spider didn't tap a female's web and wait for her to tap back?

_____

7 How does a female spider show that she doesn't want to mate?

_____

8 What does the female crocodile do when she hears peeping sounds coming from her nest?

_____

9 How does the leader of a baboon troop walk?

_____

10 What do you think might happen if animals didn't have signals to tell others who is "top dog" and must be obeyed?

_____

Insect Life

There are many thousands of different kinds of insects, making insects the largest group of animals in the world. They can live in all parts of the world, from the hottest to the coldest places. Wherever you are, you are likely to find insects.

Insects come in all sorts of shapes and sizes. The smallest ones are only the size of a pinhead, but the giants are quite large. The Goliath beetle is ten centimetres long and weighs one hundred grams. The Atlas moth has a wingspan of thirty centimetres.

Insects have a range of mouth parts to suit the sort of food they eat. Some have very strong jaws for crushing food. The dragonfly and the preying mantis catch and chew up other insects. Caterpillars use their strong jaws to chew up leaves and even wood.

The mouth part of some insects is a sharp tube, which they use to pierce and suck. Mosquitoes and fleas pierce the skin of animals and suck a meal of blood. Some bugs pierce the skin of leaves to suck out the sap inside.

Those insects that feed on nectar don't need strong mouth parts, but have a very long tongue instead. In butterflies this long tongue is rolled up into a ball. They uncurl the tongue into a flower to reach the nectar at the bottom and then lap it up. Honey bees also eat their food in this way.

One thing all insects have in common is the way they are born. They all start out as eggs and hatch into grubs. These grubs grow very fast, but their skin is stiff and can't stretch. They have to grow a new skin many times before they are adults. When the new skin is ready, the old one splits and falls off. This is called moulting.

Once the grub stops growing, it spins a cocoon. Because we can't see inside the cocoon, we might think nothing is going on, but in fact the insect is growing a new body. The grub that went into the cocoon comes out as a moth, a beetle, a fly, or any of the other many shapes an insect can have.

Insect Life

Name: _____

1 Why do insects have a range of mouth parts?

_____

2 How do the dragonfly and the preying mantis manage to chew up other insects?

_____

3 Why do some insects have a sharp tube as a mouth?

_____

4 What do some bugs do after they pierce the skin of leaves?

_____

5 Why do you think insects that feed on nectar have a long tongue instead of strong mouth parts?

_____

6 How do honey bees eat their food?

_____

7 How are all insects born?

_____

8 Why do grubs have to grow a new skin many times before they are adults?

_____

9 What is going on inside the cocoon?

_____

10 Why do you think that not all insects around plants in the garden are pests?

_____

Bats

Bats are the only mammals that can fly. Some other mammals can glide through the air, but they don't have wings and so can't really fly. Bats have four legs, with the front ones much longer than the back ones. Four long fingers on their front legs are stretched out to support the thin skin covering their wings. The thumb is a hook that they use for climbing and for holding food. Bats use their short back legs for hanging upside down while they sleep.

Most bats eat insects and hunt during the night. Bats aren't blind, but they can't see their prey in the dark. They have to use a sort of radar to find them. They send out pulses of sound too high for us to hear, which bounce off any object in their way and bounce back to the bat as an echo. Bats can hear these echoes with their large ears. Changes in the echo tell them how fast an insect is moving, so that they can quickly fly at it and snap it up.

In bats that send out sound pulses through their noses, the skin around the nose grows into unusual shapes, called leaf-noses. Some bats have a long, thin leaf-nose, like a sword, and are called sword-nosed bats. Some are called horse-shoe bats because of the shape of their leaf-noses. The leaf-noses make the bats look rather strange and ugly.

Some bats kill prey larger than insects. Some fly through forests at night, a bit like owls, to catch mice and frogs. Some bats fly over water and pull their back claws through the water to catch fish.

Another type of bat that feeds off large animals is the vampire. It scrapes a small wound on its victim and then laps up the blood. Legends make the vampire seem quite ferocious, but it is a very small bat. Its wings stretch only six inches, and it can drink only a small quantity of blood.

The largest of the bats, the flying foxes, are quite different. They still feed at night, but they eat fruit rather than other animals. They don't have leaf-noses, but use their large eyes instead to locate food.

Bats

Name: _____

1 What part of their front legs do bats use for climbing and holding food?

_____

2 At what time do bats feed?

_____

3 Why do bats have to use a sort of radar to find prey?

_____

4 What happens to the sound pulses that bats send out when they hit an object?

_____

5 What are leaf-noses for?

_____

6 What is one animal larger than an insect that some bats kill?

_____

7 What does the vampire bat do to feed off large animals?

_____

8 Why is the vampire not really ferocious or dangerous?

_____

9 Why do you think flying foxes only need their big eyes to find food and do not need leaf-noses?

_____

10 Why do you think people use the saying "blind as a bat"?

_____

Animals that Use Tools

Many animals apart from humans use tools to do a job.

One kind of ant uses a "shopping basket" to collect food. When these ants find a tasty meal, such as ripe fruit, they drop leaves on it and go away. They come back later, when the leaves have soaked up the sweet juices, and pull the leaves into their nest. Other ants then feed off the juice in these leaves.

A number of birds use tools to get food. The woodpecker finch climbs along tree trunks looking for insects. When it finds one hiding in a crack, it breaks off a twig with its beak and jiggles it in and out of the crack. When the insect rushes out, the finch can eat it.

Egyptian vultures like eating ostrich eggs, but these eggs are big and have a hard shell. The vultures pick up a stone in their beak and smash it down on the egg. It may take ten or twelve hits before the egg breaks.

The green heron picks up little pieces of food and drops them into the water. When fish come up to nibble at this bait, the heron snaps them up. It is using food as a tool to catch fish.

Elephants break off branches with their trunk and use them to scratch their back. Without using a tool, the elephant would find it very difficult to scratch bites from the many insects that live on its skin.

Sea otters eat shell fish, but to get their meal they have to open the hard shell. The otters float on their back and put a stone on their chest. They bang the shell against this stone until the shell opens.

Chimpanzees are among the best tool users. In the wild they use pieces of grass or twigs to poke into termite mounds. Termites hold onto these sticks, and the chimps can pull them out and eat them.

With captive chimps, scientists have hung up bunches of bananas too high for them to reach. They also left a few boxes lying around in the same room. The chimps stacked one box on top of the other, until they could climb up to reach the bananas.

Animals that Use Tools

Name: _____

1 Why do some ants drop leaves on a tasty meal and then go away for a while?

_____

2 How do other ants in the nest then use these leaves?

_____

3 What does the woodpecker finch do when it finds an insect hiding in a crack?

_____

4 What does the Egyptian vulture like to eat?

_____

5 Why does the green heron drop food into the water?

_____

6 What do elephants do to scratch their back?

_____

7 How do sea otters open hard shells?

_____

8 What do termites do when chimps poke twigs into their mounds?

_____

9 How do you think chimps knew how to stack up boxes to reach bananas?

_____

10 Why do you think any animal needs to use tools instead of just parts of its body?

_____

Animals in the Desert

At first glance it looks like no animal could survive in a hot, dry desert. In fact, all sorts of creatures make their home there.

Honeypot ants have a special way of storing food. After a rain when times are good, they overfeed some of the workers with sweet nectar from flowers. The bellies of these workers swell to the size of a grape. These ants are so fat that they can't move and just lie on the floor of the nest like little tubs. When times are bad and there isn't much food, the ants stroke the honeypots with their antennae. The honeypots bring up a drop of nectar for the ants to eat.

It is surprising that quite a lot of frogs can live in the desert, because they need to keep their skin moist. They only spend about forty days a year above the ground. During this time they catch insects. The rest of the time they spend below ground. The waterholding frog drinks a huge amount of water when it rains, then it burrows into the mud. As the mud dries out, the frog has a hard ball of clay around it. It can keep the inside of the ball moist with its stored water, so it has its own "swimming pool" until the next rain.

Many snakes and lizards live in the desert. They don't move about much during the day because they would get too hot on the burning sand. Mostly they stay in their burrows to keep cool, and come out to hunt in the evening. They have thick, leathery skins to stop them losing water. Reptiles have to bury their eggs deep in the ground. Otherwise the eggs would get too hot and dry out, or they could make a meal for a hungry animal.

Small mammals also stay in burrows during the day. Hopping mice dig deep burrows where the air stays cool and moist. They only go out to collect seeds and shoots at night. They have a few other tricks that let them survive without drinking any water. To conserve water they make very little urine. They can lose heat through their large ears, so they don't have to sweat much.

Animals in the Desert

Name: _____

1 How do honeypot ants store food after a rain when times are good?

_____

2 What do the honeypots do when an other ant strokes them with its antennae?

_____

3 Why do you think honeypot ants need to have a way of storing the *particular kind* of food they eat?

_____

4 Why is it surprising that quite a lot of frogs can live in the desert?

_____

5 What does the waterholding frog do to the hard ball of clay around it?

_____

6 Where do lizards and snakes mostly stay during the day?

_____

7 Why do lizards and snakes have thick, leathery skin?

_____

8 What is the air like in the deep burrows of hopping mice?

_____

9 What is one of the tricks that lets hopping mice survive without drinking any water?

_____

10 Why do you think you would see very few animals when you walk through a desert?

_____

Fish

There are about 25,000 different kinds of fish living in oceans, rivers and ponds all over the world. They range from tiny little animals to the giant whale shark, which can be fourteen metres long. Fish existed long before the dinosaurs, so they have been around for a very long time.

Some fish have strange ways of getting a meal. The archer fish spits at insects and catches them when they fall into the water. The angler fish has lines dangling off its head. When a small fish nibbles at a line, the angler fish opens its mouth and sucks in the victim. The electric eel uses an electric charge to stun its prey. The biggest fish, the whale shark, sucks in water and filters out plants and animals too small for us to see.

Instead of having lungs like we do, all fish have gills to breathe. The gills are openings on the side of the body behind the head. Fish suck water into their mouths when the gills are closed. Then they close their mouths and open the gills to push the water out. As the water flows through the gills, fish take air from the water into their bodies. This is why fish open and close their mouths all the time.

Fish have a nose, but they use it only to smell, not to breathe. Fish have a good sense of smell. For example, sharks can smell blood from a long way away. Fish don't have ears, but they can feel even small movements in the water. They use this ability to find out where other fish are, so that they know where to catch a meal.

Most fish lay their eggs and just leave them, but some fish do look after them. The male stickleback builds a nest and lures a female inside to lay her eggs. He then chases her off and guards the nest until the eggs hatch. Some fish carry their eggs in their mouths to protect them. They may even take the baby fish into their mouths if they are in danger. The female seahorse lays her eggs into a pouch on the male. He carries them around until they hatch.

Fish

Name: _____

1 Why can we say that fish have been around for a long time?

_____

2 How do the lines work that the angler fish has dangling off its head?

_____

3 What sort of food does the giant whale shark eat?

_____

4 Where are the gills of a fish?

_____

5 What are fish doing when they open and close their mouths all the time?

_____

6 What is one example of fish having a good sense of smell?

_____

7 What do most fish do with their eggs when they are laid?

_____

8 What does the male stickleback fish do with the eggs once the female has laid them?

_____

9 Why do you think fish choke to death when they are on land?

_____

10 Why do you think someone thrashing about in the water might attract sharks?

_____

Appendix D:

Comprehension, Knowledge and Vocabulary Tests

Why Lions Live in Groups

There are about forty kinds of cats in the world. The lion is the only one that lives and hunts in groups. Each group is called a pride, and the pride works together to protect the cubs and catch prey.

Lion cubs are very small and weak when they are born. They are blind and can't hunt for themselves for about a year. If they were left alone, they could be killed by hyenas. Some of the females stay behind to guard them while the others hunt.

Lions also need to work together to hunt. They have two main problems to deal with. To get enough food to eat, they have to be able to kill large animals. They also can't run as well as most of the animals they chase.

Lions can run at a top speed of about forty miles an hour. They can't keep up this speed for very long, and even at top speed they are slower than their prey. But lions do have big, strong muscles that help them take off very quickly. Their only chance of success is to get very close to the prey and pounce. If they don't get close enough, the hunted animal will outrun them. About half the lions' hunts end this way.

Lions have some tricks to help them hunt. Sometimes a few lions hide in bushes or tall grass, while other lions chase prey towards them. The hidden lions have to lie very still until the prey is close enough. Of course, the wind has to be blowing in the right direction. Another trick is to form a circle around a herd of animals and to approach them from all sides at once.

Being in a group also helps lions kill large prey. One lion pulls the victim to the ground, while another lion kills it. A large buffalo could use its horns and hooves to attack one lion, but there isn't much it can do against a group.

The male lion is supposed to be "King of the Jungle", but he hardly ever hunts for himself. This "king" eats what the females kill, or he steals from hyenas.

Why Lions Live in Groups

Name: _____

1 How long is it before lion cubs can hunt for themselves?

_____

2 Why do the females who stay behind have to guard the cubs?

_____

3 What is the main problem lions have when *chasing* animals?

_____

4 What do lions have to help them take off very quickly?

_____

5 What is the only chance lions have of catching prey?

_____

6 How often does a hunted animal outrun the lions?

_____

7 What happens while a few of the lions hide in tall grass or bushes?

_____

8 Why do you think the wind has to be blowing in the right direction when lions stalk prey?

_____

9 In what way does being in a group help lions *kill* large prey?

_____

10 Why do you think the male lion gets away with stealing from other animals?

_____

Name: _____

KNOWLEDGE

Think carefully about the following questions. Think about *what you know* about this subject and put a circle around the answer you think is **best**. For example:

What sort of an animal is a cat?

a a bird

b an insect

c a mammal

d I don't know

1 In what way are lions different to other cats?

a they don't climb trees

b they can run faster

c they live and hunt together

d I don't know

2 Why do lions have to catch large animals?

a they are big and need a lot of food

b they need energy to run around a lot

c they store the leftovers in a tree for another day

d I don't know

3 Where do lions live?

a in thick jungle

b on grassy plains

c on rocky mountains

d I don't know

4 What happens when lions have caught prey?

a females give bits to their cubs

b they share it around the group

c the biggest lions eat first

d I don't know

5 How often do lions catch the prey they chase?

a less than half the time

b most of the time

c all the time

d I don't know

6 What are lion cubs like when they are born?

a they are curious and follow the adults

b they are helpless and blind

c they are playful and start learning to hunt

d I don't know

7 Which lions do most of the hunting?

a all the adults

b the big males

c the females

d I don't know

8 How fast can lions run?

a not very far, but they can reach their top speed quickly

b very well and they are almost as fast as cheetahs

c not that fast, but they can run a long way

d I don't know

9 Why do some of the females always guard the cubs?

a they might hurt themselves, like fall down a cliff

b they might get lost in the jungle

c hyenas could catch and eat them

d I don't know

10 How does hunting in a group help lions?

a they can circle their prey and get close enough to pounce

b they can chase their prey until it gets too tired to run

c the fastest lions in the group can do all the chasing

d I don't know

WORD MEANINGS:

Read the following sentences carefully. Think about what the word in *slanted* writing means in that sentence. Put a circle around the meaning you think is best. For example:

The boy was *disobedient*. *Disobedient* means:

a collected beads

b wouldn't obey rules

c very well behaved

d I don't know

1 He is *supposed to* be clever. *Supposed to* means:

a it is a fact that he is clever

b some people think that he is clever

c he thinks he is clever

d I don't know

2 They were watching for *prey*. *Prey* means:

a animals to hunt and eat

b dead animals

c animals that might hunt them

d I don't know

3 We saw her *approach*. *Approach* means:

a leave

b suddenly appear

c come closer

d I don't know

4 She earned her *success*. *Success* means:

a getting more pocketmoney

b doing well at something

c getting time off

d I don't know

5 Their *hooves* hit the ground. *Hooves* means:

a paws

b shoes

c hard feet

d I don't know

6 We followed the *herd*. *Herd* means:

a faint sounds

b group of animals

c tracks of animals

d I don't know

7 There was still a *chance* of stopping it. *Chance* means:

a it is possible

b it won't happen

c it is certain

d I don't know

8 He showed us his *muscles*. *Muscles* means:

a kinds of shellfish

b top part of the arm

c what moves parts of the body

d I don't know

9 We have a lot of things to *deal with*. *Deal with* means:

a play with

b work on

c sell

d I don't know

10 They *guard* the sheep. *Guard* means:

a keep safe

b care for

c keep in a pen

d I don't know

Spiders as Hunters

All spiders poison their prey. They use their two fangs to grasp their prey and to inject the poison. Spiders eat mostly insects, but they don't eat by chewing and swallowing the whole insect. They only suck out the inside of its body. Around a spider's home you may find the dried up remains of its victims.

Spiders can spin fine threads, a bit like silk. The silk is very fine, but it is stronger than steel of the same thickness.

Some spiders use their silk webs to trap insects. The spider sits in the middle of its web, or spins a thread to a hiding place so that it can feel any movement in the web. The web is sticky, so that any insect that touches it gets stuck. As the insect struggles to get free, the spider feels the web move and rushes out to poison its prey. Spiders don't stick to their web because they have oil on their feet.

Some spiders don't make webs, but lie in hiding and wait for their prey instead. Trapdoor spiders make a burrow in the ground. They use silk to make a hinge for the lid, which they hold down with their front legs. When they feel an insect walking past, they rush out to grab their victim.

Crab spiders have a different way of hiding. These colourful spiders look almost exactly like part of the flowers on which they hide. They sit very still, waiting for insects to visit the flower. The insects can't see them and come close enough for the spiders to pounce.

Some spiders don't make webs or traps, but just hunt their prey. For example, the wolf spider and the huntsman chase insects. All these spiders must be able to run fast and have very good eyes.

The jumping spider creeps up to its victim and then takes off in a long leap to pounce on it. It is also called the flying spider because it can jump so far. The spitting spider is rather lazy and can't move very fast. It creeps up to its prey and then spits a stream of sticky liquid from its fangs. The insect is glued to the spot.

Spiders as Hunters

Name: _____

1 What part of an insect do spiders eat?

_____

2 How strong is the spider's silk?

_____

3 How does a spider know when an insect is caught in its web?

_____

4 How do you think oil stops spiders sticking to their web?

_____

5 Where do trapdoor spiders hide?

_____

6 Why can't insects see crab spiders?

_____

7 How do the wolf spider and the huntsman catch insects?

_____

8 How does the flying spider get its name?

_____

9 What happens to the prey when the spitting spider spits at it?

_____

10 In what way do you think spiders are useful to us?

_____

Name: _____

KNOWLEDGE

Think carefully about the following questions. Think about *what you know* about this subject and put a circle around the answer you think is best. For example:

What sort of an animal is a cat?

a a bird

b an insect

c a mammal

d I don't know

1 How are spiders different to insects?

a they are always bigger than insects

b they don't lay eggs like insects

c they have more legs than insects

d I don't know

2 Where do spiders live when they're not hunting?

a they always stay on their webs

b they find all sorts of hiding places

c they stay in holes in the ground

d I don't know

3 Do spiders have any enemies?

a yes, birds eat spiders

b yes, snakes eat spiders

c no, because they have poison

d I don't know

4 Why don't spiders get caught in their own webs?

a they have hair on their feet to stop them sticking

b they have oil on their feet to stop them sticking

c they know which parts are sticky and only walk there

d I don't know

5 How do spiders eat their prey?

a they wrap them in silk and swallow them whole

b they chew them up with their strong jaws

c they only suck out the inside

d I don't know

6 How do spiders know when insects are caught in their web?

a they feel the web moving

b they can see them land

c they hear the insects buzzing

d I don't know

7 What are the eyes of a spider like?

a they have three eyes in a row

b they can have up to eight eyes

c they have two big eyes like blowflies

d I don't know

8 What does the spitting spider do when it spits at an insect?

a sticks down the wings so that it can't fly very well

b hits it in the eyes so that it can't see

c sticks it to the ground so that it can't move

d I don't know

9 How do spiders kill their prey?

a the strong jaws snap their necks

b the fangs inject poison into them

c they wrap them up in silk and choke them

d I don't know

10 How do trapdoor spiders know when an insect is near their burrow?

a they peep out from under the lid

b they hear the insects moving about

c they feel movements in the ground

d I don't know

WORD MEANINGS

Read the following sentences carefully. Think about what the word in *slanted* writing means in that sentence. Put a circle around the meaning you think is **best**. For example:

The boy was *disobedient*. *Disobedient* means:

a collected beads

b wouldn't obey rules

c very well behaved

d I don't know

1 They *pounced* on the thief. *Pounced* means:

a pounded with fists

b jumped on and grabbed

c jumped up and down

d I don't know

2 John was the *victim* of the robbery. *Victim* means:

a the one who was robbed

b the one who helped the robber

c the one who did the robbery

d I don't know

3 She *injected* the medecine. *Injected* means:

a refused to take

b swallowed

c squeezed through the skin

d I don't know

4 They found his *remains*. *Remains* means:

a footprints on the ground

b what was left of his body

c things that were left behind

d I don't know

5 They *struggled* through the jungle. *Struggled* means:

a swung easily from tree to tree

b walked very quickly

c moved with a lot of effort

d I don't know

6 We found a *burrow* in the garden. *Burrow* means:

a a small hole for animals to live in

b a tray with a wheel and handles for carrying things

c a tunnel through the ground

d I don't know

7 They *creep* along the branch. *Creep* means:

a run quickly and smoothly

b slide easily and quickly

c move slowly and quietly

d I don't know

8 The heat turned it into a *liquid*. *Liquid* means:

a something that is burned

b something that is watery

c something that is sticky

d I don't know

9 We could see its *fangs*. *Fangs* means:

a sharp, pointed teeth

b flat, grinding teeth

c sharp claws

d I don't know

10 He *grasped* the rope. *Grasped* means:

a touched lightly

b slid it through the hands

c got hold of and held tightly

d I don't know

Carnivorous Plants

Most plants use only two ways to get all the food they need to live and grow. They use their roots to get minerals, like iron, from the soil. They use their leaves to make sugar from water, air and sunlight. But the carnivorous plants are unusual because they also eat small animals, mostly insects.

In most ways, carnivorous plants are the same as ordinary plants. They still need air, water, soil and sunlight, and they still use their leaves and roots to get food. The difference is that they grow in very poor soil, where they can't get all the minerals they need through their roots. They have to find a different way to get all the food they need, so they digest animals to make up for the poor soil.

Carnivorous plants have changed some of their leaves into special traps to catch prey. There are three main ways that these traps work. Some lie open and snap shut as soon as an insect touches them. Some use a sticky glue to catch a meal. Others make funnels and wait for insects to fall in.

The plant with the most famous trap that opens and shuts is called the Venus Flytrap. Its leaves have two parts connected by a spine, a bit like an open book. It has spikes all around the edge of the leaf. When an insect touches the leaf, it snaps shut. The spikes around the edge lock together and the insect is trapped.

One plant that uses glue to catch its prey is the Sundew. It has leaves covered with long hairs, with a drop of very sticky glue on the end of each hair. When an insect lands on one of the hairs, it gets stuck. The leaf curls up around the victim to digest it.

The Pitcher Plant has a little jug, called a pitcher, on the end of its leaves. Each jug has sweet nectar around the edge to attract insects. When an insect climbs over this edge, it falls down the slippery sides of the jug. It falls straight into the strong, smelly juices at the bottom of the jug, where it is digested.

Carnivorous Plants

Name: _____

1 In what way are the carnivorous plants unusual?

_____

2 What do carnivorous plants need, the same as ordinary plants?

_____

3 Why do carnivorous plants have to find different ways to get all the food they need?

_____

4 How many main kinds of insect traps are there? (write a number)

_____

5 What type of trap does the Venus Flytrap have?

_____

6 What sort of leaves does the Sundew have to catch its prey?

_____

7 How does the Pitcher Plant attract insects to its little jugs?

_____

8 What does an insect fall into when it falls down the slippery sides of these jugs?

_____

9 What do you think carnivorous plants digest only small animals?

_____

10 How do you think you could keep a carnivorous plant healthy if it didn't digest small animals?

_____

Name: _____

KNOWLEDGE

Think carefully about the following questions. Think about *what you know* about this subject and put a circle around the answer you think is best. For example:

What sort of an animal is a cat?

a a bird

b an insect

c a mammal

d I don't know

1 Why do carnivorous plants eat insects?

a they haven't got leaves to make their own food

b they need the water out of the insect's body

c they grow on very poor soil and need extra food

d I don't know

2 Why do insects go near carnivorous plants when they can be eaten?

a they need something to rest on when they've been flying around

b they think they are flowers and go looking for food

c they aren't very smart and don't look where they are going

d I don't know

3 How do the traps of the Venus Flytrap work?

a lie open and snap shut around insects

b are little pots that insects fall into

c use glue to stick insects to leaves

d I don't know

4 Why do plants need to grow in *rich* soil?

a so they don't get blown away

b so they get enough minerals through their roots

c so they get enough water through their roots

d I don't know

5 Could carnivorous plants grow in the dark?

a no, because plants need sunlight for their leaves to make food

b yes, because they could still eat insects

c no, because insects would never fall into the traps

d I don't know

6 How could plants digest an insect?

a they have thorny bits to crush them

b they have strong leaves that can crush them

c they make special chemicals to break them up

d I don't know

7 What happens if plants don't have enough minerals?

a their leaves are yellow and small

b their leaves go droopy

c they grow long and spindly

d I don't know

8 How do plants get sugar for food?

a from the ground through their roots

b they make it out of minerals in their leaves

c they make it out of air and water in their leaves

d I don't know

9 What are the leaves of carnivorous plants like?

a they don't have any leaves

b green like other leaves so they can use sunlight

c thick and hard so that they can crush insects

d I don't know

10 What are minerals?

a things like iron

b all the things in plant sap

c different kinds of sugar

d I don't know

WORD MEANINGS

Read the following sentences carefully. Think about what the word in *slanted* writing means in that sentence. Put a circle around the meaning you think is **best**. For example:

The boy was *disobedient*. *Disobedient* means:

a collected beads

b wouldn't obey rules

c very well behaved

d I don't know

1 They went to the *edge*. *Edge* means:

a the end of a road

b the top of something

c where something finishes

d I don't know

2 We wanted to *attract* people. *Attract* means:

a get them to come

b make them look nicer

c keep them away

d I don't know

3 The track was *slippery*. *Slippery* means:

a makes you stumble

b makes you slide

c makes you trip

d I don't know

4 The hose wasn't *connected*. *Connected* means:

a taken apart

b cut into parts

c joined

d I don't know

5 The woman was *famous*. *Famous* means:

a rich and owns expensive things

b well known by a lot of people

c good at acting

d I don't know

6 The oil flowed into the *funnel*. *Funnel* means:

a sloping sides joined to a pipe

b long, thin pipe

c carrying can with a handle

d I don't know

7 He hurt his *spine*. *Spine* means:

a back

b backbone

c neck

d I don't know

8 She bought an *ordinary* jumper. *Ordinary* means:

a on special order

b costly

c plain

d I don't know

9 We have to *digest* our food. *Digest* means:

a break into tiny pieces

b chew carefully

c swallow mouthfuls

d I don't know

10 We could see the *difference*. *Difference* means:

a things being alike

b things being apart

c things not being the same

d I don't know

The Ant Colony

Ants live in nests with an enormous number of other ants, so they have to be very well organised.

Each nest has one queen who lays all the eggs. Most ants in a nest are workers. They are females who can't produce eggs, but who do all the housekeeping jobs. The most important job of the workers is to collect food and to look after the queen and the young ants. Workers have several ways of making sure that as much food as possible gets back to the nest.

When a worker finds food, it makes a scent trail back to the nest with special glands on its belly. It can make the trail more or less strong to tell other ants how much food there is. The other ants can read this message with their antennae. That's why you might see lines of ants in the kitchen when you spill something sweet.

Ants also have two stomachs. One is the real stomach for digesting their own food, the other one is a crop for storing food. Workers fill their crop before they return to the nest and use this stored food to feed the queen and the young ants.

Ants can also store food in the nest. They carry seeds, berries or dead insects into a special store room. If the food is too big, they chew it into smaller pieces and carry them into their pantry.

Ants use their antennae when they meet another ant. They smell its antennae to see if it comes from the same nest or not. If it doesn't, there may be a fight. Ants can bite and squirt acid from their back end. No strange ant is allowed into the nest.

In each nest there are several males and young queens, who all have wings. They don't do anything and are fed by the workers. When the young queens fly away, the males follow them to mate. After mating the males die, but the females lose their wings and each make a small nest to lay their eggs. These eggs hatch into workers, and when they become adults they start building a bigger nest for the queen and her eggs.

The Ant Colony

Name: _____

1 What sort of ants are the workers?

_____

2 Who do the workers look after?

_____

3 With what do workers make a scent trail?

_____

4 With what do ants read a scent trail?

_____

5 Why do ants have two stomachs?

_____

6 What sorts of food do ants store in the nest?

_____

7 How do ants tell whether an other ant comes from the same nest or not?

_____

8 What do ants do if a strange ant comes into the nest?

_____

9 Why do you think the young queens leave the nest before they lay their eggs?

_____

10 Why do you think ants have to store food, but most other insects don't?

_____

Name: _____

KNOWLEDGE

Think carefully about the following questions. Think about *what you know* about this subject and put a circle around the answer you think is **best**. For example:

What sort of an animal is a cat?

a a bird

b an insect

c a mammal

d I don't know

1 How do ants use their antennae?

a to feel where they are going

b to send sound messages

c to smell with

d I don't know

2 Why do ants store food, but grasshoppers don't?

a because ants can't always find food and grasshoppers can

b because ants look after the young ants and grasshoppers don't

c because ants eat more than grasshoppers

d I don't know

3 Why don't queen ants get their own food?

a they are too big to run around

b they are too busy giving orders

c they think queens shouldn't have to work

d I don't know

4 What makes ants different to most other insects?

a they eat different food

b they can bite

c they work together

d I don't know

5 What do ants live on?

a insects that are smaller than they are

b a lot of different things

c only food that is sweet

d I don't know

6 Where do ants lay their eggs?

a in a special part of their nest where they look after them

b on the underside of leaves where birds won't see them

c under rocks or bark so that they won't get too hot

d I don't know

7 Which ants lay eggs?

a all the females

b only the young females

c only the queen

d I don't know

8 How do ants tell each other where food is?

a they have special dances to tell other ants where to go

b they make a smelly trail for other ants to follow

c they send out special sounds from their antennae

d I don't know

9 How can ants tell whether another ant comes from the same nest or not?

a they smell its body with their nose

b they use a special tapping signal with their antennae

c they smell its antennae with their antennae

d I don't know

10 What do male ants in a nest do?

a they do nothing until they leave with the young queens

b they fight off enemies that come into the nest

c they carry most of the food into the nest

d I don't know

WORD MEANINGS:

Read the following sentences carefully. Think about what the word in *slanted* writing means in that sentence. Put a circle around the meaning you think is **best**. For example:

The boy was *disobedient*. *Disobedient* means:

a collected beads

b wouldn't obey rules

c very well behaved

d I don't know

1 The netball team was well *organised*. *Organised* means:

a was popular with fans

b was very fit

c worked well together

d I don't know

2 We saw two *antennae*. *Antennae* means:

a ants nests

b feelers on the head

c ant-eating animals

d I don't know

3 They found it in the *pantry*. *Pantry* means:

a where food is stored

b where pans are stored

c where clothes are stored

d I don't know

4 The hens chased away the *strange* bird. *Strange* means:

a unfriendly

b not nice

c unknown

d I don't know

5 She liked the *scent*. *Scent* means:

a smell

b present

c parcel

d I don't know

6 He left a *trail* of rubbish. *Trail* means:

a pile

b line

c load

d I don't know

7 He has *several* problems. *Several* means:

a difficult

b a lot

c a few

d I don't know

8 The boy *squirted* the wall. *Squirted* means:

a hit with something soft and mushy

b wet with a jet of liquid

c wet with a bucket of liquid

d I don't know

9 We were given a *special* job. Special means:

a different

b well paid

c nice

d I don't know

10 That isn't *allowed*. *Allowed* means:

a good manners

b loud enough

c in the rules

d I don't know

The Moon

The moon is our nearest neighbour in space and is clearly visible in the night sky.

As the moon revolves around the earth, the shape that we can see changes. When there is a Full Moon, we see the whole face of the moon as a big circle. When there is a New Moon, we can't see the moon at all. This is because the side being lit up by the sun is facing away from the earth. We only see a Full Moon and a New Moon once a month. Most of the time we see part of the moon's face.

When the sun, moon and earth are all in a straight line, we see what is called an eclipse. A solar eclipse happens when the moon comes between the sun and the earth. The moon covers up the sun. When the earth comes between the sun and the moon, the earth's shadow falls on the moon. The shadow covers the moon, and this is called a lunar eclipse.

In 1969 the first people landed on the moon. There have been other landings since then. From these expeditions we have found out much more about the moon.

On earth a force called gravity holds us on the ground and gives us our weight. No matter how hard we try, we can never jump far off the ground because of gravity. The moon is much smaller than the earth and has much less gravity. So, a sixty kilogram person would weigh only ten kilograms. The men on the moon could bounce high off the ground, even though they were wearing heavy space suits, backpacks and boots.

The moon is composed of the same kinds of rocks as the earth. There are enormous craters in these rocks, some of them large enough for us to observe from earth. We see dark patches that make it look like the moon has a face. The craters were made by huge meteorites smashing into the surface.

The moon has no air or water, so it is a dry, lifeless place. In the sunshine the temperature reaches boiling point, but on the dark side of the moon it is three times colder than the North Pole.

The Moon

Name: _____

1 What changes do we see as the moon revolves around the earth?

_____

2 Why can't we see the moon when there is a New Moon?

_____

3 How often do we see a New Moon and a Full Moon?

_____

4 What is it called when the moon covers up the sun?

_____

5 What is it called when the earth's shadow falls on the moon?

_____

6 Why does a person weigh less on the moon than on earth?

_____

7 What is the moon composed of?

_____

8 How were the craters on the moon made?

_____

9 Why do you think people imagine "a man in the moon"?

_____

10 What sort of things do you think you would have to take to live on the moon?

_____

Name: _____

KNOWLEDGE

Think carefully about the following questions. Think about *what you know* about this subject and put a circle around the answer you think is **best**. For example:

What sort of an animal is a cat?

a a bird

b an insect

c a mammal

d I don't know

1 What are meteorites?

a bits of rock with a fiery tail

b chunks of rock or iron in space

c small, bright stars

d I don't know

2 Why do we usually see the moon in the sky at *night*?

a most of the time it isn't in the sky during the day

b it doesn't have any light during the day

c the sunlight during the day is too bright to see it

d I don't know

3 Why does it get very hot and very cold on the moon?

a the moon is closer to the sun than the earth

b there isn't a blanket of air around it to control the heat

c there aren't any trees to give shade and shelter

d I don't know

4 How often is there a Full Moon?

a every four weeks

b every two weeks

c every six weeks

d I don't know

5 What is one difference between a moon and a planet?

a moons are much hotter than planets

b moons circle around planets and planets circle around the sun

c planets have life on them and moons don't

d I don't know

6 What happens when there is a New Moon?

a we can only see the moon as a thin, curved shape

b we can see the whole face of the moon shining brightly

c we can't see the moon because the bright side is facing away from us

d I don't know

7 What is a solar eclipse?

a when the earth, the sun and the moon are in a straight line

b when the moon comes between the earth and the sun and blocks out the sun

c when the earth's shadow falls on the moon and blocks it out

d I don't know

8 What is the moon made of?

a the same material as the sun

b rocks that are different and puzzle scientists

c the same kinds of rocks as the earth

d I don't know

9 Why are people lighter on the moon than on earth?

a the moon has less gravity than the earth

b they lose a lot of weight in space ships

c there is no air to weigh them down

d I don't know

10 Why does the moon have big craters on it?

a there used to be rivers of ice that scraped out big holes

b large meteorites have crashed into the moon's surface

c there are earthquakes on the moon that have left big holes

d I don't know

WORD MEANINGS:

Read the following sentences carefully. Think about what the word in *slanted* writing means in that sentence. Put a circle around the meaning you think is best. For example:

The boy was *disobedient*. *Disobedient* means:

a collected beads

b wouldn't obey rules

c very well behaved

d I don't know

1 The rocks hit the *surface*. *Surface* means:

a big wave

b top layer

c small hill

d I don't know

2 The bushfire was *visible*. *Visible* means:

a could be seen

b caused damage

c very hot

d I don't know

3 They prepared for the *expedition*. *Expedition* means:

a hard test

b camping holiday

c scientific trip

d I don't know

4 We *observed* the ducks. *Observed* means:

a were fond of

b watched closely

c fed

d I don't know

5 What was the monster *composed of? Composed of* means:

a what it was made of

b what it wanted to do

c what it ate

d I don't know

6 Doors on hotels sometimes *revolve. Revolve* means:

a open by themselves

b swing backwards and forwards

c go round in circles

d I don't know

7 He opened the door *by force. By force* means:

a with strength

b with a clever trick

c with a tool

d I don't know

8 The country was *lifeless. Lifeless* means:

a was a good place for plants and animals

b had new kinds of plants and animals

c had no plants and animals

d I don't know

9 We saw the *crater. Crater* means:

a a kind of box

b a hole in the ground

c a kind of animal

d I don't know

10 They *bounced* on the bed. *Bounced* means:

a sprang up and down

b tossed around

c flopped down

d I don't know

Fungi: the Strange Plants

When we think of plants, we think of grass, trees, flowers and vegetables. These plants have a special green chemical in their leaves, which lets them use sunlight to make their own food. They have seeds that we can see and even collect to grow more plants.

Fungi are different sorts of plants. They don't have the special green chemical, so they can't make their own food. Instead, they live on rotting materials, such as wood or animal manure. Fungi also don't have flowers and normal seeds.

Some fungi are very small and hundreds of them have to get together before we can see them. The furry mould on old bread is made up of small fungi. But the best known fungi are mushrooms and toadstools.

Mushrooms grow from a fine powder made up of spores, the special type of seed that fungi make. When a spore sprouts, it grows into a thin thread under the ground. This thread splits into many branches and grows into a tangled mass.

When the fungus is ready, it sends up a "fruit" to make new spores. This is when we see a mushroom. At first a small ball comes through the ground, then it grows up into a stem with a cap. On the under side of the cap there are small furrows called gills. When the new spores drop from these gills, the mushroom has done its job and it rots away.

Some fungi have their own ways of spreading around their spores. The puffball is a small ball sitting on the ground. When anything touches it, it splits and puffs out a cloud of spores.

One fungus that never grows above the ground is the truffle. People in France pay a lot of money for these mushrooms, but they are very hard to find. Some farmers use pigs to sniff out the truffles, but when the pigs start digging, the farmers have to hold them back because pigs like truffles too!

There are many sorts of mushrooms that are good to eat, but many are also very poisonous. One mushroom has a red cap with white spots. It is hard to believe that such a pretty fungus could be deadly, but it is.

Fungi - the Strange Plants

Name: _____

1 What can normal plants do with the green chemical in their leaves?

_____

2 What do fungi do instead of making their own food?

_____

3 In what other ways are fungi different to normal plants?

_____

4 How is it that we can see the very small fungi on mouldy old bread?

_____

5 What are spores like and what are they for?

_____

6 What is the job of a mushroom?

_____

7 How does the puffball spread around its spores?

_____

8 Why are truffles hard to find?

_____

9 When you go out looking for field mushrooms, why do you think you often find them in groups?

_____

10 Why do you think you often find mushrooms in fields where there are horses and cows?

_____

Name: _____

KNOWLEDGE

Think carefully about the following questions. Think about *what you know* about this subject, and put a circle around the answer that you think is **best**. For example:

What sort of an animal is a cat?

a a bird

b an insect

c a mammal

d I don't know

1 Why do plants make fruit?

a to make food for animals

b to protect and spread their seeds

c to attract insects

d I don't know

2 Why do you have to grow most plants in a sunny place?

a plants need sunlight to make food

b it isn't warm enough in the shade

c bees wouldn't find the flowers in the shade

d I don't know

3 Why do plants have special ways of spreading their seeds?

a to make nice patches of the same kinds of plants

b to take over places where other plants are growing

c there wouldn't be enough room if they all grew together

d I don't know

4 How are fungi different to most other plants?

a they don't make their own food

b they come in all sorts of colours

c they grow in damp places

d I don't know

5 What do the underground parts of a mushroom look like?

a roots like other plants

b there aren't any

c bundles of threads

d I don't know

6 What sort of small fungi could you see around the house?

a moss on the garden paths

b furry mould on old bread

c small groundcover plants

d I don't know

7 What are mushrooms for?

a they attract animals, who carry their seeds to other places

b they make food for the plants in their caps

c they drop a fine powder to make new mushrooms

d I don't know

8 What are truffles?

a poisonous toadstools that are red with white spots

b mushrooms that grow buried under the ground

c balls on the ground that puff their seeds into the air

d I don't know

9 How do fungi get their food?

a they live on rotting things

b they make their own in their caps

c they get everything from any soil

d I don't know

10 How do new mushrooms start growing?

a they grow from mushroom caps that fall on the ground

b they grow from seeds like other plants

c they grow from a fine powder called spores

d I don't know

WORD MEANINGS

Read the following sentences carefully. Think about what the word in *slanted* writing means in that sentence. Put a circle around the meaning that you think is best. For example:

The boy was *disobedient*. *Disobedient* means:

a collected beads

b wouldn't obey rules

c very well behaved

d I don't know

1 She wasn't a *normal* person. *Normal* means:

a not very smart

b nice and friendly

c like everyone else

d I don't know

2 The plants were *rotting*. *Rotting* means:

a turning to mush

b turning yellow

c making new roots

d I don't know

3 We raked up the *manure*. *Manure* means:

a soil

b animal droppings

c grass clippings

d I don't know

4 There was a *mass* of colour. *Mass* means:

a little spots

b messy bits

c big clumps

d I don't know

5 He covered over the *furrow*. *Furrow* means:

a deep hole

b long, narrow hole

c vegetable seed

d I don't know

6 The seed *split*. *Split* means:

a broke apart

b went soft

c started growing

d I don't know

7 I *believed* the story. *Believed* means:

a made it up

b thought it was true

c lied about it

d I don't know

8 We all used the same *material*. *Material* means:

a the way of putting something together

b a sort of tool

c the stuff that something is made of

d I don't know

9 The plant was *sprouting*. *Sprouting* means:

a making lots of roots

b growing out of the seed

c drying out and going droopy

d I don't know

10 He had *different* problems. *Different* means:

a unlike the others

b a lot of

c hard to solve

d I don't know

The Earth and Other Planets

The Earth is only one of the many bodies that exist in space. It is one of the nine planets that circle the sun in what is called our solar system. Each planet has its own path around the sun, called an orbit. Heat and light on the planets come from the sun.

The sun is a huge ball of hot gas. It gets its heat from nuclear reactions in its centre, where the temperature is much, much hotter than anything on Earth. It is much bigger than the Earth, but it doesn't look this big because it is so far away. If we had a super fast plane that could travel at one thousand kilometres an hour, it would still take us seventeen years to get to the sun.

The Earth is just the right distance from the sun to support life. Living things need water to survive. Venus is a planet closer to the sun. It is so hot that any water there turns straight into steam. Mars is further way from the sun. It is colder than the Earth and all its water is frozen into ice caps. Plants and animals would not have any water on Venus or Mars.

Living things also need air to breathe. Some planets, like Mars, don't have much air around them at all. Venus is covered with clouds, but they contain deadly sulphuric acid. All life forms would be poisoned straight away.

Plants and animals need weather that isn't too hot or too cold. The planets closer to the sun are so hot that anything would burn to a crisp in a split second. On any of the planets further away from the sun any form of life would be frozen solid. Venus also has strong winds and lightning storms most of the time. Mars can have dust storms that last for months. Living things could not survive in this dreadful weather.

We know that there are no little green men on Mars, but our sun is only one of the many stars in space. It may be that some of these other stars have planets that could support life.

The Earth and Other Planets

Name: _____

1 What is our solar system?

_____

2 What is the sun made of?

_____

3 What happens to the water on planets closer to the sun than the Earth?

_____

4 Why couldn't plants and animals get water on Mars?

_____

5 Why couldn't living things breathe on Venus?

_____

6 What would happen to life forms on planets closer to the sun than the Earth?

_____

7 What would happen to life forms on planets further away from the sun than the Earth?

_____

8 What dreadful weather does Venus have?

_____

9 What do you think it would be like on Neptune and Pluto, the most distant planets in the solar system?

_____

10 How do you think we know there are no little green men on Mars when no-one has ever been there?

_____

Name: _____

KNOWLEDGE

Think carefully about the following questions. Think about *what you know* about this subject and put a circle around the answer you think is best. For example:

What sort of an animal is a cat?

a a bird

b an insect

c a mammal

d I don't know

1 What is one difference between planets and stars?

a stars are closer together

b stars are smaller

c stars make their own heat

d I don't know

2 What things can you find in space apart from planets and stars?

a shooting stars, or meteors

b what we see in the Milky Way

c the sun

d I don't know

3 What is the best way of finding out whether there are intelligent beings very far out in space?

a send robots to explore

b send people to explore

c send and pick up sound messages

d I don't know

4 What is the sun made of?

a rocks like on the moon

b a ball of hot gas

c rocks like in volcanoes

d I don't know

5 What is the solar system?

a the sun and nine planets

b the band of stars across the sky

c the sun, the Earth and the moon

d I don't know

6 What sort of water would you find on Mars?

a there isn't any water at all

b it is all frozen into icecaps

c rivers run across the planet

d I don't know

7 Why couldn't animals breath on the planet Venus?

a there isn't any gas around it

b there is too much steam

c the gas there is poisonous

d I don't know

8 What sort of weather could you find on Venus and Mars?

a very much like on earth

b storms much worse than on earth

c it would be colder than on earth

d I don't know

9 Is there life on the other planets in the solar system?

a no, there isn't any life

b there could be on a few planets

c there probably is on Mars

d I don't know

10 Which planets in the solar system could people land on?

a none of them

b any one of them in the near future

c probably only Mars

d I don't know

WORD MEANINGS

Read the following sentences carefully. Think about what the word in *slanted* writing means in that sentence. Put a circle around the meaning you think is best. For example:

The boy was *disobedient*. *Disobedient* means:

a collected beads

b wouldn't obey rules

c very well behaved

d I don't know

1 That film was *dreadful*. *Dreadful* means:

a exciting

b very bad

c very funny

d I don't know

2 We don't know what *exists* over there. *Exists* means:

a is hidden

b lives

c happens

d I don't know

3 We worked out the *distance*. *Distance* means:

a a space between things

b a hard problem

c a sum in maths

d I don't know

4 My auntie is *huge*. *Huge* means:

a very big

b very strong

c hugs a lot

d I don't know

5 That place can *support life*. *Support life* means:

a has firm ground for things to live on

b has all the things needed by living things

c doesn't have any danger for living things

d I don't know

6 The *temperature* was too low. *Temperature* means:

a tool for measuring heat

b hot weather

c level of hotness

d I don't know

7 That plant is *deadly*. *Deadly* means:

a can kill people

b is dying

c is already dead

d I don't know

8 We could see the *steam*. *Steam* means:

a water flowing along the ground

b the gas coming off water

c a kind of boat

d I don't know

9 They didn't *survive*. *Survive* means:

a die

b keep on living

c grow really well

d I don't know

10 The ball rolled to the *centre*. *Centre* means:

a outside

b middle

c inside

d I don't know

A for Aardvark

The aardvark is a strange African animal. Its name means "earth pig", and its long snout makes it look a lot like a pig. Unlike a pig, it has a long, thick tail, a bit like a kangaroo. It feeds only on ants and termites, and its body is well made for this job.

Termite mounds on the African plains are very large and may be up to ten feet tall. The outside of the mound is made of sunbaked mud and termite saliva. It is rock hard and protects the insects from most enemies.

The aardvark has a very good sense of smell, which it uses to find termite mounds. To rip open a mound, the aardvark has big, sharp claws on each of its strong legs. It sits on its back legs and big tail. This leaves the front legs free to rip into the mound.

Once the outside wall is ripped open, the aardvark sticks its long snout into the mound. It has a tongue nearly half a metre long. When it flicks this sticky tongue into the mound, many termites at a time stick to it and are flicked back into the aardvark's mouth. The aardvark doesn't have any front teeth, but it has flat back teeth for crushing the termites.

It could be a risky business sticking your head into a termite mound, but the aardvark is well protected. It has a very thick skin to stop insect bites. It has hairy bristles in its nose and can also close its nostrils to stop dust and termites getting in. The aardvark can fold its ears back, closing the ear openings to stop termites crawling in.

The aardvark itself could make a good meal for African hunting animals. One way it avoids hunters is to feed at night. It also has very good ears to warn it of danger. If it is frightened, it buries itself in the ground. The aardvark's legs and claws are so strong, that it can dig faster than six men with shovels.

But the biggest danger for aardvarks is not hunting animals, but people. More and more land is being cleared for farming, so the aardvark will have trouble finding food.

A for Aardvark

Name: _____

1 What sort of a tail do aardvarks have?

_____

2 How do aardvarks find termite mounds?

_____

3 Why does the aardvark need sharp claws to rip into termite mounds to get food?

_____

4 What does the aardvark do to leave its front legs free to rip into a termite mound?

_____

5 How does the aardvark flick ants into its mouth?

_____

6 What does the aardvark do to protect its nose when it sticks its head into a termite mound?

_____

7 Why does the aardvark fold its ears back?

_____

8 How is the aardvark able to bury itself quickly when it is frightened?

_____

9 What is one animal that you think might hunt the aardvark?

_____

10 Why do you think aardvarks have trouble finding food when land is cleared for farming?

_____

Name: _____

KNOWLEDGE

Think carefully about the following questions. Think about *what you know* about this subject and put a circle around the answer you think is best. For example:

What sort of an animal is a cat?

a a bird

b an insect

c a mammal

d I don't know

1 Why are ant-eating animals not eaten very often by hunters?

a the ants make the meat taste bad

b they can run too fast

c they have claws for digging and scratching

d I don't know

2 What is an aardvark?

a a South American animal that has armour plates on its back

b an African animal that looks a bit like a pig

c an American animal that has long spines on its back

d I don't know

3 What are termites?

a crawling insects that eat woody things

b crawling insects that chew holes in leaves

c flying insects that bite people

d I don't know

4 Why do termites make mounds?

a to hide from aardvarks

b to protect their eggs and grubs from the sun

c to keep out the cold rain and wind

d I don't know

5 What does an aardvark do when it is in danger?

a it rolls itself up into a hard ball

b it runs away as quickly as possible

c it digs itself into the ground

d I don't know

6 In the country where aardvarks live, what are the termite mounds like?

a holes in the ground with lots of tunnels

b big hills that are made of hard mud

c rotten logs that are hollow on the inside

d I don't know

7 How do aardvarks catch termites?

a they keep their nose close to the ground and lap them up

b they scratch the ground and scoop them up with their sharp claws

c they flick them into their mouth with a long, sticky tongue

d I don't know

8 Why don't aardvarks feel termite bites?

a the aardvarks' skin is very thick

b the aardvarks' armour is too hard

c the aardvarks' fur is very thick

d I don't know

9 How can aardvarks eat termites when other animals would have trouble?

a the eyes of other animals aren't good enough to find small insects

b other animals don't have flat teeth to squash the termites

c other animals can't close their nose and ears to stop termites crawling in

d I don't know

10 How do aardvarks get into a termite mound?

a they use their strong snout to dig up the nest

b they sit next to the mound and use their strong front claws

c they stand on the mound and scratch with their strong back claws

d I don't know

WORD MEANINGS:

Read the following sentences carefully. Think about what the word in *slanted* writing means in that sentence. Put a circle around the meaning you think is best. For example:

The boy was *disobedient*. *Disobedient* means:

a collected beads

b wouldn't obey orders

c very well behaved

d I don't know

1 He dug through the leaves with his *snout*. *Snout* means:

a small spade

b hole-digging machine

c nose

d I don't know

2 There was a big *mound* in the backyard. *Mound* means:

a kind of animal

b heap

c hole

d I don't know

3 The *bristles* felt rough. *Bristles* means:

a small rocks

b spiky flowers

c stiff hair

d I don't know

4 I want to *avoid* that person. *Avoid* means:

a stay out of their way

b tell them off

c get even with them

d I don't know

5 It was mixed with *saliva*. *Saliva* means:

a plant sap

b spit

c glue

d I don't know

6 We should love our *enemies*. *Enemies* means:

a people who hate us

b friends

c people around us

d I don't know

7 It was a *risky* trip. *Risky* means:

a exciting

b cost a lot of money

c full of danger

d I don't know

8 She *flicked* her hair back. *Flicked* means:

a combed

b quickly tossed

c curled

d I don't know

9 She *protected* the animal. *Protected* means:

a kept away danger

b made it well again

c was nice to it

d I don't know

10 They *crawled* along. *Crawled* means:

a walked slowly

b tip-toed quietly

c moved along the ground

d I don't know

Why Bees Dance

There is a very good reason for the saying "busy as a bee". When there are flowers near a beehive, bees work all day long picking up nectar to make honey. Honey keeps for a long time. It is a good way to store food for times when there are no flowers. Normally bees get their food by sipping nectar from flowers with their long tongues. At some times of the year, though, there are very few flowers. Without honey the bees might starve.

Bees send out scouts to look for good places to find flowers. When a scout comes back to the hive, it has to be able to tell the other bees where the food is. Flowers don't last very long, so all the bees have to collect nectar as quickly as possible. The scout uses special dances to send them to the right places.

When the flowers are very close to the hive, the scout does a round dance. It walks round and round in small circles outside the hive. When the nectar is further away, the scout does a waggle dance. It walks in the direction of the food, waggling its body. Then it turns back to the starting point and waggles towards the food again.

The waggle dance shows other bees only which way to go. How long the waggle dance lasts shows them how far away the food is. If the dance takes between one and two seconds, the flowers are one hundred metres away. If the dance takes two seconds, the food will be two hundred metres away, and if the dance goes on for five seconds, the food is over one kilometre away.

Bees don't only dance about food. They have a special dance to tell other bees to get water to cool the hive when it gets too hot. Otherwise, the queen and the babies inside might die.

When the group gets too big, a young queen and some of the bees leave and start a new hive. This move is called swarming. Scouts first go out looking for good places, then they use another special dance to tell the other bees where to find their new home.

Why Bees Dance

Name: _____

1 Why is honey a good way to store food?

_____

2 How do bees *normally* get their food?

_____

3 Why might bees starve without honey?

_____

4 Why do bees have to collect nectar as quickly as possible?

_____

5 When does the scout do a round dance to tell other bees where to find flowers?

_____

6 How does the scout show other bees *which way to go* in the waggle dance?

_____

7 How does the waggle dance show bees *how far away* the food is?

_____

8 Why do bees have to get water to cool the hive when it gets too hot?

_____

9 Why do you think some bees have to leave when the group gets too big?

_____

10 Why do you think bees use dancing rather than other ways to tell other bees what to do?

_____

Name: _____

KNOWLEDGE

Think carefully about the following questions. Think about *what you know* about this subject and put a circle around the answer you think is best. For example:

What sort of an animal is a cat?

a a bird

b an insect

c a mammal

d I don't know

1 What is pollen?

a something bees can make

b a fine dust made by flowers

c a sweet syrup made by flowers

d I don't know

2 How do bees send messages to each other?

a do special dances for different messages

b send sound signals with their antennae

c tap the front legs of other bees with a special code

d I don't know

3 How do wild honeybees make a hive?

a they make a mud ball on a tree

b they make a hole in a tree

c they make it in anything hollow

d I don't know

4 How do bees store honey in a hive?

a they make special wax "tubs"

b they make a pool in a special part of the hive

c they spread it around the walls of the hive

d I don't know

5 What is the nectar in flowers?

a a fine dust

b a sweet syrup

c little seeds

d I don't know

6 What does it mean when bees are *swarming*.

a a queen and some worker bees leave to start a new hive

b lots of bees fly out together to go looking for food

c lots of bees come out to attack animals near the hive

d I don't know

7 How do bees know where to find flowers?

a the queen bee tells them where to go

b they all go out and fly around until they find some

c scouts go out and tell the others where flowers are

d I don't know

8 What do bees eat most of the time?

a nectar from flowers

b honey from the hive

c pollen from flowers

d I don't know

9 Why do bees waggle their bodies?

a to cool down their bodies when they are too hot

b to show other bees where to find food

c to show other bees that they are ready to mate

d I don't know

10 Why do bees make honey?

a it is a store for when they can't find food

b they have to feed their queens honey

c that is what they eat most of the time

d I don't know

WORD MEANINGS

Read the following sentences carefully. Think about what the word in *slanted* writing means in that sentence. Put a circle around the meaning you think is best. For example:

The boy was *disobedient*. *Disobedient* means:

a collected beads

b wouldn't obey rules

c very well behaved

d I don't know

1 She gave us a *reason*. *Reason* means:

a making up a story

b a dried fruit

c explaining why

d I don't know

2 They *normally* come late. *Normally* means:

a sometimes

b most of the time

c very rarely

d I don't know

3 The people on the island *starved*. *Starved* means:

a were sick from too much fruit

b died of thirst

c died of hunger

d I don't know

4 He *sipped* the juice. *Sipped* means:

a took small mouthfuls

b sucked through a straw

c gulped down

d I don't know

5 The *scout* showed us where to go. *Scout* means:

a someone who likes camping

b the leader who gives the orders

c someone who goes ahead to look around

d I don't know

6 What makes you think it is *possible? Possible* means:

a could happen

b probably will happen

c couldn't happen

d I don't know

7 The insect *waggled* across the floor. *Waggled* means:

a moved by springing up and down

b moved its body from side to side

c moved by pulling its body along

d I don't know

8 We knew the *direction* of the house. *Direction* means:

a which streets to follow

b the address

c which way to face

d I don't know

9 Mum told us a *saying*. *Saying* means:

a a funny story

b something lots of people say

c a story with a message

d I don't know

10 We had to *collect* our clothes. *Collect* means:

a get together

b sort through

c put into neat piles

d I don't know

Appendix E:

Transfer Tests

Name: _____

How to Fill in Missing Words

Your job is to write *ONE* word on each blank space. The word has to make sense in the sentence, so you have to think carefully about what you are reading.

Examples:

1 They jumped _____ the hurdle.

2 Mary got her bag. She put her books in the _____

3 John lifted the _____ The dog was very heavy.

The Mallee Fowl

Male birds usually do their share of the work in looking after their young. _____ male Mallee Fowl, though, _____ a particularly hard-working father. _____ scratches a big hole _____ the ground for the _____ Once the eggs are _____, the female's job is _____ and she leaves.

The _____ bird covers the eggs _____ leaves and soil _____ make a big mound. _____ the leaves start rotting, _____ inside of the heap _____ warm. The male makes _____ that the temperature is _____ right for the eggs. _____ the heap is too _____, he opens the mound _____ let cool air in. _____ uncovers the eggs in _____ hottest part of the afternoon to let the sun shine on them if the mound is too cold.

Does Everything Make Sense?

When you are reading, you have to think carefully about whether everything makes sense. In these paragraphs there may be some things that don't make sense. Your job is to *underline* these things.

Examples

1 The cat was grey and white. We couldn't see her in the dark because she was all black.

2 Our dog hated bones. On his birthday we bought him some bones as a special treat.

The Return of the Salmon

Salmon are fish that spend most of their life in the ocean. When they are ready to breed, though, they make the difficult journey up the river where they were born. They have to swim up waterfalls, and on the way they may be eaten by eagles or bears. The journey back home is a happy time for the salmon.

When the salmon arrive at the top of the river, they are exhausted by their long journey. As soon as they mate and lay their eggs, they die. After several weeks, the little fish hatch out of the eggs. The parent fish show the babies the way back to the ocean. The young fish will live in the sea until they come back up the river to breed.

The Slow Sloth

Sloths are strange animals that live in the thick jungle of south America. You can't hurry a sloth; they are famous for moving very, very slowly. They spend their life high up in the trees and use their long claws to crawl along branches. You can see the sloths swinging from tree to tree in the jungle.

Sloths feed only on the leaves of trees. They have shaggy, rough fur with an unusual colour. The fur is really grey, but it looks greenish because tiny plants, called algae, grow on it. The green colour is useful for the sloths because it lets them hide in the tree tops. Sloths often come down from the trees to eat the grass and juicy berries on the ground. They always carry their babies with them.

Seaweed

In some ways, seaweed isn't like a normal plant. It doesn't have roots like other plants. Instead it has a flat pad that holds it onto rocks. It gets minerals, such as iron, from the sea water instead of from the soil. Seaweed has to grow on sand so that it can send down its deep roots.

Seaweed doesn't have flowers, but makes its small, powdery seeds, called spores, in pods on the end of its leaves. When the pods burst, the water carries the spores to different places.

Like ordinary plants, seaweed needs sunlight to make sugar in its leaves. The leaves of seaweed have to be very tough to stop them being smashed by the waves. Seaweed often grows in the dark because it can get sugar from the sea water.

Name: _____

How to Fill in Missing Words

Your job is to write *ONE* word on each blank space. The word has to make sense in the sentence, so you have to think carefully about what you are reading.

Examples:

1 They jumped _____ the hurdle.

2 Mary got her bag. She put her books in the _____

3 John lifted the _____ The dog was very heavy.

Shooting Stars

Most people have seen the streak of light in the night sky that we call a shooting star. _____ fact, these bright flashes _____ not stars at all. _____ are really meteors, which _____ bits of rock, sometimes _____ the size of a _____ of sand. As they _____ through the Earth's atmosphere, _____ burn up. This burning _____ is what we see _____ a shooting star. The _____ the meteor, the longer _____ takes to burn.

Sometimes _____ meteors that fall through _____ atmosphere are so large _____ they don't burn up _____ they reach the ground. _____ are called meteorites, and _____ they hit the Earth, _____ can leave big craters. _____ for us, large meteorites don't often hit the Earth.

Does Everything Make Sense?

When you are reading, you have to think carefully about whether everything makes sense. In these paragraphs there may be some things that don't make sense. Your job is to *underline* these things.

Examples

1 The cat was grey and white. We couldn't see her in the dark because she was all black.

2 Our dog hated bones. On his birthday we bought him some bones as a special treat.

Sea Mammals

Most mammals live on the land or only go into the water sometimes. But dolphins, whales and dugongs spend their whole lives in the water. They never come onto the land and even give birth to their young in the water.

Dugongs are also called sea cows because they graze on seaweed from the ocean floor. Like all mammals, they feed their young on milk. They have to find a sandy beach to give birth to their calves.

Dolphins are well suited to life in the sea, but like all mammals they have to come to the surface to breathe air. Dolphins are friendly and like to leap out of the water near boats. They are good divers because they can breathe underwater. Dolphins catch fish in the water for food.

Fossils

Most fossils are found in rocks. When an animal dies, the soft parts of its body, like the skin and muscles, are eaten by other animals or rot away. Hard parts like the bones, teeth or shells, are left behind. They may be covered by sand or mud. Over a very long time, the weight of this mud presses the bones into the rocks. You can tell where an animal lived by studying the skin and fur on a fossil.

Some fossils are found in ice. When these animals died, their bodies were frozen so quickly that there was no time for the soft parts to rot away. Scientists found a whole baby mammoth in the ice. Even though they only found its bones, they could still work out what it looked like.

Pumas - the American Big Cats

The puma or mountain lion is one of the biggest cats in north America. Its body is about one metre long. It is a greyish colour all over. It can be found from the snowy mountains in the north to the dry mountains in the south.

Pumas have three or four cubs at a time. When they are born they have spots, just like their cousins, the leopards. As they grow up, they keep these spots. They eat whatever they can catch, from deer and rabbits, right down to mice and grasshoppers when times are hard. Pumas grow to the size of a large house cat.

Pumas have a problem because a lot of the places where they live have been turned into farms, and farmers often shoot pumas.

Name: _____

How to Fill in Missing Words

Your job is to write *ONE* word on each blank space. The word has to make sense in the sentence, so you have to think carefully about what you are reading.

Examples:

1 They jumped _____ the hurdle.

2 Mary got her bag. She put her books in the _____

3 John lifted the _____ The dog was very heavy.

Cuckoos - the Lazy Parents

Most birds are busy and caring parents, but cuckoos don't bother raising their own young. _____ female cuckoo lays her _____ in an other bird's _____ and leaves.

The foster _____ sit on the eggs _____ keep them warm, and _____ the hungry chicks when _____ hatch. The cuckoo chick _____ before the other chicks. _____ is also bigger than _____ other chicks. Its big _____ lets it grab most _____ the food the adults _____ back. Often the cuckoo _____ pushes other eggs or _____ out of the nest.

_____ don't have it all _____ own way. Some birds _____ very good eyes and _____ any egg that doesn't _____ the same out of the nest. Some birds chase cuckoos away as soon as they see them.

Does Everything Make Sense?

When you are reading, you have to think carefully about whether everything makes sense. In these paragraphs there may be some things that don't make sense. Your job is to *underline* these things.

Examples

1 The cat was grey and white. We couldn't see her in the dark because she was all black.

2 Our dog hated bones. On his birthday we bought him some bones as a special treat.

Starfish

Starfish are very simple animals that have five arms around a central disc. Their head only has a big mouth, but no eyes or ears, and they have a strange way of feeding. When they see food, such as shellfish, they push their stomach out through their mouth. They cover the food with this stomach and start to digest it before they swallow it. Starfish don't have very good table manners!

Starfish have lots of small tubes or suckers on their arms that let them crawl along the ocean floor. This is the only way they move about. If they lose one of their arms, they can grow another one in its place. They use these arms to swim through the water. Starfish come in many different colours.

Water on the Earth

On high mountains a blanket of snow covers the land all winter long. When the weather gets warmer in spring, some of the snow melts, runs down into creeks, and finally into rivers in the valley. From there it flows out into the ocean. Sometimes these mountains are a long way from the sea. They are too far away for the melting snow to run into the ocean.

When rain falls on the land, some of the water sinks into the ground. There are big pools under the ground. Sometimes this water comes up by itself to make springs. This is where we get mineral water from. Sometimes people dig wells in the ground to get the water out. Once water is trapped underground, it can't come out again.

Sea Snakes

Most snakes can swim, but sea snakes are specially designed to live in the ocean. They have a flattened tail to help them steer through the water. They don't have to lay eggs like other snakes. Instead, they give birth to live young in the water. They still have to come to the surface to breathe air. When a female sea snake is ready to lay her eggs, she looks for a safe place on the ocean floor.

Sea snakes have a very strong poison. Luckily for us, they don't have a nasty temper and bite mainly to catch fish. They also have short fangs, which wouldn't go very far into our skin. If you are diving, you have to be on the look-out for sea snakes because they often attack divers.

Survey of Opinions About Reading

Example: *I love homework.*

First decision: Which side of the double line, true or not true?
Second decision: Which box, *how* true (or untrue) is it for you?

Example: *I don't like icecream.*

First decision: Which side of the double line, true or not true?
Second decision: Which box, *how* true (or untrue) is it for you?

1 I am a good reader.

2 The only time I read is when the teacher makes me.

3 I think I will have problems reading in high school.

4 Being a good reader is important for doing well at school.

5 Reading takes a lot of effort for me.

6 You don't have to be a good reader when you leave school.

7 I really enjoy reading.

8 I won't be able to improve my reading.

9 Reading lets me find out lots of interesting things.

10 I want to learn more about being a good reader.

TRUE		NOT TRUE '	
Very true	Sort of true	Not really true	Not at all true

Bibliography

Adams, A., Carnine, D. & Gersten, R. (1982), 'Instructional strategies for studying content area texts in the intermediate grades', *Reading Research Quarterly* **18**, 27–55.

Afflerbach, P. (1990), 'The influence of prior knowledge on expert readers' main idea construction strategies', *Reading Research Quarterly* **25**, 31–46.

Anderson, R. (1980), Schema-directed processes in language comprehension, *in* J. Hartley, ed., 'The Psychology of Written Communications', Kogan Page, London.

Anderson, R., Reynolds, R., Schallert, D. & Goetz, E. (1977), 'Frameworks for comprehending discourse', *American Educational Research Journal* **14**, 367–381.

Andre, M. & Anderson, T. (1978-9), 'The development and evaluation of a self-questioning study technique', *Reading Research Quarterly* **14**, 605–623.

Andre, T. (1979), 'Does answering higher-level questions while reading facilitate productive learning?', *Review of Educational Research* **49**, 280–318.

Armbruster, B., Anderson, T. & Ostertag, J. (1987), 'Does text structure/summarization training instruction facilitate learning for expository text?', *Reading Research Quarterly* **22**, 331–346.

Armbruster, B., Anderson, T., Armstrong, J., Wise, M., Janish, C. & Meyer, L. (1991), 'Reading and questioning in content area lessons', *Journal of Reading Behaviour* **23**, 35–59.

August, D., Flavell, J. & Clift, R. (1984), 'Comparison of comprehension monitoring of skilled and less skilled readers', *Reading Research Quarterly* **20**, 39–53.

Baker, L. (1979), 'Comprehension monitoring: identifying and coping with text confusions', *Journal of Reading Behaviour* **11**, 365–374.

Baker, L. (1984*a*), 'Children's effective use of multiple standards for evaluating their comprehension', *Journal of Educational Psychology* **76**, 588–597.

Baker, L. (1984*b*), 'Spontaneous versus instructed use of multiple standardss for evaluating comprehension: effects of age, reading proficiency, and type of standard', *Journal of Experimental Child Psychology* **38**, 289–311.

Baker, L. (1985*a*), 'Differences in the standards used by college students to evaluate their comprehension of expository prose', *Reading Research Quarterly* **20**, 297–313.

Baker, L. (1985*b*), How do we know when we don't understand? standards for evaluating text comprehension, *in* D. Forrest, G. MacKinnon & T. Waller, eds, 'Metacognition, Cognition and Human Performance (vol. 1)', Academic Press, New York.

Baker, L. & Anderson, R. (1982), 'Effects of inconsistent information on text processing: evidence for comprehension monitoring', *Reading Research Quarterly* **17**, 281–294.

Baker, L. & Brown, A. (1984*a*), Cognitive monitoring in reading, *in* J. Flood, ed., 'Understanding Reading Comprehension', International Reading Association, Newark Del.

Baker, L. & Brown, A. (1984*b*), Metacognitive skills and reading, *in* P. Pearson, ed., 'Handbook of Reading Research', Longman, New York.

Baker, L. & Zimlin, L. (1989), 'Instructional effects on children's use of two levels of standards for evaluating their comprehension', *Journal of Educational Psychology* **81**, 340–346.

Barclay, J., Bransford, J., Franks, J., McCarrell, N. & Nitsch, K. (1974), 'Comprehension and semantic flexibility', *Journal of Verbal Learning and Verbal Behaviour* **13**, 471–481.

Baumann, J. (1984), 'The effectiveness of a direct instruction paradigm for teaching main idea comprehension', *Reading Research Quarterly* **20**, 93–115.

Beal, C. (1990), 'The development of text evaluation and revision skills', *Child Development* **61**, 247–258.

Beck, I., Perfetti, C. & McKeown, M. (1982), 'Effect of long-term vocabulary instruction on lexical access and reading comprehension', *Journal of Educational Psychology* **74**, 506–521.

Berkowitz, S. (1986), 'Effects of instruction in text organization on sixth-grade students' memory for expository reading', *Reading Research Quarterly* **21**, 161–178.

Biemiller, A. (1970), 'The development of the use of graphic and contextual information as children learn to read', *Reading Research Quarterly* **6**, 75–90.

Biemiller, A. (1977–8), 'Relationships between oral reading rates for letters, words, and simple text in the development of reading achievement', *Reading Research Quarterly* **13**, 223–253.

Book, C., Duffy, G., Meloth, L. R. M. & Vavrus, L. (1985), 'A study of the relationship between teacher explanation and student metacognitive awareness during reading instruction', *Communication Education* **34**, 29–36.

Borkowski, J. & Kurtz, B. (1987), Metacognition and executive control, *in* J. Borkowski & J. Day, eds, 'Intelligence and Cognition in Special Children: Perspectives on Retardation, Learning Disabilities and Giftedness', Ablex, Norwood N.J.

Borkowski, J., Carr, M. & Pressley, M. (1987), "spontaneous' strategy use: perspectives from metacognitive theory', *Intelligence* **11**, 61–75.

Borkowski, J., Weyhing, R. & Carr, M. (1988), 'Effects of attributional retraining on strategy-based reading comprehension in learning disabled students', *Journal of Educational Psychology* **80**, 46–53.

Bransford, J. & Johnson, M. (1972), 'Contextual prerequisites for understanding: some investigations of comprehension and recall', *Journal of Verbal Learning and Verbal Behaviour* **11**, 717–726.

Bransford, J. & McCarrell, N. (1974), A sketch of a cognitive approach to comprehension: some thoughts about understanding and what it means to comprehend, *in* W. Weiner & D. Palermo, eds, 'Cognition and the Symbolic Processes', Erlbaum, Hillsdale N.J.

Bransford, J., Stein, B., Nye, N., Franks, J., Auble, P., Meyzinski, K. & Perfetto, G. (1982), 'Differences in approaches to learning: an overview', *Journal of Experimental Psychology: General* **111**, 390–398.

Brewer, W. (1977), 'Memory for the pragmatic implications of sentences', *Memory and Cognition* **5**, 673–678.

Bridge, C. & Winograd, P. (1982), 'Readers' awareness of cohesive relationships during cloze comprehension', *Journal of Reading Behaviour* **14**, 299–312.

Bridge, C., Belmore, S., Moskow, S., Cohen, S. & Matthews, P. (1984), 'Topicalization and memory for main ideas in prose', *Journal of Reading Behaviour* **16**, 61–80.

Brien, O., Shank, D., Myers, J. & Rayner, K. (1988), 'Elaborative inferences during reading: do they occur on-line', *Journal of Experimental Psychology: Learning, Memory and Cognition* **14**, 410–420.

Brooks, L. & Dansereau, D. (1983), 'Effects of structural schema training and text organization on expository prose processing', *Journal of Educational Psychology* **75**, 811–820.

Brooks, L., Dansereau, D., Holley, C. & Spurlin, J. (1983), 'Generation of descriptive text headings', *Contemporary Educational Psychology* **8**, 103–108.

Brooks, L., Dansereau, D., Spurlin, J. & Holley, C. (1983), 'Effects of headings on text processing', *Journal of Edcuational Psychology* **75**, 292–302.

Brown, A. (1980), Metacognitive development and reading, *in* R. Spiro, B. Bruce & W. Brewer, eds, 'Theoretical Issues in Reading Comprehension', Erlbaum, Hillsdale N.J.

Brown, A. & Day, J. (1983), 'Macrorules for summarizing texts: the development of expertise', *Journal of Verbal Learning and Verbal Behaviour* **22**, 1–14.

Brown, A. & DeLoache, J. (1978), Skills, plans and self-regulation, *in* R. Siegler, ed., 'Children's Thinking - What Develops?', Erlbaum, Hillsdale N.J.

Brown, A. & Palincsar, A. (1982), 'Induced strategic learning from texts by means of informed, self-control training', *Topics in Learning and Learning Disabilities* **2**, 1–17.

Brown, A. & Smiley, S. (1977), 'Rating the importance of structural units of prose passages: a problem of metacognitive development', *Child Development* **48**, 1–8.

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

Brown, A., Campione, J. & Day, J. (1981), 'Learning to learn: on training students to learn from texts', *Educational Researcher* **10**, 14–21.

Brown, A., Day, J. & Jones, R. (1983), 'The development of plans for summarizing texts', *Child Development* **54**, 968–979.

Butkowsky, I. & Willows, D. (1980), 'Cognitive motivational characteristics of children varying in reading ability: evidence for learned helplessness in poor readers', *Journal of Educational Psychology* **72**, 408–422.

Capelli, C. & Markman, E. (1982), 'Suggestions for training comprehension monitoring', *Topics in Learning and Learning Disabilities* **2**, 87–96.

Carpenter, P. & Daneman, M. (1981), 'Lexical retrieval and error recovery in reading: a model based on eye fixations', *Journal of Verbal Learning and Verbal Behaviour* **20**, 137–160.

Carpenter, P. & Just, M. (1977), Reading comprehension as eyes see it, *in* M. Just & P. Carpenter, eds, 'Cognitive Processes in Comprehension', Erlbaum, Hillsdale N.J.

Carpenter, P. & Just, M. (1989), The role of working memory in language comprehension, *in* D. Klahr & K. Kotovsky, eds, 'Complex Information Processing - the Impact of H.A. Simon', Erlbaum, Hillsdale N.J.

Carr, E., Dewitz, P. & Patberg, J. (1983), 'The effect of inference training on children's comprehension of expository text', *Journal of Reading Behaviour* **15**, 1–18.

Carr, M., Borkowski, J. & Maxwell, S. (1991), 'Motivational components of underachievement', *Developmental Psychology* **27**, 108–118.

Chabot, R., Petros, T. & McCord, G. (1983), 'Developmental and reading ability differences in accessing information from semantic memory', *Journal of Experimental Child Psychology* **35**, 128–142.

Chan, L. (1988), 'Comprehension monitoring in students with reading difficulties', *Australian Educational Researcher* **15**, 19–39.

Chan, L. & Cole, P. (1986), 'The effects of comprehension monitoring training on the reading competence of learning disabled and regular class students', *Remedial and Special Education* **7**, 33–40.

Chang, F. (1980), 'Active memory processes in visual sentence comprehension: clause effects and pronominal reference', *Memory and Cognition* **8**, 58–64.

Chiesi, H., Spilich, G. & Voss, J. (1979), 'Acquisition of domain-related information in relation to high and low domain knowledge', *Journal of Verbal Learning and Verbal Behaviour* **18**, 257–273.

Clark, H. & Clark, E. (1977), *Psychology and Language*, Harcourt Brace Jovanovich, New York.

Clark, H. & Sengul, C. (1979), 'In search of referents for nouns and pronouns', *Memory and Cognition* **7**, 35–41.

Collins, A. & Loftus, E. (1975), 'A spreading activation theory of semantic processing', *Psychological Review* **82**, 407–428.

Cook, L. & Mayer, R. (1988), 'Teaching readers about the structure of scientific text', *Journal of Educational Psychology* **80**, 448–456.

Corbett, A. & Chang, F. (1983), 'Pronoun disambiguation: accessing potential antecedents', *Memory and Cognition* **11**, 283–294.

Corbett, A. & Dosher, B. (1978), 'Instrumental inferences in sentence encoding', *Journal of Verbal Learning and Verbal Behaviour* **17**, 479–491.

Cunningham, J. & Moore, D. (1986), The confused world of main idea, *in* J. Baumann, ed., 'Teaching Main Idea Comprehension', International Reading Association, Newark Del.

Curtis, M. (1980), 'Development of components of reading skill', *Journal of Educational Psychology* **72**, 656–669.

Daneman, M. (1984), Why some people are better readers than others: a process and storage account, *in* R. Sternberg, ed., 'Advances in the Psychology of Human Intelligence', Erlbaum, Hillsdale N.J.

Daneman, M. & Carpenter, P. (1980), 'Individual differences in working memory and reading', *Journal of Verbal Learning and Verbal Behaviour* **19**, 450–466.

Daneman, M. & Carpenter, P. (1983), 'Individual differences in integrating information between and within sentences', *Journal of Experimental Psychology: Learning, Memory and Cognition* **9**, 561–583.

Danner, F. (1976), 'Children's understanding of intersentence organization in the recall of short descriptive passages', *Journal of Educational Psychology* **68**, 174–183.

Davey, B. & McBride, S. (1986), 'Effects of question-generation training on reading comprehension', *Journal of Educational Psychology* **78**, 256–262.

Day, J. (1986), 'Teaching summarization skills: influences of student ability level and strategy difficulty', *Cognition and Instruction* **3**, 193–210.

Dewitz, P., Carr, E. & Patberg, J. (1987), 'Effects of inference training on comprehension and comprehension monitoring', *Reading Research Quarterly* **22**, 99–121.

Doctorow, M., Wittrock, M. & Marks, C. (1978), 'Generative processes in reading comprehension', *Journal of Educational Psychology* **70**, 109–118.

Dole, J., Duffy, G., Roehler, L. & Pearson, D. (1991), 'Moving from the old to the new: research on reading comprehension instruction', *Review of Educational Research* **61**, 239–264.

Dole, J., Valencia, S., Greer, E. & Wardrop, J. (1991), 'Effects of two types of prereading instruction on the comprehension of narrative and expository text', *Reading Research Quarterly* **26**, 142–159.

Dooling, D. & Mullet, R. (1973), 'Locus of thematic effects in retention of prose', *Journal of Experimental Psychology* **97**, 404–6.

Duffy, G., Roehler, L., Meloth, M., Vavrus, L., Book, C., Putnam, J. & Wesselman, R. (1986), 'The relationship between explicit verbal explanations during reading skill instruction and student awareness and achievement: a study of reading teacher effects', *Reading Research Quarterly* **21**, 237–252.

Duffy, S., Morris, R. & Rayner, K. (1988), 'Lexical ambiguity and fixation times in reading', *Journal of Memory and Language* **27**, 429–446.

Dweck, C. & Reppucci, N. (1973), 'Learned helplessness and reinforcement responsibility in children', *Journal of Personality and Social Psychology* **25**, 109–116.

Ehrlich, K. (1983), Eye movements in pronoun assignment: a study of sentence integration, *in* K. Rayner, ed., 'Eye Movements in Reading - Perceptual and Language Processes', Academic Press, New York.

Ehrlich, K. & Rayner, K. (1983), 'Pronoun assignment and semantic integration during reading: eye movements and immediacy of processing', *Journal of Verbal Learning and Verbal Behaviour* **22**, 75–87.

Englert, C. & Hiebert, E. (1984), 'Children's developing awareness of text structures in expository materials', *Journal of Educational Psychology* **76**, 65–74.

Englert, C., Hiebert, E. & Stewart, S. (1988), 'Detecting and correcting inconsistencies in the monitoring of expository prose', *Journal of Educational Research* **81**, 221–227.

Epstein, W., Glenberg, A. & Bradley, M. (1984), 'Coactivation and comprehension: contribution of text variables to the illusion of knowing', *Memory and Cognition* **12**, 355–360.

Fincher-Kiefer, R., Post, T., Greene, T. & Voss, J. (1988), 'On the role of prior knowledge and task demands in the processing of text', *Journal of Memory and Language* **27**, 416–428.

Flavell, J. (1981), Cognitive monitoring, *in* W. Dickson, ed., 'Children's Oral Communication Skills', Academic Press, New York.

Fleisher, L., Jenkins, J. & Pany, P. (1979), 'Effects on poor readers' comprehension of training inrapid decoding', *Reading Research Quarterly* **15**, 30–48.

Frazier, L. & Rayner, K. (1982), 'Making and correcting errors during sentence comprehension: eye movements in the analysis of structurally ambiguous sentences', *Cognitive Psychology* **14**, 178–210.

Frederiksen, J. (1981), Sources of process interactions in reading, *in* A. Lesgold & C. Perfetti, eds, 'Interactive Processes in Reading', Erlbaum, Hillsdale N.J.

Friedman, F. & Rickards, J. (1981), 'Effect of level, review and sequence of inserted questions on text processing', *Journal of Educational Psychology* **73**, 427–436.

Gambrell, L. & Heathington, B. (1981), 'Adult disabled readers' metacognitive awareness about reading tasks and strategies', *Journal of Reading Behaviour* **13**, 215–222.

Garner, R. (1980), 'Monitoring of understanding: an investigation of good and poor readers' awareness of induced miscomprehension of text', *Journal of Reading Behaviour* **12**, 55–63.

Garner, R. (1981), 'Monitoring of passage inconsistency among poor comprehenders: a preliminary test of the 'piecemeal processing' explanation', *Journal of Educational Research* **74**, 159–162.

Garner, R. & Gillingham, M. (1987), 'Students' knowledge of text structure', *Journal of Reading Behaviour* **19**, 247–259.

Garner, R. & Kraus, C. (1981–2), 'Good and poor comprehender differences in knowing and regulating reading behaviours', *Educational Research Quarterly* **6**, 5–12.

Garner, R. & Reis, R. (1981), 'Monitoring and resolving comprehension obstacles: an investigation of spontaneous text lookbacks among upper-grade good and poor comprehenders', *Reading Research Quarterly* **16**, 569–582.

Garner, R., Alexander, P., Slater, W., Hare, V., Smith, T. & Reis, R. (1986), 'Children's knowledge of structural properties of expository text', *Journal of Educational Psychology* **78**, 411–416.

Garner, R., Hare, V., Alexander, P., Haynes, J. & Winograd, P. (1984), 'Inducing use of a text lookback strategy among unsuccessful readers', *American Educational Research Journal* **21**, 789–798.

Garner, R., Wagoner, S. & Smith, T. (1983), 'Externalizing question-answering strategies of good and poor comprehenders', *Reading Research Quarterly* **18**, 439–447.

Garrod, S. & Sanford, T. (1981), Bridging inferences and the extended domain of reference, *in* J. Long & A. Baddeley, eds, 'Attention and Performance (vol. 9)', Erlbaum, Hillsdale N.J.

Garvey, C., Caramazza, A. & Yates, J. (1975), 'Factors influencing assignment of pronoun antecedents', *Cognition* **3**, 227–243.

Glenberg, A., Wilkinson, A. & Epstein, W. (1982), 'The illusion of knowing: failure in the self-assessment of comprehension', *Memory and Cognition* **10**, 597–602.

Glover, J., Plake, B., Roberts, B., Zimmer, J. & Palmere, M. (1981), 'Distinctiveness of encoding: the effects of paraphrasing and drawing inferences on memory from prose', *Journal of Educational Psychology* **73**, 736–744.

Goetz, E. (1979), 'Inferring from text: some factors influencing which inferences will be made', *Discourse Processes* **2**, 179–195.

Goldman, S., Hogaboam, T., Bell, L. & Perfetti, C. (1980), 'Short-term retention of discourse during reading', *Journal of Educational Psychology* **72**, 647–655.

Gordon, C. & Pearson, P. (1983), The effects of instruction in metacomprehension and inferencing on children's comprehension abilities, Technical Report 277, Centre for the Study of Reading, University of Illinois.

Gough, P. (1972), One second of reading, *in* J. Kavanagh & F. Meltingly, eds, 'Language by Ear and by Eye', M.I.T. Press, Cambridge Mass.

Grabe, M. (1980), 'Utilization of imposed structure: the impact of reading competence and grade level', *Journal of Reading Behaviour* **12**, 31–40.

Grabe, M., Antes, J., Kahn, H. & Kristjanson, A. (1991), 'Adult and adolescent readers' comprehension monitoring performance: an investigation of monitoring accuracy and related eye movements', *Contemporary Educational Psychology* **16**, 45–60.

Graesser, A., Hoffman, N. & Clark, L. (1980), 'Structural components of reading time', *Journal of Verbal Learning and Verbal Behaviour* **19**, 135–151.

Gronlund, N. (1981), *Measurement and Evaluation in Teaching*, fourth edition, Macmillan, New York.

Guindon, R. & Kintsch, W. (1984), 'Priming macropropositions: evidence for the primacy of macropropositions in the memory for text', *Journal of Verbal Learning and Verbal Behaviour* **23**, 508–518.

Haberlandt, K. & Bingham, G. (1978), 'Verbs contribute to the coherence of brief narratives: reading related and unrelated triples', *Journal of Verbal Learning and Verbal Behaviour* **17**, 419–425.

Haberlandt, K. & Graesser, A. (1985), 'Component processes in text comprehension and some of their interactions', *Journal of Experimental Psychology: General* **114**, 357–374.

Haberlandt, K. & Graesser, A. (1989), 'Processing of new arguments at clause boundaries', *Memory and Cognition* **17**, 186–193.

Hansen, J. (1981), 'The effects of inference training and practice on young children's comprehension', *Reading Research Quarterly* **16**, 391–417.

Hansen, J. & Pearson, P. (1983), 'An instructional study: improving the inferential comprehension of good and poor fourth-grade readers', *Journal of Educational Psychology* **75**, 821–829.

Hare, V. & Borchardt, K. (1984), 'Direct instruction of summarization skills', *Reading Research Quarterly* **20**, 62–78.

Hare, V., Rabinowitz, M. & Schieble, K. (1989), 'Text effects on main idea comprehension', *Reading Research Quarterly* **24**, 72–88.

Harris, P., Kruithof, A., Terwogt, M. & Visser, T. (1981), 'Children's detection and awareness of textual anomaly', *Journal of Experimental Child Psychology* **31**, 212–230.

Haviland, S. & Clark, H. (1974), 'What's new? acquiring new information as a process in comprehension', *Journal of Verbal Learning and Verbal Behaviour* **13**, 512–521.

Hays, W. (1963), *Statistics for Psychologists*, Holt, Rinehart and Winston, New York.

Helfeldt, J. & Lalik, R. (1976), 'Reciprocal student-teacher questioning', *The Reading Teacher* **December**, 283–287.

Henk, W. (1982), 'A response to shanahan, kamil and tobin: the case is not yet clozed', *Reading Research Quarterly* **17**, 591–595.

Herdman, C. & LeFevre, J. (1992), 'Individual differences in the efficiency of word recognition', *Journal of Educational Psychology* **84**, 95–102.

Hiebert, E., Englert, C. & Brennan, S. (1983), 'Awareness of text structure in recognition and production of expository discourse', *Journal of Reading Behaviour* **15**, 63–79.

Holley, C., Dansereau, D., Evans, S., Collins, K., Brooks, L. & Larson, D. (1981), 'Utilizing intact and embedded headings as processing aids with non-narrative text', *Contemporary Educational Psychology* **6**, 227–236.

Holmes, B. (1987), 'Children's inferences with print and pictures', *Journal of Educational Psychology* **79**, 14–18.

Howell, M. & Manis, F. (1986), 'Developmental and reader ability differences in semantic processing efficiency', *Journal of Educational Psychology* **78**, 124–129.

Jacobs, J. & Paris, S. (1987), 'Children's metacognition about reading: issues in definition, measurement and instruction', *Educational Psychologist* **22**, 255–278.

Johnson, H. & Smith, L. (1981), 'Children's inferential ability in the context of reading to understand', *Child Development* **52**, 1216–1223.

Johnson, M., Bransford, J. & Solomon, S. (1973), 'Memory for tacit implications of sentences', *Journal of Experimental Psychology* **98**, 203–205.

Johnston, P. & Winograd, P. (1985), 'Passive failure in reading', *Journal of Reading Behaviour* **17**, 279–301.

Just, M. & Carpenter, P. (1980), 'A theory of reading from eye fixations to comprehension', *Psychological Review* **87**, 329–353.

Just, M. & Carpenter, P. (1987), *The Psychology of Reading and Language Comprehension*, Allyn and Bacon, Boston.

Kaufman, N., Randlett, A. & Price, J. (1985), 'Awareness of the use of comprehension strategies in good and poor college readers', *Reading Psychology* **6**, 1–11.

Kavale, K. & Schreiner, R. (1979), 'The reading processes of average and above average readers: a comparison of the use of reasoning strategies in responding to standardized comprehension measures', *Reading Research Quarterly* **15**, 102–128.

Kieras, D. (1980), 'Initial mention as a signal to thematic content in technical passages', *Memory and Cognition* **8**, 345–353.

Kintsch, W. (1988), 'The role of knowledge in discourse comprehension: a construction-integration model', *Psychological Review* **95**, 163–182.

Kintsch, W. & Mross, E. (1985), 'Context effects in word identification', *Journal of Memory and Language* **24**, 336–349.

Kintsch, W. & van Dijk, T. (1978), 'Toward a model of text comprehension and production', *Psychological Review* **85**, 363–394.

Kintsch, W. & Yarborough, J. (1982), 'Role of rhetorical structure in text comprehension', *Journal of Educational Psychology* **74**, 828–834.

Kletzien, S. (1991), 'Strategy use by good and poor comprehenders reading expository text of differing levels', *Reading Research Quarterly* **26**, 67–85.

Kuhn, D. & Phelps, E. (1982), 'The development of problem-solving strategies', *Advances in Child Development and Behaviour* **17**, 1–44.

Kurtz, B. & Borkowski, J. (1987), 'Development of strategic skills in impulsive and reflective children: a longitudinal study of metacognition', *Journal of Experimental Child Psychology* **43**, 129–148.

LaBerge, D. & Samuels, S. (1974), 'Toward a theory of automatic information processing in reading', *Cognitive Psychology* **6**, 293–323.

Lipson, M. (1983), 'The influence of religious affiliation on children's memory for text information', *Reading Research Quarterly* **18**, 448–457.

Loman, N. & Mayer, R. (1983), 'Signaling techniques that increase the understanding of expository prose', *Journal of Educational Psychology* **75**, 402–412.

Lorch, R. & Lorch, E. (1985), 'Topic structure representation and text recall', *Journal of Educational Psychology* **77**, 137–148.

Lorch, R., Lorch, E. & Inman, W. (1993), 'Effects of signaling topic structure on text recall', *Journal of Educational Psychology* **85**, 281–290.

Lorch, R., Lorch, E. & Klusewitz, M. (1993), 'College students' conditional knowledge about reading', *Journal of Educational Psychology* **85**, 239–352.

Lorch, R., Lorch, E. & Matthews, P. (1985), 'On-line processing of the topic structure of a text', *Journal of Memory and Language* **24**, 350–362.

Lysynchuk, L., Pressley, M., d'H. Ailly, Smith, M. & Cake, H. (1989), 'A methodological analysis of experimental studies of comprehension strategy instruction', *Reading Research Quarterly* **24**, 458–470.

MacDonald, J. (1986), 'Self-generated questions and reading recall: Does training help?', *Contemporary Educational Psychology* **11**, 290–304.

Manelis, L. & Yekovich, F. (1976), 'Repetitions of propositional arguments in sentences', *Journal of Verbal Learning and Verbal Behaviour* **15**, 301–312.

Manzo, A. (1969), 'The request procedure', *Journal of Reading* **November**, 123–126.

Markman, E. (1979), 'Realizing that you don't understand: elementary school children's awareness of inconsistencies', *Child Development* **50**, 643–655.

Markman, E. (1981), Comprehension monitoring, *in* W. Dickson, ed., 'Children's Oral Communication Skills', Academic Press, New York.

Markman, E. (1985), Comprehension monitoring: developmental and educational issues, *in* S. Chipman, J. Segal & R. Glaser, eds, 'Thinking and Learning Skills (vol. 2)', Erlbaum, Hillsdale N.J.

Markman, E. & Gorin, L. (1981), 'Children's ability to adjust their standards for evaluating comprehension', *Journal of Educational Psychology* **73**, 320–325.

McConkie, G. & Zola, D. (1981), Language constraints and the functional stimulus of reading, *in* A. Lesgold & C. Perfetti, eds, 'Interactive Processes in Reading', Erlbaum, Hillsdale N.J.

McCormick, S. (1992), 'Disabled readers' erroneous responses to inferential comprehension questions: description and analysis', *Reading Research Quarterly* **27**, 54–77.

McGee, L. (1982), 'Awareness of text structure: effects on children's recall of expository text', *Reading Research Quarterly* **17**, 581–595.

McKoon, G. & Ratcliff, R. (1980), 'The comprehension processes and memory structures involved in anaphoric reference', *Journal of Verbal Learning and Verbal Behaviour* **19**, 668–682.

McKoon, G. & Ratcliff, R. (1981), 'The comprehension processes and memory structures involved in instrumental inference', *Journal of Verbal Learning and Verbal Behaviour* **20**, 671–682.

McKoon, G. & Ratcliff, R. (1988), 'Contextually relevant aspects of meaning', *Journal of Experimental Psychology: Learning, Memory and Cognition* **14**, 331–343.

McKoon, G. & Ratcliff, R. (1992), 'Inference during reading', *Psychological Review* **99**, 440–466.

McLain, K., Gridley, B. & McIntosh, D. (1991), 'Value of a scale used to measure metacognitive reading awareness', *Journal of Educational Research* **85**, 81–87.

Medin, D. & Shoben, E. (1988), 'Context and structure in conceptual combination', *Cognitive Psychology* **20**, 158–190.

Medin, D. & Smith, E. (1984), 'Concepts and concept formation', *Annual Review of Psychology* **35**, 113–138.

Meyer, B. (1984), Organizational aspects of text: effects on reading comprehension and applications for the classroom, *in* J. Flood, ed., 'Promoting Reading Comprehension', International Reading Association, Newark Del.

Meyer, B., Brandt, D. & Bluth, G. (1980), 'Use of top-level structure in text: key for reading comprehension of ninth-grade students', *Reading Research Quarterly* **16**, 72–103.

Miller, G. (1987), 'The influence of self-instruction on the comprehension monitoring performance of average and above average readers', *Journal of Reading Behaviour* **19**, 303–317.

Miller, G., Giovenco, A. & Rentiers, K. (1987), 'Fostering comprehension monitoring in below average readers through self-instruction training', *Journal of Reading Behaviour* **19**, 379–393.

Miller, R. (1980), 'Inference at comprehension: fact or artifact?', *Journal of Reading Behaviour* **12**, 279–297.

Mitchell, D. & Green, D. (1978), 'The effects of context and content on immediate processing in reading', *Quarterly Journal of Experimental Psychology* **30**, 609–636.

Miyake, N. & Norman, D. (1979), 'To ask a question, one must know enough to know what is not known', *Journal of Verbal Learning and Verbal Behaviour* **18**, 357–364.

Murphy, G. & Medin, D. (1985), 'The role of theories in conceptual coherence', *Psychological Review* **92**, 289–316.

Myers, M. & Paris, S. (1978), 'Children's metacognitive knowledge about reading', *Journal of Educational Psychology* **70**, 680–690.

Norusis, M. (1988), *SPSS-X Introductory Statistics Guide (for SPSS-X Release 3)*, SPSS Inc., Chicago.

Oakhill, J. (1984), 'Inferential and memory skills in children's comprehension of stories', *British Journal of Educational Psychology* **54**, 31–39.

Ohlhausen, M. & Roller, C. (1988), 'The operation of text structure and content schemata in isolation and in interaction', *Reading Research Quarterly* **23**, 70–88.

Olshavsky, J. (1976–7), 'Reading as problem-solving: an investigation of strategies', *Reading Research Quarterly* **12**, 654–674.

Palincsar, A. (1984), A quest for meaning from expository text: a teacher-guided journey, *in* G. Duffy, L. Roehler & J. Mason, eds, 'Comprehension Instruction: Perspectives and Suggestions', Longman, New York.

Palincsar, A. (1985), 'The unpacking of a multi-component, metacognitive training package', Paper: annual meeting of the American Educational Research Association.

Palincsar, A. (1986), 'Metacognitive strategy instruction', *Exceptional Children* **53**, 118–124.

Palincsar, A. (1987), 'Collaborating for collective learning of text comprehension', Paper: annual conference of the American Educational Research Association.

Palincsar, A. & Brown, A. (1984), 'Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities', *Cognition and Instruction* **1**, 117–175.

Palincsar, A. & Brown, A. (1986), 'Interactive teaching to promote independent learning from text', *The Reading Teacher* **39**, 771–777.

Palincsar, A., Brown, A. & Martin, S. (1987), 'Peer interaction in reading comprehension instruction', *Educational Psychologist* **22**, 231–253.

Paris, S. & Jacobs, J. (1984), 'The benefit of informed instruction for children's reading awareness and comprehension skills', *Child Development* **55**, 2083–2093.

Paris, S. & Lindauer, B. (1976), 'The role of inference in children's comprehension and memory for sentences', *Cognitive Psychology* **8**, 217–227.

Paris, S. & Myers, M. (1981), 'Comprehension monitoring, memory, and study strategies of good and poor readers', *Journal of Reading Behaviour* **13**, 5–22.

Paris, S. & Oka, E. (1986*a*), 'Children's reading strategies, metacognition and motivation', *Developmental Review* **6**, 25–56.

Paris, S. & Oka, E. (1986*b*), 'Self-regulated learning among exceptional children', *Exceptional children* **53**, 103–108.

Paris, S. & Upton, L. (1976), 'Children's memory for inferential relationships in prose', *Child Development* **47**, 660–668.

Paris, S., Cross, D. & Lipson, M. (1984), 'Informed strategies for learning: a program to improve children's reading awareness and comprehension', *Journal of Educational Psychology* **76**, 1239–1252.

Paris, S., Lindauer, K. & Cox, G. (1977), 'The development of inferential comprehension', *Child Development* **48**, 1728–1733.

Paris, S., Lipson, M. & Wixson, K. (1983), 'Becoming a strategic reader', *Contemporary Educational Psychology* **8**, 293–316.

Paris, S., Wasik, B. & van der Westhuizen, G. (1987), 'Meta-metacognition: a review of research on metacognition and reading', Paper: National Reading Conference, St Petersburg Florida.

Pearson, P. (1984), Direct explicit teaching of reading comprehension, *in* G. Duffy, L. Roehler & J. Mason, eds, 'Comprehension Instruction: Perspectives and Suggestions', Longman, New York.

Pearson, P. & Johnson, D. (1978), *Teaching Reading Comprehension*, Holt, Rinehart and Winston, New York.

Pearson, P., Hansen, J. & Gordon, C. (1979), 'The effect of background knowledge on young children's comprehension of explicit and implicit information', *Journal of Reading Behaviour* **11**, 200–209.

Perfetti, C. (1984), 'Reading acquisition and beyond: decoding includes cognition', *American Journal of Education* **93**, 40–59.

Perfetti, C. & Goldman, S. (1976), 'Discourse memory and reading comprehension skill', *Journal of Verbal Learning and Verbal Behaviour* **14**, 33–42.

Perfetti, C. & Hogaboam, T. (1975), 'Relationship between single word decoding and reading comprehension skill', *Journal of Educational Psychology* **67**, 461–469.

Perfetti, C. & Roth, S. (1981), Some of the interactive processes in reading and their role in reading skill, *in* A. Lesgold & C. Perfetti, eds, 'Interactive Processes in Reading', Erlbaum, Hillsdale N.J.

Perfetti, C., Finger, E. & Hogaboam, T. (1978), 'Sources of vocalization latency differences between skilled and less skilled young readers', *Journal of Educational Psychology* **70**, 730–739.

Perfetti, C., Goldman, S. & Hogaboam, T. (1979), 'Reading skill and the identification of words in context', *Memory and Cognition* **7**, 273–282.

Phifer, S. & Glover, J. (1982), 'Don't take students' word for what they do while reading', *Bulletin of the Psychonomic Society* **19**, 194–196.

Pichert, J. & Anderson, R. (1977), 'Taking different perspectives on a story', *Journal of Educational Psychology* **69**, 309–315.

Posner, M. & Synder, C. (1975), Attention and cognitive control, *in* R. Solso, ed., 'Information Processing and Cognition - the Loyola Symposium', Erlbaum, Hillsdale N.J.

Pressley, M., Ghatala, E., Woloshyn, V. & Pirie, J. (1990), 'Sometimes adults miss the main ideas and do not realize it: confidence in responses to short-answer and multiple-choice comprehension questions', *Reading Research Quarterly* **25**, 232–249.

Pressley, M., Johnson, C., Symons, S., McGoldrick, J. & Kurita, J. (1989), 'Strategies that improve children's memory and comprehension of text', *Elementary School Journal* **90**, 3–32.

Rabinowitz, M. (1988), 'On teaching cognitive strategies: the influence of accessibility of conceptual knowledge', *Contemporary Educational Psychology* **13**, 229–235.

Raphael, T. (1982), 'Question answering strategies for children', *The Reading Teacher* **November**, 186–190.

Raphael, T. (1984), 'Teaching learners about sources of information for answering comprehension questions', *Journal of Reading* **January**, 303–311.

Raphael, T. & McKinney, J. (1983), 'An examination of fifth- and eighth-grade children's question-answering behaviour: an instructional study in metacognition', *Journal of Reading Behaviour* **15**, 67–86.

Raphael, T. & Pearson, P. (1982), The effect of metacognitive training on children's question-answering behaviour, Technical Report 238, Centre for the Study of Reading, University of Illinois.

Raphael, T. & Wonnacott, C. (1985), 'Heightening fourth-grade students' sensitivity to sources of information for answering comprehension questions', *Reading Research Quarterly* **20**, 282–296.

Rayner, K. & Frazier, L. (1989), 'Selection mechanisms in reading lexically ambiguous words', *Journal of Experimental Psychology: Learning, Memory and Cognition* **15**, 779–790.

Rayner, K., Carlson, M. & Frazier, L. (1983), 'The interaction of syntax and semantics during sentence processing: eye movements in the analysis of semantically biased sentences', *Journal of Verbal Learning and Verbal Behaviour* **22**, 358–374.

Recht, D. & Leslie, L. (1988), 'Effect of prior knowledge on good and poor readers' memory of text', *Journal of Educational Psychology* **80**, 16–20.

Reid, M. & Borkowski, J. (1987), 'Causal attributions of hyperactive children: implications for teaching strategies and self-control', *Journal of Educational Psychology* **79**, 296–307.

Reis, R. & Spekman, N. (1983), 'The detection of reader-based versus text-based inconsistencies and the effects of direct training of comprehension monitoring among upper-grade poor comprehenders', *Journal of Reading Behaviour* **15**, 49–60.

Rinehart, S., Stahl, S. & Erickson, L. (1986), 'Some effects of summarization training on reading and studying', *Reading Research Quarterly* **21**, 422–438.

Roller, C. (1990), 'The interaction between knowledge and structure variables in the processing of expository prose', *Reading Research Quarterly* **25**, 29–39.

Rosch, E. (1973), On the internal structure of perceptual and semantic categories, *in* T. Moore, ed., 'Cognitive Development and the Acquisition of Language', Academic press, New York.

Rosch, E. & Mervis, C. (1975), 'Family resemblances: studies in the internal structure of categories', *Cognitive Psychology* **7**, 573–605.

Rosenshine, B. & Stevens, R. (1984), Classroom instruction in reading, *in* P. Pearson, ed., 'Handbook of Reading Research', Longman, New York.

Rumelhart, D. & McClelland, J. (1981), Interactive processing through spreading activation, *in* A. Lesgold & C. Perfetti, eds, 'Interactive Processes in Reading', Erlbaum, Hillsdale N.J.

Rumelhart, D. & Ortony, A. (1977), The representation of knowledge in memory, *in* R. Anderson, R. Spiro & W. Montague, eds, 'Schooling and the Acquisition of Knowledge', Erlbaum, Hillsdale N.J.

Salomon, G., Globerson, T. & Guterman, E. (1989), 'The computer as a zone of proximal development: internalizing reading-related metacognitions from a reading partner', *Journal of Educational Psychology* **81**, 620–627.

Sampson, M., Valmont, W. & van Allen, R. (1982), 'The effects of instructional cloze on the comprehension, vocabulary and divergent production of third-grade students', *Reading Research Quarterly* **17**, 389–399.

Samuels, S., Begy, G. & Chen, C. (1975–6), 'Comparison of word recognition speed and strategies of less skilled and more highly skilled readers', *Reading Research Quarterly* **11**, 72–86.

Samuels, S., LaBerge, D. & Bremer, C. (1978), 'Units of word recognition: evidence for developmental changes', *Journal of Verbal Learning and Verbal Behaviour* **17**, 715–720.

Samuels, S., Miller, N. & Eisenberg, P. (1979), 'Practice effects on the unit of word recognition', *Journal of Educational Psychology* **71**, 514–520.

Scardamalia, M. & Bereiter, C. (1984), Development of strategies of text processing, *in* H. Mandl, N. Stein & T. Trabasso, eds, 'Learning and Comprehension of Text', Erlbaum, Hillsdale N.J.

Schallert, D. (1976), 'Improving memory for prose: the relationship between depth of processing and context', *Journal of Verbal Learning and Verbal Behaviour* **15**, 621–632.

Schank, R. (1982), *Dynamic Memory: a Theory of Reminding and Learning in Computers and People*, Cambridge University Press, New York.

Schank, R. & Abelson, R. (1977), *Scripts, Plans, Goals and Understanding*, Erlbaum, Hillsdale N.J.

Schneider, W., Koerkel, J. & Weinert, F. (1989), 'Domain-specific knowledge and memory performance: a comparison of high- and low-aptitude children', *Journal of Educational Psychology* **81**, 306–312.

Schommer, M. & Surber, J. (1986), 'Comprehension-monitoring failure in skilled adult readers', *Journal of Educational Psychology* **78**, 353–357.

Schustack, M., Ehrlich, S. & Rayner, K. (1987), 'Local and global sources of contextual facilitation in reading', *Journal of Memory and Language* **26**, 322–340.

Schwarz, M. & Flammer, A. (1981), 'Text structure and title effects on comprehending and recall', *Journal of Verbal Learning and Verbal Behaviour* **20**, 61–66.

Seidenberg, M., Tanenhaus, M., Leiman, J. & Bienkowski, M. (1982), 'Automatic access of the meanings of ambiguous words in context: some limitations of knowledge-based processing', *Cognitive Psychology* **14**, 489–537.

Shanahan, T., Kamil, M. & Tobin, A. (1982), 'Cloze as a measure of intersentential comprehension', *Reading Research Quarterly* **17**, 229–255.

Short, E. & Ryan, E. (1984), 'Metacognitive differences between skilled and less skilled readers: remediating deficits through story grammar and attribution training', *Journal of Educational Psychology* **76**, 225–235.

Singer, H. (1978), 'Active comprehension: from answering to asking questions', *The Reading Teacher* **31**, 900–908.

Singer, M. (1979*a*), 'Processes of inference during sentence encoding', *Memory and Cognition* **7**, 192–200.

Singer, M. (1979*b*), 'Temporal locus of inference in the comprehension of brief passages: recognizing and verifying implications about instruments', *Perceptual and Motor Skills* **49**, 539–550.

Singer, M. & Ferreira, F. (1983), 'Infering consequences in story comprehension', *Journal of Verbal Learning and Verbal Behaviour* **22**, 437–448.

Singer, S. & Donlan, D. (1982), 'Active comprehension: problem-solving schema with question generation for comprehension of complex short stories', *Reading Research Quarterly* **17**, 166–186.

Smiley, S., Oakley, D., Worthen, D., Campione, J. & Brown, A. (1977), 'Recall of thematically relevant material by adolescent good and poor readers as a function of written versus oral presentation', *Journal of Educational Psychology* **69**, 381–387.

Spilich, G., Vesonder, G., Chiesi, H. & Voss, J. (1979), 'Text processing of domain-related information for individuals with high and low domain knowledge', *Journal of Verbal Learning and Verbal Behaviour* **18**, 275–290.

Stahl, S., Hare, V., Sinatra, R. & Gregory, J. (1991), 'Defining the role of prior knowledge and vocabulary in reading comprehension: the retiring of number 41', *Journal of Reading Behaviour* **23**, 487–508.

Stanovich, K. (1980), 'Toward an interactive-compensatory model of individual differences in the development of reading fluency', *Reading Research Quarterly* **16**, 32–71.

Stanovich, K. & West, R. (1983), 'On priming by a sentence context', *Journal of Experimental Psychology: General* **112**, 1–36.

Stanovich, K., West, R. & Feeman, D. (1981), 'A longitudinal study of sentence context effects in second-grade children: tests of an interactive-compensatory model', *Journal of Experimental Child Psychology* **32**, 185–199.

Stein, B., Bransford, J., Franks, J., Owings, R., Vye, N. & McGaw, W. (1982), 'Differences in the precision of self-generated elaborations', *Journal of Experimental Psychology: General* **111**, 399–405.

Stevens, R. (1988), 'Effects of strategy training on the identification of the main idea of expository passages', *Journal of Educational Psychology* **80**, 21–26.

Stevens, R., Slavin, R. & Farnish, A. (1991), 'The effects of cooperative learning and direct instruction in reading comprehension strategies on main idea identification', *Journal of Educational Psychology* **83**, 8–16.

Sullivan, O. & Pressley, M. (1984), 'Completeness of instruction and strategy transfer', *Journal of Experimental Child Psychology* **38**, 275–288.

Tabossi, P. (1988*a*), 'Accessing lexical ambiguity in different types of sentential contexts', *Journal of Memory and Language* **27**, 324–340.

Tabossi, P. (1988*b*), 'Effects of context on the immediate interpretation of unambiguous nouns', *Journal of Experimental Psychology: Learning, Memory and Cognition* **14**, 153–162.

Tabossi, P. & Laird, J.-P. (1980), 'Linguistic context and the priming of semantic information', *Quarterly Journal of Experimental Psychology* **32**, 595–603.

Taylor, B. (1980), 'Children's memory for expository text after reading', *Reading Research Quarterly* **15**, 399–411.

Taylor, B. (1982), 'Text structure and children's comprehension and memory for expository material', *Journal of Educational Psychology* **74**, 323–340.

Taylor, B. (1986), Teaching middle grade students to summarize content textbook material, *in* J. Baumann, ed., 'Teaching Main Idea Comprehension', International Reading Association, Newark Del.

Taylor, B. & Beach, R. (1984), 'The effects of text structure instruction on middle grade students' comprehension and production of expository text', *Reading Research Quarterly* **19**, 134–146.

Taylor, K. (1986), 'Summary writing by young children', *Reading Research Quarterly* **21**, 193–207.

Terry, P., Samuels, S. & LaBerge, D. (1976), 'The effects of letter degradation and letter spacing on word recognition', *Journal of Verbal Learning and Verbal Behaviour* **15**, 577–585.

Thorpe, K. & Satterly, D. (1990), 'The development and interrelationship of metacognitive components among primary school children', *Educational Psychology* **10**, 5–21.

Till, R., Mross, E. & Kintsch, W. (1988), 'Time course of priming for associate and inference words in a discourse context', *Memory and Cognition* **16**, 283–298.

Tunmer, W., Nesdale, A. & Pratt, C. (1983), 'The development of young children's awareness of logical inconsistencies', *Journal of Experimental Child Psychology* **36**, 97–108.

van Dijk, T. (1977), Semantic macrostructures and knowledge frames in discourse comprehension, *in* M. Just & P. Carpenter, eds, 'Cognitive Processes in Comprehension', Erlbaum, Hillsdale N.J.

van Dijk, T. & Kintsch, W. (1983), *Strategies of Discourse Comprehension*, Academic Press, New York.

Vosniadou, S., Pearson, D. & Rogers, T. (1988), 'What causes children's failure to detect inconsistencies in text? representation versus comparison difficulties', *Journal of Educational Psychology* **80**, 27–39.

Walczyk, J. (1990), 'Relations among error detection, sentence verification and low-level reading skills of fourth graders', *Journal of Educational Psychology* **82**, 491–497.

Walczyk, J. (1993), 'Are general resource notions still viable in reading research?', *Journal of Educational Psychology* **85**, 127–135.

Walczyk, J. & Hall, V. (1991), 'A developmental study of children's ability to adopt perspectives and find errors in text', *Contemporary Educational Psychology* **16**, 2–17.

Walczyk, J. & Raska, L. (1992), 'The relationship between low- and high-level reading skills in children', *Contemporary Educational Psychology* **17**, 38–46.

West, R. & Stanovich, K. (1978), 'Automatic contextual facilitation in readers of three ages', *Child Development* **49**, 717–727.

Williams, J. (1984), 'Categorization, macrostructure, and finding the main idea', *Journal of Educational Psychology* **76**, 874–879.

Williams, J., Taylor, M. & de J. Cani (1984), 'Constructing macrostructure for expository text', *Journal of Educational Psychology* **76**, 1065–1075.

Williams, J., Taylor, M. & Ganger, S. (1981), 'Text variations at the level of individual sentence and the comprehension of simple expository paragraphs', *Journal of Educational Psychology* **73**, 851–865.

Winograd, P. (1984), 'Strategic difficulties in summarizing texts', *Reading Research Quarterly* **19**, 404–425.

Winograd, P. & Bridge, C. (1986), The comprehension of important information in written prose, *in* J. Baumann, ed., 'Teaching Main Idea Comprehension', International Reading Association, Newark Del.

Winograd, P. & Johnston, P. (1982), 'Comprehension monitoring and the error detection paradigm', *Journal of Reading Behaviour* **14**, 61–75.

Wixson, S. P. K. & Palincsar, A. (1986), 'Instructional approaches to reading comprehension', *Review of Research in Education* **13**, 91–128.

Wong, B. (1985), 'Self-questioning instructional research: a review', *Review of Educational Research* **55**, 227–268.

Wong, B. & Jones, W. (1982), 'Increasing metacomprehension in learning disabled and normally achieving students through self-questioning training', *Learning Disability Quarterly* **5**, 228–240.

Wong, B. & Wong, R. (1986), 'Study behaviour as a function of metacognitive knowledge about critical task variables: an investigation of above average, average, and learning disabled readers', *Learning Disability Research* **1**, 101–111.

Wong, B., Wong, R., Perry, N. & Sawatsky, D. (1986), 'The efficacy of a self-questioning summarization strategy for use by underachievers and learning disabled adolescents in social studies', *Learning Disability Focus* **2**, 20–35.

Yuill, N. & Oakhill, J. (1988), 'Understanding of anaphoric relations in skilled and less skilled comprehenders', *British Journal of Psychology* **79**, 173–186.

Yussen, S. & Smith, C. (1990), 'Detecting general and specific errors in expository texts', *Contemporary Educational Psychology* **15**, 224–240.

Zabrucky, K. & Commander, N. (1993), 'Rereading to understand: the role of text coherence and reader proficiency', *Contemporary Educational Psychology* **18**, 442–454.

Zabrucky, K. & Moore, D. (1989), 'Children's ability to use three standards to evaluate their comprehension of text', *Reading Research Quarterly* **24**, 336–354.

Zabrucky, K. & Ratner, H. (1986), 'Children's comprehension monitoring and recall of inconsistent stories', *Child Development* **57**, 1401–1418.