

**THE ROLE OF ILLUSTRATIONS
IN
CHILDREN'S ORAL READING ACCURACY,
STRATEGIES AND COMPREHENSION
AT
DIFFERENT DEVELOPMENTAL AND
PROGRESS LEVELS :
A PSYCHOLINGUISTIC INVESTIGATION**

DAVID R. DONALD MA (Natal) M.Ed (Exeter)

**Thesis submitted in partial fulfilment of the requirements for
the degree of Doctor of Philosophy :
Department of Psychology, Faculty of Social Science,
University of Cape Town.**

**Rondebosch,
South Africa.
September, 1981.**

The University of Cape Town has been given
the right to reproduce this thesis in whole
or in part. Copyright is held by the author.

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

"The great value of showing illustrative pictures with sentences, aside from the interest aroused, is in their making the child feel the sentence's meaning as he reads it"

E.B. HUEY, 1908: The Psychology and Pedagogy of Reading, (p.322; Huey's emphasis)

ACKNOWLEDGEMENTS

The author would like to express his appreciation of the help, time and advice given by his supervisor, Dr. Richard Oxtoby as well as that so generously and freely offered by Gillian Finchilescu and Lester Gilbert of the Psychology Department, University of Cape Town.

Perhaps the greatest sufferers through the production of a Doctoral Thesis are members of the author's family.

In very sincere appreciation of this, the author would like to express his gratitude for the patience, support and forbearance of his wife and three daughters through this extended period.

There are many others without whose willing co-operation the research and production of the thesis would have been impossible. To all these, not least the teachers and children involved in the research, the author would like to express his thanks.

Finally, the author would like to thank the University of Cape Town and the Human Sciences Research Council for financial assistance.

ABSTRACT

The effects of illustrations on early reading development have been subject to considerable controversy. Results and interpretations under the 'focal attention hypothesis' indicate that illustrations have a distracting effect on the learning of responses to orthographic cues in the process of isolated word recognition. Conversely, considerable although inconclusive evidence suggests that illustrations may be facilitative as contextual information in the process of reading and comprehending continuous prose. Within a psycholinguistic model of the reading process, the contextual hypothesis, that illustrations constitute a source of contextual redundancy which facilitates word identification accuracy, strategy and comprehension, was tested. Given the results of an earlier experiment that had confirmed the hypothesis for seven-year-old, average readers, the aim was to test the hypothesis over high and low progress readers at reading ages seven and nine. From 1868 grades I, III and V children screened on the D. Young Group Reading Test, 120 subjects at the respective reading age and progress levels were selected. Within a matched samples, 2 x 2 x 2 factorial design, subjects read 320 word narrative stories at instructional level of difficulty, with or without illustrations. Results in general confirmed the hypothesis. In particular, analysis of variance revealed that the illustration effect was strong and significant for RA7, high progress and for RA9, low progress readers; moderate and significant for RA7, low progress readers; and consistent but generally non-significant for RA9, high progress readers. This significant interactive pattern held over word identification accuracy; literal comprehension; use of semantic information (error acceptability) and rate of self-correction. Use of syntactic information was moderately and significantly facilitated across combined groups. Use of orthographic information, as predicted, was moderately and significantly reduced across combined groups. Inferential comprehension was non-significantly affected. It was concluded

that, in the process of contextual reading, illustrations facilitate access to meaning; that the strength of the effect depends on the need for extra-textual contextual information and processing capabilities of the respective groups; and that the 'focal attention' effect on isolated word recognition is a particular processing case within the more general, practically relevant case of contextual reading.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	ii
ABSTRACT	iii
TABLE OF CONTENTS	v
LIST OF TABLES	x
LIST OF FIGURES	xv
SUMMARY	xvii

PART A: THEORETICAL DEVELOPMENT
and RESEARCH REVIEW

	PAGE
I <u>Introduction and Overview</u>	1
1. The Context	1
2. The Issues	3
3. Aims	16
II <u>The Process of Reading: Reading as Synthesis</u>	20
1. The Traditional View	21
2. Reading as Synthesis	24
3. Summary	34
III <u>Reading Development: Models of the Trans- ition to Fluent Reading</u>	37
1. Goodman's (1968) Model	37
2. Smith's (1971) Model	40
3. An Accelerated Progression Model	44
4. Implications of the Model.	51
IV <u>Evaluating Reading Development</u>	56
Theoretical justification of dependent variables and their relevance in evaluating the effects of	

	illustrations under the contextual hypothesis.	
	1. Accuracy	61
	2. Strategy	63
	i) Semantic Acceptability of Errors	71
	ii) Syntactic Acceptability of Errors	78
	iii) Orthographic Acceptability of Errors	85
	iv) Self-Corrections.	90
	3. Comprehension	98
	i) Literal Comprehension	102
	ii) Inferential Comprehension	105
V	<u>Illustrations and Reading: Research Review</u>	107
	1. The Focal Attention Hypothesis	107
	2. The Contextual Hypothesis	124
	3. Semantic Access: Word and Image Interaction	144
	4. Theoretical Derivations/	150
VI	<u>Experiment I: The effects of illustrations on the oral reading accuracy, strategies and comprehension of second-year, average readers: a test of the contextual hypothesis.</u>	156
	1. Method	157
	2. Results	164
	3. Discussion	166
	4. Limitations and Research Implications.	171
VII	<u>Theoretical Derivation of the Independent Variables: Developmental Level and Progress Level.</u>	175
	1. Developmental Level	176

2.	Progress Level	182
3.	Interaction	187

PART B: EXPERIMENT II

The effects of illustrations on the accuracy and efficiency of oral reading message identification and on the comprehension of high and low progress readers at seven and nine year reading age developmental levels.

VIII	<u>Aims and Experimental Design</u>	190
	1. Aims	190
	2. Design	191
IX	<u>Method</u>	200
	1. Subjects and Sampling Procedure	200
	i) Population Considerations	200
	ii) Screening	205
	iii) Sample Selection	207
	2. Materials	213
	3. Experimental Procedure	227
	4. Transcription of Oral Reading and Comprehension Responses	229
	5. Scoring of Dependent Variables	233
	i) Word Identification Accuracy	233
	ii) Semantic Acceptability of Errors	233
	iii) Syntactic Acceptability of Errors	238
	iv) Orthographic Acceptability of Errors	244
	v) Self-Corrections	246

	vi) Literal Comprehension.	247
	vii) Inferential Comprehension	252
	6. Reliability	254
X	<u>Results and Interpretation</u>	259
	1. Analyses of Variance	261
	i) Word Identification Accuracy	261
	ii) Semantic Acceptability of Errors	271
	iii) Syntactic Acceptability of Errors	281
	iv) Orthographic Acceptability of Errors	287
	v) Self-Corrections	293
	vi) Literal Comprehension	304
	vii) Inferential Comprehension	315
	viii) Summary of the Illustration Effect over Analysis of Variance Results	320
	2. Correlations	322
XI	<u>Conclusions</u>	333
	1. Accuracy	333
	2. Strategy	335
	3. Comprehension	348
	4. General Conclusions	352
XII	<u>Critique and Research Implications</u>	366
	1. Critique	366
	2. Research Implications	368
XIII	<u>Theoretical and Educational Implications</u>	375
	1. Theoretical Implications	375
	2. Educational Implications	378

LIST OF TABLES

<u>TABLE</u>	<u>PAGE</u>
1. Cell means for Word Identification Accuracy (Experiment I)	164
2. Weightings for the theoretically expected effects of illustrations by developmental and progress levels.	187
3. Hypothetical 2 x 2 x 2 factor design with illustration condition, progress level and developmental level as treatments.	192
4. Actual 2 x 2 x 2 factor design with illustration condition, progress level and reading age developmental level as treatments.	196
5. Socio-economic catchment of schools used in sample selection.	204
6. Distribution by sex and grade level for the total screened population.	207
7. Socio-economic scale used for sample matching, across illustration conditions.	210
8. Final cell samples by socio-economic rating, grade level, chronological age and sex.	211
9. Sample used for pre-testing experimental reading material.	216
10. Results of pre-testing experimental reading material for accuracy and comprehension.	217

<u>TABLE</u>	<u>PAGE</u>
11. Ratios of illustration to number of words in experimental reading material.	223
12. Number of words and illustrations per story in experimental reading material.	225
13. Transcription and coding system for oral reading errors.	230
14. Scale and criteria for scoring semantic acceptability of errors.	234
15. Scale and criteria for scoring syntactic acceptability of errors.	239
16. Scale and criteria for scoring orthographic acceptability of errors.	244
17. Scale and criteria for scoring literal comprehension: detail	248
18. Scale and criteria for scoring literal comprehension: cause-effect.	249
19. Scale and criteria for scoring literal comprehension: main idea.	250
20. Scale and criteria for scoring literal comprehension: related ideas.	251
21. Scale and criteria for scoring inferential comprehension.	252
22. Distribution within design cells of randomly selected reliability sample.	255

<u>TABLE</u>	<u>PAGE</u>
23a. Cell means/standard deviations, cell mean differences (P1-P0) and treatment means: word identification accuracy	261
b. Analysis of Variance summary table: word identification accuracy.	263
c. Simple interaction effects: word identification accuracy.	264
d. Simple-simple main effects: word identification accuracy.	265
e. Simple main effects: word identification accuracy.	266
24a. Cell means/standard deviations, cell mean differences (P1-P0) and treatment means: semantic acceptability	272
b. Analysis of Variance summary table: semantic acceptability	274
c. Simple interaction effects: semantic acceptability.	275
d. Simple-simple main effects: semantic acceptability.	276
e. Simple main effects: semantic acceptability.	277
25a. Cell means/standard deviations, cell mean differences (P1-P0) and treatment means: syntactic acceptability.	282
b. Analysis of Variance summary table: syntactic acceptability.	284
26. Comparison of cell mean differences (P1-P0) and error terms across semantic and syntactic acceptabilities of errors.	285

<u>TABLE</u>	<u>PAGE</u>
27a. Cell means/standard deviations, cell mean differences (P1-P0) and treatment means: orthographic acceptability.	288
b. Analysis of Variance summary table: orthographic acceptability.	290
28a. Cell means/standard deviations, cell mean differences (P1-P0) and treatment means: self-corrections.	293
b. Analysis of Variance summary table: self-corrections.	295
c. Simple interaction effects: self-corrections.	296
d. Simple-simple main effects: self-corrections.	297
e. Simple main effects: self-corrections.	298
29a. Cell means/standard deviations, cell mean differences (P1-P0) and treatment means: literal comprehension.	304
b. Analysis of Variance summary table: literal comprehension.	306
c. Simple interaction effects: literal comprehension.	307
d. Simple-simple main effects: literal comprehension.	308
e. Simple main effects: literal comprehension.	309
30a. Cell means/standard deviations, cell mean differences (P1-P0) and treatment means: inferential comprehension.	315
b. Analysis of Variance summary table: inferential comprehension.	317
c. Simple main effects: inferential comprehension.	318

31. Summary table of direction, significance and relative strength of illustration (P) main effects over all dependent variables. 320
32. Correlation matrix: All dependent variables over all conditions (N = 120). 322
33. Weightings for average obtained illustration effects over grade, progress and reading age developmental levels for the variables of word identification accuracy, semantic acceptability, self-correction and literal comprehension. 360

LIST OF FIGURES

<u>FIGURE</u>		<u>PAGE</u>
1a	Conceptual illustration of the accelerated progression of reading development.	49
b	Conceptual illustration of the initial stages of reading development.	50
2.	Selection of high and low progress groups at two reading age levels across hypothetical frequency distributions of reading age for grades I, III and V.	195
3.	Socio-economic distribution by geographical area of the population of Cape Town: Position and approximate catchment of schools used.	203
4.	Colour example of illustrations used.	221
5.	Graph of cell mean profiles: word identification accuracy.	263
6.	Graph of cell mean profiles: semantic acceptability.	273
7.	Graph of cell mean profiles: syntactic acceptability.	283
8.	Graph of cell mean profiles: orthographic acceptability.	289
9.	Graph of cell mean profiles: self-corrections.	295
10.	Graph of cell mean profiles: literal comprehension.	306

FIGUREPAGE

11. Graph of cell mean profiles: inferential comprehension. 317
12. Graph of P1 and P0 treatment means for orthographic and semantic acceptabilities. 340
13. Graphs of P1 and P0 cell means for orthographic and semantic acceptabilities at 7H; 7L; 9H; and 9L.
14. Extrapolated developmental progression for the relative influence of illustrations on the reading development of high, average, and low progress readers. 362

SUMMARY

Within the controversy surrounding the role of illustrations in early reading development, contradictions appear to exist between results and theoretical explanations under the focal attention hypothesis on the one hand (Samuels, 1967; 1970) and a contextual hypothesis on the other (Goodman, 1965; Smith, 1971).

The focal attention hypothesis argues that illustrations constitute attentional competition to orthographic stimuli and that the effects of illustrations are therefore negative within the process of reading development. The alternate, contextual hypothesis holds that illustrations constitute a source of contextual information that is redundant within, particularly, the semantic dimension of information: the process of message identification and comprehension is therefore facilitated through increasing the reader's use of contextual information while reducing, through the operation of redundancy, his reliance on orthographic information. Thus the focal attention hypothesis holds that reading development is dependent on attention to orthographic stimuli and the learning of the related orthographic responses. The paradigm developed to test the hypothesis has therefore relied on paired associate learning and the measure of successful reading development has been taken as the ability to recognise isolated words. By contrast, the contextual hypothesis holds that reading development is dependent on learning to process information from the semantic, the syntactic and the orthographic dimensions

of information and that the efficiency of the process relies on learning to select, within the redundancy of information available, only that which is necessary to uncertainty reduction and message identification. The only appropriate paradigm within this view of reading development is one in which the task, and the measures of successful reading development, are based on the reading of meaningful, continuous text.

The main argument of the thesis is that the apparent contradictions in the area lie in these two respective views of the nature of reading and its development and that the former view is but a special case within the more general and representative case of the normal process of reading development that the latter view espouses. Review of the relevant research as well as the results of the two experiments reported in the thesis tend to confirm this argument.

Since the contextual hypothesis had been less specifically tested than the focal attention hypothesis, the first experiment, as a previous study (Donald, 1979a) had been designed to test the effects of illustrations on the oral reading accuracy, strategies and comprehension - specifically on the reading of continuous narrative text - of seven year old, average readers. The results had confirmed that contextually relevant illustrations had a facilitative effect on word identification accuracy; that while use of semantic and syntactic information increased with illustrations, reliance on orthographic information decreased;

that the strategy of self-correction was facilitated with illustrations; and that while comprehension at the level of idea recall was facilitated, comprehension at the level of inference was not significantly affected. Experiment II, the central research study in the thesis, was designed not only to replicate the test of the contextual hypothesis under more rigorous and representative conditions but also to test the effects of illustrations for high and low progress readers at the reading age developmental levels of seven and nine. Since previous research contained the suggestion that an interactive relationship might exist between these groups, where the relative effects of illustrations were concerned, it was important to investigate this possibility.

The results of the second experiment confirmed the contextual hypothesis in general. In particular, analysis of variance revealed that for word identification accuracy, use of semantic information, rate of self-correction and literal comprehension, the illustration effect was strong and significant for high progress readers at reading age seven as well as for low progress readers at reading age nine; that it was moderate and significant for low progress readers at reading age seven; and that it was consistent but generally marginal and non-significant for high progress readers at reading age nine. The similar pattern of results on these particular variables was attributed to the common illustration effect of facilitating access to the semantic information that is central to all these aspects of reading. The differential effect on

the various groups was attributed to different patterns of need and use of illustration as extra-textual information in the process of gaining access to semantic information. For use of syntactic information the effect was moderate, significant and undifferentiated across groups. For use of orthographic information the effect was moderate, significant and, as predicted under the contextual hypothesis, negative. Across combined groups, a trend, however, suggested that the effect was considerably stronger for high progress readers at reading age seven than for other groups. For inferential comprehension the effect was marginal and non-significant across combined groups.

The most essential conclusions drawn are that the contextual hypothesis is confirmed in these results; that, given the relevant processing differences, the focal attention results are not inconsistent but merely irrelevant within these results; and that generalizations about the role of illustrations in early reading development must necessarily be qualified by differences in the way that illustrations as information appear to be used by readers at different progress and developmental levels. In particular, it would appear that readers who are able to optimise textual linguistic information (high progress readers at reading age nine) are able, relatively, to dispense with the extra-textual semantic context available in illustrations. Conversely, both high progress readers at reading age seven and low progress readers at reading age nine appear to benefit substantially from the extra-textual

context provided in illustrations to help them access the semantic information that is central to accurate and efficient (strategy) word identification as well as comprehension. The similar but more moderate effect for low progress readers at reading age seven would appear to indicate that these readers may need extra-textual support but that they are relatively less able to integrate the available information at their level of processing. Implications for further research, particularly where low progress readers are concerned, are developed from these conclusions. Theoretical implications on the question of reading development and the process of transition from initial to fluent reading are also evident. In educational terms the implications are significant not only with reference to the relative didactic and remedial value of illustrations for the respective developmental and progress levels but also with reference to the relative role that extra-textual contextual information - of whatever sort - may play in the development of reading.

1

PART A

THEORETICAL DEVELOPMENT AND RESEARCH REVIEW

CHAPTER 1

INTRODUCTION AND OVERVIEW

1. The Context

The reading materials on which children, today, are learning to read are more extensively illustrated than at any period in the history of the teaching of reading. Although the reason for this is probably mainly due to technological development in the printing process, there tends, behind it, to be the unquestioned assumption on the part of most teachers and publishers that the presence of illustrations is generally beneficial to the learners. Claims such as,

"Care has been taken to ensure a close harmony between text and illustration to achieve maximum help for the children"

which appears in a recently published reading scheme (Ginn, Reading 360, 1978), are ostensibly laudable and likely to be accepted without question by most teachers. Yet just what form of help illustrations provide and just how and where this enters into the process of learning to read are real and unanswered questions: questions that have both theoretical and practical import.

The role that illustrations might play in the early reading

process has been separated by several authors into:

a) a motivational role, and b) a cognitive role (Miller, 1938; Samuels, 1970; Gibson and Levin, 1975; Goldstein and Underwood, 1981).

The former, although not specifically researched, applies to the likely function that illustrations have in enlivening the presentation of text that, were it not illustrated, would appear dull - if not formidable - to young readers. Factors that might have a bearing on this role could be the size, frequency and lay-out of illustrations together with perhaps aesthetic considerations such as the use of colour, design, composition etc. It is somewhat artificial to distinguish these factors as 'motivational' from the content of illustrations that carry the more clearly 'cognitive' aspect of meaning: there must, inevitably, be an interaction between what a picture represents and how it is presented. Nevertheless, in so far as the mere presence of bright and attractive illustrations, whatever their content, may be expected to motivate a child to pick up a book and to view it as less daunting than pages of uninterrupted print, the distinction can probably be upheld. In support of this, Miller points out that

"... bright pictures make a book attractive to children who are beginning to read, and studies of children's choice of books verify this statement" (Miller, 1938, p.676).

Similarly Samuels maintains that

"One argument for including illustrations with basal readers is that attractive pictures

may help a child develop positive attitudes towards reading. Learning to read is a difficult task for many children, and it is possible that attractive pictures which accompany text may make the task of learning to read a bit more pleasant" (Samuels, 1970, p.405).

As Samuels himself point out, this contention has not been directly verified through research and it must therefore remain at the level of an hypothesis. Observations of the degree of interaction with, and patent enjoyment of, pictures by young readers, however, can leave little doubt that it is a very 'likely' hypothesis.

The cognitive role that illustrations might play in the development of reading is, by contrast, a more contentious question. A considerable body of research has been generated in the area (reviews: Samuels, 1970; Golstein and Underwood, 1981). Despite this, firm conclusions remain illusive. The problem, as is contended in this thesis, is that the issues, both theoretical and practical, have not been adequately clarified. Consequently, individual findings and their interpretations have not been evaluated within a consistent or rationalized framework.

2. The Issues

Within the range of research, the basic question being asked has been: Do illustrations help, hinder or have no effect in the cognitive task of learning to read? The answers to this question have not been clear - as will emerge from

the research review in Chapter V. The first part of the problem is that the issue has been seen to require a single question rather than a series of more specific questions. From the research on illustrations as well as from reading development research, it would appear that at least three factors may be instrumental in determining the relative influence of illustrations on reading. Failure to identify these factors clearly and to avoid generalizations based on particular sets of factors has led to apparent contradictions and inconclusive answers to the basic question. It is the contention of this thesis that there is no categorical answer to the basic question: there are only specific answers to a series of more specific questions. These questions can only be asked within the framework of interaction of those factors that enter into the complex pattern of reading development.

The age of readers, or more specifically the developmental level they have reached in learning to read, is the first factor that needs to be taken into account. Sufficient evidence exists from the general field of cognitive development (Flavell, 1977) and from the particular field of reading development (Goodman, 1968; Biemiller, 1970; Burke, 1976; Donald, 1980a) to suggest that childrens' processing strategies, particularly during the early years of reading, are likely to vary with development. It cannot be assumed therefore that the information available in illustrated text will be processed in the same way by children at different developmental levels. Yet this is precisely what has been

done. Samuels (1970), for instance, in reviewing a wide range of research studies, draws conclusions and makes processing interpretations on the basis of studies where the age of subjects has varied from 17-year olds (Vernon, 1953) to pre-schoolers (Samuels, 1967).

The second factor is the relative competence of the readers within their age level. Differences between good and poor readers have been identified on a variety of criteria (Gibson and Levin, 1975). More specifically, differences have been noted in the processing strategies of good and poor readers at various levels of development (Clay, 1972; Willows, 1974; Clay and Imlach, 1971; Samuels et al, 1976; Burke, 1977; Rusted and Coltheart, 1979; Thomson, 1979; Potter, 1980). Once again, the evidence suggests that these differences will be reflected in differential processing of the information available in illustrated text.

Third is the relationship between illustration and text. Where the interaction of two sources of information is being studied, it is clear that how the relationship is structured must constitute an important variable. Several studies, in fact, have deliberately varied the relationship between illustration and text (Dwyer, 1970; Ketcham and Heath, 1972; Bransford and Johnson, 1972; Peeck, 1974). What emerges from these studies is that the effects of illustrations vary with such factors as the degree of relevance of the illustration to the text content, the point at which the illustration appears in the textual sequence and the nature of the information required in the

task set for the readers. Clearly the relative effects of illustrations is something that can only be evaluated in the context of an informational relationship between text and illustration. This need not imply that all generalizations on this question are invalid. However it does imply that different orders of relationship may have very different effects. Some of the apparent contradictions in this area of research, for instance, would appear to be attributable to the difference in the order of relationship between an illustration and an isolated word on the one hand and illustrations and continuous text on the other. Equally, the way the relationship between illustration and text is structured in any one study needs to be critically examined before generalizations based on the results are accepted at face value.

The second problem in deriving clarity from the research evidence to date has related to differences in theoretical position and the related differences in views of the reading process itself and of how illustrations might influence that process. A controversy, central to the whole issue, has developed around the 'focal attention hypothesis' (Samuels, 1967; 1970) on the one hand and a contextual hypothesis (Goodman, 1965; Smith, 1971) on the other.

Briefly, the former is based on a conception of limited attention capacity (developed in La Berge and Samuels, 1974) such that learning is seen to be influenced by the degree of attentional competition to whatever constitutes the

target stimulus. In the context of illustrations and reading, illustration, on the principle of 'least effort' (Samuels, 1970) are seen as distracting attention from the more difficult orthographic cues¹. In the sense that word recognition is seen as necessarily based on accurate orthographic analysis and identification, illustrations are therefore held to constitute attentional competition in the learning of appropriate responses to target orthographic cues.

A considerable number of studies and articles have been generated specifically within this theoretical framework (Samuels, 1967; 1970; 1977; Singer, Samuels and Spirof, 1973-74; Harzem, Lee and Miles, 1976; Montare, Elman and Cohen, 1977; Arlin, Scott and Webster, 1978; Willows, 1978a; b; Singer, 1980). These, and other studies not specifically generated by the hypothesis but which can be adduced as evidence within the theoretical framework, will be reviewed in Chapter V. For the present, it is sufficient to indicate that there is considerable although not unchallenged, support for the hypothesis. Validity of the interpretation, however, must necessarily take into account the three factors discussed above as well as the central question of task definition. What is important at this point is to identify the processing assumptions that

¹ Throughout this thesis the term 'orthographic' will be used in preference to the more commonly used term 'graphic'. Not only are illustrations also 'graphic' which may lead to confusion, but the term 'orthographic' refers more properly to the structure of graphemes in written language and conceptually includes the regularities that govern their relationship one to another (Gibson and Levin, 1975).

underpin the theoretical position and which have necessarily determined the experimental paradigm that has been developed as most appropriate for testing the hypothesis.

First, since orthographic cues are regarded as the attentional 'target' for effective learning, reading development is necessarily seen as a process of acquiring accurate responses to those orthographic cues. The criteria of effective reading development therefore become limited to the speed and accuracy with which individual words can be recognised on the basis of orthographic cues alone. Although there has been considerable variation in the structure of learning trials, all studies within this theoretical framework have limited the test trial to the recognition of target words in isolation. In other words the assumption has been that one learns to read through learning to recognise, accurately, the orthographic structure of individual words. Whether this view of reading development can be held as representative of the normal process of learning to read will constitute much of the argument of this thesis. That illustrations may interfere with learning accurate responses to isolated words may be a legitimate conclusion in its own right. As will be argued, however, reading development in its wider and more normal sense involves the reader in learning to use a far more complex process of informational interaction. Whether it is legitimate to equate this process with the more restricted process of learning to recognise isolated words, and to generalize the effects of illustrations in the one context to the other is, therefore, questionable.

Second, the hypothesis is based on the idea that attentional competition occurs between the pictorial cue on the one hand and the orthographic cues within the word on the other. According to the hypothesis, this competition is resolved, inefficiently for learning, through the principle of least effort such that the more easily accessible pictorial cue is attended at the cost of the orthographic cues. This conception clearly assumes an 'either-or' attentional function. What is significant is that the paired-associate paradigm, that has most consistently been used to test the hypothesis, necessitates such a dichotomised attentional function. Whether such an attentional function is characteristic of the normal reading process under natural reading conditions is, however, dubious. It will be argued, for instance, that in continuous, illustrated text there is a variety of sources of information available to the reader. Moreover those sources, including information in illustrations, overlap and are redundant within one another. The demands on attention under these conditions cannot be seen as simply dichotomised. Where there is overlap or redundancy of information, the most likely attentional function, even given the conception of limited attention capacity, is an integrative one. In other words, attention will be focussed most naturally on clusters of informational features that, in their redundancy, most reliably and efficiently predict the on-going textual message (Smith, 1971). Thus, whether the results of the paired-associate studies support the hypothesis or not, it is at least questionable whether they or the attentional

interpretation have relevance for understanding the process of reading development in the more normal context of learning to read continuous text.

Third, the paired-associate paradigm necessitates an informational relationship that is limited to the specific cueing function of a particular picture in relation to a particular word. Under these conditions the picture is semantically non-specific (any isolated picture can elicit a variety of verbal labels) while the word, by contrast, is highly specific. Thus, as the hypothesis predicts, attention to the cueing function of the picture is likely to yield semantically non-specific information which is inefficient in that it distracts from the only other, and more reliable, source of information: the orthographic cues. In other words, within the paradigm, the assumption is that pictures necessarily constitute an unreliable source of information and the informational relationship is structured so that they will necessarily be used as such. By contrast, where a series of illustrations are set in relation to continuous text, an informational relationship of another order is set up. The relationship is more freely structured so that the illustrations may suggest a general context and expectation for certain sorts of event to follow rather than only having a specific cueing function for particular words. Moreover, since cumulative semantic and syntactic information is also available to the reader from the text itself (i.e., he is not limited to orthographic information as the only alternative), it follows that an

interaction is possible between contextual information available through the text and through the illustrations. Under these conditions interpretation of illustrations is likely to be far more specific and relevant to the text. Most significantly, however, illustrations under these conditions constitute only one of several sources of overlapping information. This not only increases the reliability of the information itself but it also releases the reader from the constraint, in the paired-associate paradigm, of having to rely unduly on pictorial information. Thus, both the information value of an illustration, as well as the weight of the information in the processes involved, is radically different for the two task conditions and it is doubtful if conclusions from one context can be applied to the other.

The contextual hypothesis can be traced most directly to a psycholinguistic view of the reading process. Since this view and its implications for a theory of reading development are regarded as central to understanding the role of illustrations it will be fully discussed in subsequent chapters. In essence, however, the psycholinguistic view holds that reading is a process of predicting and confirming the textual message on the basis of a variety of overlapping sources of information, the most essential of which are the orthographic, the syntactic and the semantic (Goodman, 1967; Smith, 1971). Within this, the contextual hypothesis holds that the more contextual information (essentially syntactic and semantic) the reader has access to, the more he can integrate this with only selected

orthographic information in order to achieve efficient and accurate word identification. Thus, Goodman(1965) and Potter (1980) for instance, have confirmed a common teacher's observation that children can identify words in context more accurately than those same words in isolation where a total reliance on orthographic information is necessary. Likewise, the tendency of readers to use orthographic information in inverse proportion to their use of contextual information has emerged through error analyses (Tulving and Gold, 1963; Weber, 1970b; Donald, 1979a). Applied to illustrations and reading, the hypothesis maintains that illustrations, linked to continuous text, provide an enriched contextual set for the reader. In other words, illustrations constitute a source of contextual information that overlaps with, particularly, the semantic source of information in the text such that prediction and confirmation of the textual message may be enhanced through a total contextual enrichment. It is important to stress that orthographic information is not unattended in this conception of information usage. As stated earlier, it is not an 'either-or' construct but rather an integrative one such that information (both contextual and orthographic) is selected on an efficiency basis according to the needs of prediction and confirmation (Denberg, 1976-77). Since reading is undeniably a linguistic activity, however, and concerned with generating meaning, the more contextual information the reader has access to the more efficient can his selection of orthographic cues be.

The underlying process assumptions in this theoretical framework are essentially incompatible with those in the focal attention framework. As a direct consequence of this, the experimental paradigm - including task definition, the selection of dependent variables and the structuring of the materials - developed to test the focal attention hypothesis is regarded as inappropriate to test the effects of illustrations under the contextual hypothesis. As will be argued, the two hypotheses should not be seen, as they have tended to be, as competing explanations of a single process but rather as two different explanations of two different processes. The question of which task is 'pragmatically' the most valid (Oller, 1979) and which process is the most representative of normal reading development is, however, highly relevant as this must determine the didactic implications of the conclusions. As will be developed more fully in subsequent chapters, the first assumption is that reading development is primarily a process of developing and refining strategies of information processing rather than primarily a process of acquiring accurate responses to orthographic cues. This is not regarded as absolute but rather as a matter of emphasis. Learning accurate responses to orthographic cues is viewed as a necessary but not a sufficient condition in the process of learning to read. In the light of this view, therefore, evaluation of the effectiveness of reading development cannot be made on the basis of isolated word recognition where the only information available is ortho-

graphic and the reader's competence in selecting and integrating alternative sources of information is precluded. Conversely, the alternative theoretical framework demands that the effectiveness of reading development can only be evaluated in terms of the reader's competence in selecting and integrating from those sources of information that are available under normal reading conditions. In this framework, the accuracy of the reader's word identification, in context, as well as his strategies for word identification as revealed through error patterns, are regarded as forms of reading behaviour that legitimately reflect the effectiveness of the development of the reader's total information processing strategy. Moreover, since comprehension is the ultimate goal and outcome of reading, the effectiveness of reading development cannot be evaluated without reference to this variable as well.

Second, since reading is regarded as an essentially linguistic process - a process where information from a variety of sources is available for selection, prediction and confirmation of a linguistically cumulative message - it follows that any experimental paradigm that reduces the linguistic information available under normal reading conditions must also be modifying the very process involved. In other words, the 'reading' that takes place where only isolated words are available may be a legitimate cognitive process to analyse but interpretations based on behaviour under such task conditions cannot be related to the normal reading of continuous, linguistically constrained text.

The processes are simply different. Where the object is to clarify the role of illustrations in the normal process of reading development, therefore, there is no alternative but to locate such an analysis in a paradigm that reflects that process. Thus, an essential assumption in this theoretical framework is that the experimental task requires the reading of normal, continuous, linguistically constrained text such as may be found in the books on which children learn to read.

Third, as has repeatedly been emphasised, the theoretical framework presupposes the normal availability of contextual information (information other than only orthographic) to the reader. Furthermore, it is assumed that illustrations normally provide a source of contextual information that is relevant to, and therefore redundant within, the content and structure of the text. The question at issue then is whether readers are helped, hindered or not affected by the availability of this information. Where either the nature of the illustration or its relation to the text preclude a meaningful contextual interaction, however, the question cannot legitimately be asked. This need not imply that other questions are not legitimate. The roles of a picture paired with an isolated word or of an irrelevant picture in relation to continuous text are legitimate questions but they are not the same question that is at issue here. Thus, whether the relationship of illustration to text that is structured into the experimental task fulfills the assumption of contextual relevance

is crucial to answering the question at issue. Furthermore, if the aim is to clarify what happens in the normal context of children learning to read, then this question, if not the only one, is certainly the most pertinent.

As yet only a few studies have been directly concerned with testing the effects of illustrations under the contextual hypothesis (Singer, Samuels and Spiroff, 1973-74; Denberg, 1976-77; Donald, 1979a)¹. Only the latter two, in fact, come close to meeting the task requirements for an adequate test of the contextual hypothesis and both of these have structural limitations. On the other hand, there is a large group of studies that have used the reading of continuous text to test the effects of illustrations on comprehension (Vernon, 1953; 1954; Dwyer, 1970; Bransford and Johnson, 1972; Peeck, 1974; Rasco, Tennyson and Boutwell, 1975; Rowher and Matz, 1975; Haring, 1978; Rusted and Coltheart, 1979a; b; Bianco, 1980). Although these studies have not been argued within the theoretical framework of the contextual hypothesis and although the crucial questions of word identification competence and strategies have not been broached within them, many of the results can be interpreted with at least partial relevance to the contextual hypothesis.

3. Aims

1. The wider purpose of the thesis is to develop

¹ These studies and the others quoted below will be reviewed in Chapter V.

and articulate a process framework for understanding and evaluating the development of reading.

2. Within this framework, the more specific purpose is to clarify the cognitive effects of illustrations in the development of reading. This encompasses three sub-goals:

i) First, resolution of apparent contradictions in research findings on the effects of illustrations both within and between the focal attention hypothesis and the contextual hypothesis will be attempted.

This is to be approached not as testing one hypothesis against the other, which would be regarded as irrelevant, but, through clarifying the question of processing differences as revealed in the structure of different studies and paradigms.

ii) Second, with respect to the effects of illustrations on reading development, the contextual hypothesis with its related processing assumptions, has not been specifically enough tested. With regard to these processing assumptions, therefore, the aims of both, but most specifically the second, experimental studies reported in this thesis include the following:

a) On the first assumption, that reading acquisition involves the development of strategies for processing information from a variety of sources, the aim is to test the effects of illustrations on a range of reading

behaviours that together reflect the effectiveness of the reader's information processing under normal reading conditions. These include measures of contextual word identification accuracy, strategies of information selection (orthographic, semantic and syntactic) and self-correction as well as measures of different levels of comprehension.

b) On the second assumption that reading is a linguistic process and that the linguistic information available to readers in the experimental task should parallel what is available under normal reading conditions, the aim is to test the effects of illustrations on reading material that is as representative as possible of normal learning-to-read conditions. This involves a variety of continuous, narrative passages, of reasonable length for the age of the readers, and representative of the content, linguistic structure and format of a typical basal reader.

c) On the third assumption that, under normal reading conditions, illustrations provide a source of contextual information that is relevant to, and therefore redundant within, the content and structure of the text, the aim is to test the effects of illustrations that fulfill this assumption.

The experimental material, therefore, comprises illustrations that are relevant, are sequentially related to the text and that, to avoid criticism of artificiality, are taken from the same basal readers as the text.

iii) Third, as has been pointed out, the effects of illustrations appear likely to vary with the developmental level of readers, the relative competence of readers and the relationship of illustration to text. Most specifically therefore, the aim of the second experimental study, within the framework of the contextual hypothesis and control of the relationship of illustration to text, is to test the consistency of the illustration effect over two developmental levels (seven-year-old and nine-year-old reading age levels) and two competence levels (high and low progress readers).

3. Finally, developing practical implications for the teaching and evaluation of reading within the contexts of both the wider and the more specific conclusions of the thesis is seen as an aim in itself.

CHAPTER II

THE PROCESS OF READING: READING AS SYNTHESIS

As long ago as 1908 - and relatively unrecognised for its worth - Huey said:

"... even if the child substitutes words of his own for some that are on the page, provided that these express the meaning, it is an encouraging sign that the reading has been real, and recognition of details will come as it is needed. The shock that such a statement will give to many a practical teacher of reading is but an accurate measure of the hold that a false ideal has taken of us, viz., that to read is to say just what is upon the page, instead of to think, each in his own way, the meaning that the page suggests, ... for reading is always of the nature of translation and, to be truthful, must be free" (Huey, 1908, p.399: author emphasis).

For 1908 this was a radical and profound statement; and it has only been in the past decade - primarily through the influence of psycholinguistics - that reading has become recognised for what Huey said it was: a process that is constructive, where the "fluent reader reads for meaning, not for the identificaton of letters, words or phrases" (Cooper and Petrosky, 1976, p.186).

This view of the process of reading has been labelled as 'analysis-by-synthesis' (Gibson and Levin, 1975) following the lead of Neisser's (1967) influential cognitive theory; and also as 'psycholinguistic' (Goodman, 1967; Smith, 1971;

1973). Essentially, whatever the label, it is a constructive or synthetic approach and must be seen in opposition to the more widely and commonly held notion that reading is essentially an analytic process. To appreciate this view of the reading process, an overview of the more traditional, analytic, standpoint needs to be given.

1. The Traditional View

A simple statement that summarizes the traditional view is given by Goodman (1967, p.259):

"Reading is a precise process. It involves exact, detailed, sequential perception and identification of letters, words, spelling patterns and large language units."

Such precise, sequential identification represents just that 'false ideal' of which Huey speaks. It has led teachers of reading to focus almost exclusively on the skills of precise identification. Phonic approaches are directed at letter and letter group identification, while whole word and sentence approaches are directed at precise identification of sight words or larger language units. Whatever the approach, it is assumed that the process of reading is the process of learning and using such identifications precisely. It is really only with the advent of language-experience approaches to teaching reading (Stauffer, 1970; Reid, 1974) that the idiosyncratic and constructive nature of learning to read has found some acceptance in teaching methodology.

Beyond reading teachers, research in the area has, until recently, focussed almost exclusively on the questions that surround the problem of precise identification. As Brown has put it (1970, p.164):

"For some decades, reading research concentrated on just two processes: the identification of letters and words as visual forms and the translation of such forms into speech sounds."

In cognitive psychology this research was based very much on the activity of fluent readers and was perhaps more concerned with elucidating the nature of perceptual processes than with the cognitive activity of reading per se (Neisser, 1967). In education the focus was on the most effective methods for teaching reading: that is, the most effective methods for developing the skills of precise identification. Chall (1967), in a comprehensive review of this conflicting research, concludes that methods of teaching reading that emphasise code analysis appear to be marginally more effective than methods that emphasise word or sentence recognition. Ironically all such methods are concerned essentially with precise identification and, in so far as this has been regarded as the process of reading, the research has really begged the question.

Gibson has consistently, and perhaps to the most sophisticated extent, tackled the ultimate analytic question of what, in the orthographic display, constitutes the basic unit of identification in the process of reading. Her basic conclusion, through a variety of experiments, is that:

"... the smallest component units in written English are spelling patterns ... a functional unit of one or more letters, in a given position within the word, which is in correspondence with a specified pronunciation" (Gibson, 1965, p.1071).

Despite Gibson's intention, and it is clear from later publications (Gibson, 1971; 1972; Gibson and Levin, 1975) that she has a far from simple, analytic view of reading, the conclusion on spelling patterns has been interpreted by reading methodologists as requiring yet another code of precise identification (Atkinson et al, 1970).

At this point it is as well to re-emphasise what was said in the introduction. Orthographic analysis does have its place in the reading process but it does not, itself, constitute the reading process. The real significance of Gibson's spelling patterns illustrates this: they are basic units in written English not because words cannot be read without analysing them but because they are the most likely units to be selected, on an efficiency of prediction basis, from the available orthographic display. The argument so far should therefore not be taken as rejection of the role of orthographic analysis. The issue is relative and not absolute and, as will be developed in the following chapter, the distinction between initial reading and fluent reading emphasises further the relativity.

2. Reading as Synthesis

The alternative view of reading as synthesis has been developed largely, although not exclusively, from the perspective of fluent reading. A number of theorists and researchers, from various theoretical standpoints, have contributed to this view. Some of the most influential of these have been Goodman (1965; 1967; 1968; 1969; 1970; 1973; 1974), Smith (1971; 1973; 1975; 1978), Kolars (1966; 1969; 1970) and Hochberg (1970).

At the root of the view is the realization that reading is a linguistic process. As Kolars (1970) puts it:

"... just as recognising words does not occur by the piecemeal recognition of their letters, reading connected discourse does not proceed by the piecemeal recognition of words ... any theory that attempts to account for reading in terms of translating graphemes into phonemes, in terms of the discrimination of individual letters, or in terms of a sensitivity to the morphemic structure of single words, is hopelessly insensitive to even the simplest kinds of linguistic processing the reader engages in" (Kolars, 1970, pp.100, 108).

In support of this view Kolars devised a number of ingenious experiments in which adult, fluent readers were required to read continuous text that had been transformed in various ways. Analysis of their responses yielded impressive evidence of the degree to which linguistic constraints operate in the process of reading. In one set of experiments (1969) in which text was geometrically transformed (mirror

transformations, vertical reversal of letters, etc.) subjects demonstrated that their ability to read the text was strongly influenced by syntactic expectancy. In terms of parts of speech, substitution errors for the various categories (noun, verb, article etc.) were in the same category as their stimulus word between 50-75% of the time as against a chance expectancy of 12%. Moreover, in terms of relative sensitivity to syntactic constraints as opposed to orthographic constraints, it was found that 89% of errors that were syntactically acceptable within their respective clause and were orthographically dissimilar to their respective stimulus word were left uncorrected. Conversely only 23% of orthographically similar errors that were not syntactically acceptable were left uncorrected. In another experiment (1966), bilingual subjects read passages where the text contained an almost equal distribution of French and English words, e.g.,

His horse followed de deux bassets, faisait
la terre resonner under its even tread. ...

Under these conditions subjects demonstrated the remarkable degree to which semantic and logical constraints operate directly in the process of reading. For instance, subjects who had been given the same amount of time to read similar passages in French only or English only showed no difference in comprehension of the various passages. This was interpreted as demonstrating that where a reader knows the words of a language he perceives them directly in terms of meaning and does not necessarily translate from graphemes to

phonemes to words in language A, from words in language A to words in language B, and from words in language B to meaning. This process would necessarily have been more time consuming than the reading of only one language and comprehension would have been affected. Subjects' errors in a similar oral reading task supported this interpretation. It was found, here, that a high proportion of 'errors' (textual inaccuracies) were translations of the printed word to its equivalent in the other language. In other words, not only was meaning being directly accessed but the expectation of meaning was over-riding the orthographic cues altogether.

The analysis of reading errors as the means for gaining insight into the process of reading has been developed, to the most sophisticated extent, by Kenneth Goodman.

As he argues:

"All responses to the graphic display are caused and are not accidental or capricious. In every act of reading, the reader draws on the sum total of prior experience and learning. Every response results from the interaction of the reader with the graphic display. Responses which correspond to expected responses mask the process by which they are produced. But observed responses (OR's) which do not correspond to expected responses (ER's) are generated through the same process as expected ones. By comparing the ways these miscues differ from the expected responses we can get direct insights into how the reading process is functioning in a particular reader at a particular point in time" (Goodman, 1969, p.12).

On the basis of this argument Goodman has made a detailed study of 'miscues' produced in oral reading performance (Goodman and Burke, 1973). This has been achieved through the development and application of a detailed taxonomy for recording and analysing the nature of miscues and the information selection on which they are based (1969). Using such analyses as his basic evidence, Goodman has developed a model of the reading process (1967) that attempts to come to terms not only with the degree to which linguistic constraints determine the process of reading but also with how the reader efficiently accommodates information from a variety of sources. Thus, briefly:

"Reading is a selective process. It involves partial use of available minimal language cues selected from perceptual input on the basis of the reader's expectations. As this partial information is processed, tentative decisions are made to be confirmed, rejected, or refined as reading progresses" (Goodman, 1967, p.127).

This process of selection, prediction and confirmation, Goodman maintains, is made effective through the variety of sources of information that the reader has available to him. Basically he sees these as the graphophonic, the syntactic and the semantic (1969) but he adds to these information sources that the reader brings from within himself (including his linguistic, experiential and conceptual background) as well as information that may be external to the text and the reader (including illustrations, charts, external prompts etc.) (1968).

Within this broad structure of information, therefore, the reader at any one point is engaged in a process of selecting what he judges to be the most reliable and efficient elements of information to help him continue generating the message of the text. The more the sources of information overlap and re-inforce one another, however, the more easily and efficiently can the process of selection take place.

Thus;

"Since the value of any bit of the (three) types of information must be related to the other available information, the choice of which bit to select can only be made in full context and the strategies for making those selections can only be learned in response to real language materials" (Goodman, 1969, p.18).

In the context of the potential value of illustrations as an extra-textual source of contextual information as well as the importance of reading development being evaluated on continuous text, Goodman's view has obvious relevance.

Frank Smith's contribution to the constructive or synthetic view of the reading process has been largely through his compelling synthesis of the theories and findings of information processing, psycholinguistics and perceptual theory as applied to the process of reading (1971 in particular). In the present context what is most significant is his explicit development of the concepts of 'informational redundancy' and of 'reduction of uncertainty' as applied specifically to the act of reading. These concepts are

not new but Smith's achievement has been in clarifying and integrating them in a direct and plausible account of the complex cognitive act of reading.

A full development of Smith's thesis would not be appropriate here. What follows, therefore, is a summary statement of his view in his own words, together with a brief explication of the two concepts.

"... reading is not a passive activity - the reader must make an active contribution if he is to acquire the available information. All information acquisition in reading, from the identification of individual letters or words to the comprehension of entire passages, can be regarded as the reduction of uncertainty. Skilled reading utilizes redundancy - of information from a variety of sources - so that, for example, knowledge of the world and of language will reduce the need for visual information from the printed page" (Smith, 1971, p.12).

The concept of informational redundancy is clearly central to Smith's argument. It is also a concept that will be drawn on frequently in this thesis. According to Smith,

"Redundancy exists whenever information is duplicated by more than one source ... or ... there is redundancy whenever the same alternatives can be eliminated in more than one way" (Smith, 1971, p.19).

In the context of reading, any section of text contains a variety of potentially overlapping or redundant sources of information. These correspond essentially with those

identified by Goodman (above). The functioning of redundancy within these sources of information is best understood through taking an example. Consider the following sentence:

'Some hungry cows trampled over my garden.'

For each word in the sentence, beyond the initial 'Some' there are degree of syntactic, semantic and orthographic redundancy. If the word 'cows' is taken, then, without decoding any of the orthographic information, syntactic constraints to that point determine that it is likely to be a plural noun and semantic constraints determine that it is bound to be animate. If minimal orthographic information is selected, say the 'c' together with a non-specific impression of its length and lack of ascenders and descenders¹ then the likely alternatives are limited to 'cows' or 'crows', and remotely, 'crews' or 'czars'. At this point there is a strong but not total overlap of information. If more orthographic information is selected, say a more specific estimate of length or even simply the curve of the letter 'o', then there is total redundancy. In other words, the word 'cows' can now be read either through combining the selected orthographic information with the syntactic and semantic information or through selecting the remaining orthographic information: in either case the alternatives to 'cows' are eliminated.

¹ Although not central to point being made, it is important to realise that orthographic information comprises more than visual cues. For instance any letter or letter combination constrains what sorts of letter are likely to follow it, thus further limiting alternatives.

Put differently, the remaining orthographic features are redundant within the semantic and syntactic information. Clearly far more of the orthographic information in the word 'garden' would be redundant as the syntactic, and, particularly, the semantic constraints increase cumulatively with the amount of preceding textual information.

The concept of uncertainty reduction in reading is closely related to that of redundancy. Originally it derives from signal detection theory and without developing the technicalities of its definition, it may be applied in the reading context, to the means by which a reader, as a processor of information, reduces the alternatives to any one item of information - whether letter, word or meaning. Essentially this is a matter of choices: the less constrained the target item of information, the more choices exist and the more definitive information must be sought to reduce the alternatives. This is inevitably inefficient and time consuming. Conversely, the more an item is constrained, the fewer choices are required to eliminate the alternatives and the more efficient can the process of uncertainty reduction be. Redundancy relates directly to this concept in the sense that the more overlap of information exists between the various sources of information, the more any one item of information is constrained. Referring to the example again, a reader might sense that 'cows' is a noun without knowing its meaning. In this case, reduction of uncertainty on the word 'trampled' would be relatively less efficient (it could be any action, rather

than action likely from cows) precisely because the degree of redundancy had been reduced through lack of semantic information. These two concepts have been developed in some detail as they are central to understanding the possible ways in which illustrations as information may be used by early readers.

Finally, Hochberg's model (1970) has relevance in the view of reading as synthesis in that he introduces two related concepts, one of which introduces an added dimension to what has been developed so far. In order to account for varying saccadic fixations observed, particularly between skilled and less skilled readers, Hochberg coined the complimentary concepts of peripheral search guidance (PSG) and cognitive search guidance (CSG). Briefly, in his model he maintains that there is an interaction between these two processes. PSG involves picking up low acuity and mainly orthographic, information in peripheral vision as an indication of where to locate the next information search. CSG involves predicting, on the basis of cumulative linguistic and contextual information, the next point of high or relevant information in the text. The interaction involves PSG being monitored and essentially informed by CSG while the predictions in CSG could also be modified by dissonant feedback from PSG.

In this context, several points in this model are of interest. First, CSG is to all intents and purposes equivalent to Goodman's conception of prediction. The concept of PSG, however, suggests the means by which successive

selections may be rationalized. Gibson and Levin (1975) are quite severe in their criticism of both the Goodman and Hochberg models. They maintain that the models are too diffuse and that questions such as exactly where and to what feature(s) of information the reader successively shifts his attention are unanswered. However, this can in part be answered by Smith's conception of uncertainty reduction where choices are made (i.e., that information is selected) on the bases of what alternatives remain. In the example, the alternatives 'cows/crows' demand precisely such orthographic information as will resolve that alternative. It is nevertheless unlikely that readers are consistently as systematically logical as this might suggest. What is more likely, and what is ignored by Gibson and Levin in their criticism, is that readers develop strategies for information selection. Orthographic information selection, for example, is never random in even partially skilled readers. It is well established for instance that orthographic information at the beginnings of words is more consistently selected than at the ends, and at the ends more than in the middles (Marchbanks and Levin, 1965; Shankweiler and Liberman, 1972; Weber, 1968). There is nothing that pre-determines this: readers learn that to select in this way is often successful, and it becomes a strategy. Clearly there are many other potential strategies, not only with respect to selection of orthographic information but also to the selection of syntactic, semantic and extra-textual information as well as integrative strategies that optimise redundancy in Smith's uncertainty reduction sense.

Returning to PSG, then, it is quite feasible that it is the reader's strategies that determine the initial point of selection and that this is monitored, and if necessary modified, by CSG.

The second point relates to the first. Neither PSG nor CSG are 'givens' in the reading process. They must, by definition, be learned ways of approaching text; strategies. PSG, in particular, must depend for its efficiency on a developing competence in the orthographic structure of written language. Evidence such as the finding (Gibson and Guinet, 1971) that morphological inflections like verb endings are identified and selected as unitary features increasingly with reading developmental level supports this. The questions, then, are how do the complimentary processes of PSG and CSG operate in the initial stages of reading acquisition, and most significantly for this study, on what sorts of information do they operate if competence in the structural regularities of written text, semantic, syntactic and particularly orthographic, is still being acquired? These questions will be central to the development of the next chapter.

3. Summary

In essence, then, the view of reading as synthesis could be summarized as follows:

- i) Reading is inevitably a linguistic process. In so far as it involves the generation of language meanings it must be heavily constrained by the linguistic structure

of the text in interaction with the linguistic competence of the reader.

ii) Fluent reading does not necessarily involve a precise, sequential decoding of the orthographic display to spoken language or its internal equivalent. Rather, the reader's linguistic competence together with what he expects in terms of meaning and the flow of language, operate on the total informational cue structure to generate the textual meaning in the most efficient and direct way possible.

iii) Since a variety of sources of information are potentially available at any point in a written textual sequence, the resultant informational redundancy or overlap constitutes the basis for efficient information selection and message construction in reading.

iv) The reader's competence in the structure of his language, syntactic, semantic and orthographic, together with his strategies for information selection and integration determine the efficiency of uncertainty reduction and use of redundancy.

v) As with spoken language, the process of message identification and comprehension does not involve precise, sequential identification of the perceptual elements in the signal. Rather, it involves a minimal selection from the total available cue structure in order to predict and confirm the message.

Finally, the contextual hypothesis:

vi) For optimally efficient reading, the degree of orthographic information required to be selected and processed is in inverse proportion to the degree of contextual information available and accessible to the reader.

CHAPTER III

READING DEVELOPMENT: MODELS OF THE TRANSITION TO FLUENT READING¹

The view, developed over the previous chapter, is that reading is a constructive process; that the fluent reader engages in a complex and efficient process of information selection, prediction and confirmation to generate, as directly as possible, the meanings of the text. Can learning-to-read be regarded in the same light, however? If not, and clearly there are differences, through what process does the learner graduate to this level of efficiency? Most theoretical models of the reading process have focussed either on the fluent process or on the very initial stages (Geyer, 1972; Mackworth, 1972a; Gibson and Levin, 1975). As Fugh (1978) points out, however, the transitional process - how the reader moves from one point to the other - has remained a relatively unexamined area. Where the aim is to establish a basis for understanding and evaluating the development of reading, as it is in this thesis, the question necessarily requires clarification.

1. Goodman's (1968) model

Goodman (1968) has presented, in the form of a developmental

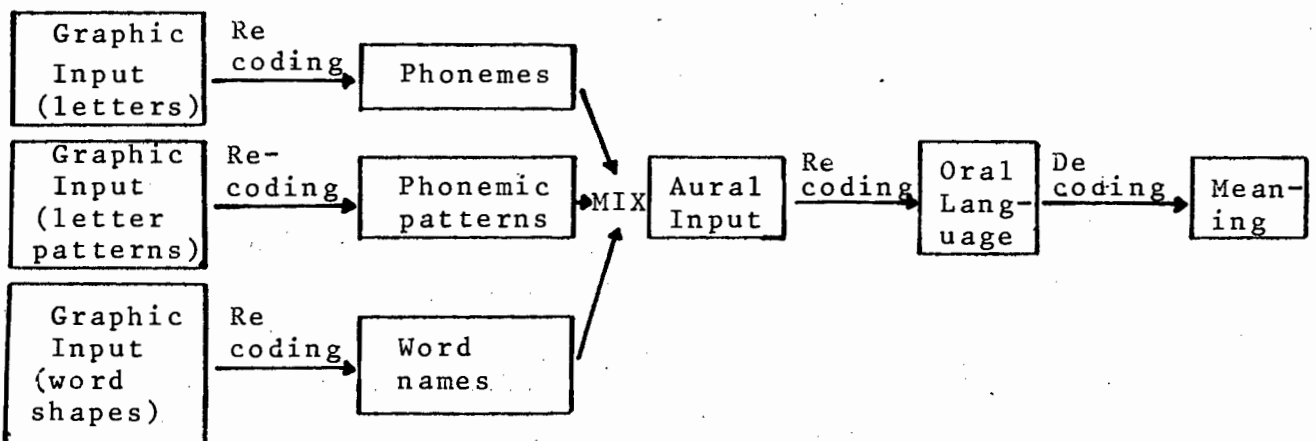
¹ This chapter is based on a recent published article by the author (Donald, 1981). Several substantial modifications and additions, however, have been effected.

model, one of the few attempts to explain the transition from initial to fluent reading. In this he suggests three levels of proficiency within which he makes a distinction between 're-coding' and 'de-coding'.

"In the early stages of reading the process may involve a stretching out so that graphic input is re-coded (not decoded) into aural input which is eventually decoded for meaning. ... recoding can take the form of assigning phonemic values to letters. It can take the form of assigning patterns of phonemes to patterns of letters. It can take the form of putting oral names on written word shapes" (Goodman, 1968, p.16-17).

This distinction plays a central role in the model of developing reading that he proposes. At proficiency level 1, for instance, a protracted process of recoding is suggested, and decoding to meaning only occurs once a full oral language message is available;

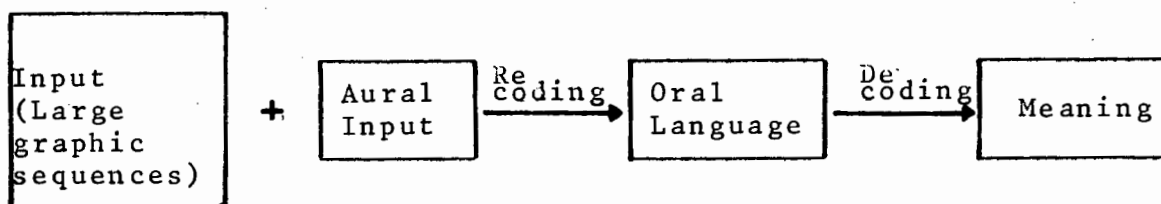
Proficiency Level 1



At proficiency level 2, Goodman suggests that,

"The aural input is supplied by the reader simultaneously with his recoding of the graphic input. To achieve this level of proficiency the reader must perceive letters and words always as parts of larger language units"

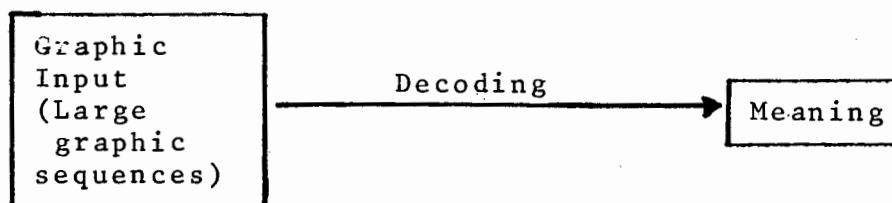
Proficiency Level 2



(op. cit. p.17-18)

Finally, at proficiency level 3, recoding and decoding occur simultaneously so that, under optimal reading conditions, the reader is decoding meaning directly from the graphic display.

Proficiency Level 3



(Op. cit. p.19)

This model comes close to elucidating the problem. Even here, however, the question of just how the learner makes the transition through the various levels is left unanswered.

2. Smith's (1971) model

Smith has also considered the difference between initial and fluent reading. Since his explanation is relevant and somewhat different to Goodman's it will be set out in full.

"The more difficulty a reader has with reading, the more he relies on the visual information; this statement applies to both the fluent reader and the beginner. In each case, the cause of the difficulty is inability to make full use of syntactic and semantic redundancy of non-visual sources of information.

This difference between the fluent and beginning reader may be epitomized in the manner in which the reader makes use of syntax, the bridge between surface structure and meaning. The fluent reader can be regarded as crossing the bridge from the meaning side, merely sampling the visual information to confirm his expectations. ... Syntax is a tool that the fluent reader uses to predict what the surface representation should be ...

The beginning reader, however, spends most of his time crossing the bridge of syntax in the opposite direction. ...he must deduce meaning from surface structure .. (and) .. this requires a maximum of visual information. Since there is no prediction of what surface structure will be, the novice reader is forced to analyse all the constituents of the surface representation, in order to be able to apply his syntactic skills. ...this is a slow and laborious process that is

almost certain to result in loss of comprehension ... (and) ... may create such a memory overload that it will in fact be impossible to apply the rules of syntax" (Smith, 1971, p.221-222).¹

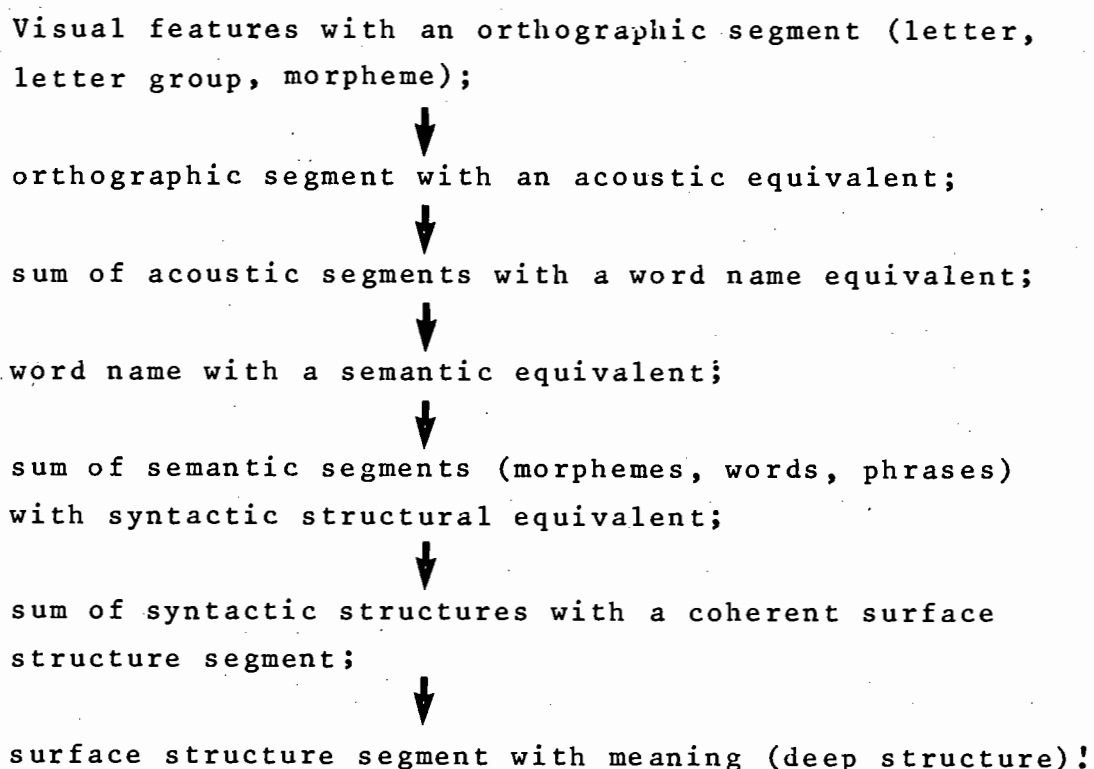
This quote contains at least two very important ideas and one inherent mystery. The important ideas, to be developed, and in part challenged, are, first, the inability of the beginner to make full use of redundancy and, second, his reliance on surface structure. The second idea matches with Goodman's protracted process at proficient level 1 but adds the notion that it is syntax in particular that provides the bridge between the surface structure and meaning. The first idea is not represented in Goodman's linear model and is an important omission. The mystery is in the bridge concept: How does the novice begin to cross the bridge in the more efficient direction? According to Smith he is 'locked' into the inefficient direction. The change of direction is surely not sudden, yet if it is gradual, what is the transitional process? Smith's partial answer to this is that reading experience provides the reader with the required distinctive features, visual-acoustic-semantic equivalences and knowledge of redundancy to make the processing change. But this really begs the question for, other

¹ Deletions in this quote - the original contains 91 words more without being different in meaning - illustrates an important aspect of redundancy at the meaning level! This is not meant as criticism of Smith's writing style: his purpose is clarificatory, where the purpose of quoting him is summary.

than in an unspecified, osmotic way, it is still difficult to know how the transition occurs.

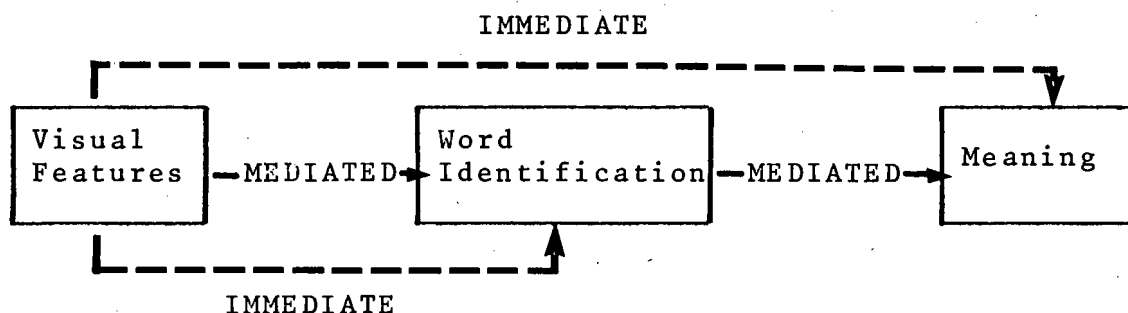
Before attempting an answer to this question, two additional concepts developed by Smith need to be introduced. In the context of access to textual information, he makes the distinction between words and meanings. Access to either may be through what he terms 'mediated identification' or through 'immediate identification'.

Mediated access refers to a process whereby a word is identified or, with more processing, meaning is attained through a protracted sequence of establishing equivalences between visual, acoustic, semantic and syntactic featural categories. Extrapolating from Smith, and in its most protracted form, the process could involve the following sequence of equivalences needing to be established:



Immediate access, on the other hand, refers to a process whereby either a word is identified directly, from selected visual features (without the intervening orthographic or acoustic analysis and synthesis) or meaning is identified directly from highly selected visual features (without establishment of any of the intervening equivalences). Whatever the level, it is informational redundancy that makes this possible.

Taking the concepts together, then, there are, according to Smith's argument, three access routes to meaning:



(adapted from Smith, 1971,
p.206)

The first and most protracted is mediated meaning access via mediated word identification. The second is mediated meaning access via immediate word identification, and the third is immediate meaning access. These bear obvious resemblances to Goodman's three proficiency levels. The problem with both conceptions lies in the initial, protracted process. Is it realistic to see this ever occurring? Mediated access, particularly when it encompasses meaning identification as well as word identification, clearly requires an impossible use of cognitive processing resources. Limited capacity systems such as short term memory and

attention, under such conditions (where the reader has to select, analyse, match, hold and synthesise cumulatively through a protracted sequence of operations), would simply not cope. La Berge and Samuels (1974), for instance, argue that unless a good proportion of these operations take place at an 'automatic level of processing', where focal attention is not required, then the process is a cognitive impossibility.¹ Yet children do comprehend meaning from continuous text even in the very initial stages. Given that this is so, there must be a misconception in how the process functions at this stage. An attempt at clarification of this misconception will in turn lead on to an explanation of how the transition from initial to fluent processing may be occurring through a process of accelerated progression.

3. An Accelerated Progression Model

The first key to the argument is the concept of informational redundancy. On the one hand it can be argued that the beginning reader comes to the task of learning to read with a developed competence in spoken language and that he can therefore make use of semantic and syntactic information despite his limited orthographic resources. On the other hand it

¹ La Berge and Samuels' concern is not to question the existence of the operations but to explain their functioning in terms of a theory of 'automatic processing'. As an explanation this has theoretical problems of its own that would be irrelevant to include at this point.

can also be argued that efficient use of redundancy is dependent on the prior establishment of equivalences at all the levels mentioned by Smith (above). The second alternative, true as it might be however, does not exclude the first. Redundancy may operate at relative degrees of efficiency. Moreover, linguistic redundancy, which is particularly subject to the prior establishment of the equivalences mentioned, is not the only form of redundancy. As mentioned in chapter II, extra textual information may also overlap with textual information. The question, then, is do beginners make use of redundancy? The evidence for this is, in fact, strong. Weber (1970a; b), Biemiller (1970), Clay (1968) and Goodman (1965) were amongst the first to present convincing evidence of the degree to which beginning readers make use of semantic and, particularly, syntactic information. There is no question today, of the validity of this. Moreover, Francis (1977) demonstrated how, in 5-year olds, information from the syntactic structure and meaning of sentences that had previously been read was used in a current reading task. In other words, this extra-textual source of information, in so far as it was redundant within the current text, facilitated reading. Finally, Denberg (1976-77) and Donald (1979a) have both demonstrated that, in first ^{and second} year readers, illustrations as an extra-textual source of information facilitate word identification in context. Again, where redundancy is available, the evidence is that beginners make use of it. The point of this is that although beginners may not yet be optimising

the redundancy that is available in the linguistic structure of the text, they are making use of whatever redundancy is accessible to them at their level of development. They are, in short, learning a redundancy strategy.

The second key to the argument lies in the questionable notion that beginners are necessarily locked into mediated access and that they are therefore denied the more efficient process of immediate access. Any reading teacher knows that children very soon acquire a remarkable facility in the rapid identification of a range of words that are common in their vocabulary - known as 'sight words'.

It could be argued that the identification of these words still demands mediated access in the sense that the visual features need to be matched to an acoustic equivalent before the semantic equivalent can be accessed. This, however, is debatable. The words are familiar and invariably of high frequency in the childrens' vocabulary so that the semantic threshold of accessibility is likely to be low (Paivio and O'Neill, 1970). The likelihood that a direct association between visual features and the semantic equivalent could be readily established, and early in the process of reading development, is therefore certainly high. Above the word level, the same could be said for a few, high frequency phrases ('Once upon a time!'). Webster, for instance, has made conscious use of this in his highly successful remedial readers (1970). The point, again, is that although beginners may be partly limited to mediated access they are likely, even from the early stages, to be

able to make use of considerable immediate access provided the words are familiar and of high frequency.

Returning to the protracted conception of mediated access and Goodman's proficiency level 1, the cognitive impossibility of these process sequences can now be modified. The beginning reader can now be seen as engaged in some mediated processing together with some immediate processing together with some use of redundancy. The cognitive load is considerably lessened in so far as the demands on short-term memory and attention are less densely cumulative. If this holds, then the question might be asked why beginning reading is nevertheless so inefficient. The answer is that the beginner is still heavily reliant on mediated access, he is not yet able to optimise the linguistic redundancy that is available, and his use of immediate access is still limited in its scope. The answer to the transition question could equally be that there is simply a gradual improvement in all three areas; more equivalences established leading to less protracted mediated access, a wider scope of immediate access and a more efficient use of available redundancy. This is similar to Smith's explanation and could be regarded as a simple cumulative model of reading acquisition. However, what is proposed is a progressive interaction of these three means of access that leads to a model of accelerated reading development.

It can be argued that immediate access occurs only where mediated access is possible, although not actively engaged. In other words, immediate access is based, and substantially dependent on, the build-up of featural equivalences established

through previous occurrences of mediated access. Equally, however, it may be argued that immediate access may occur, through the operation of redundancy, when the full featural equivalences are not available. Thus, meaning may be directly identified with only partial orthographic information, provided sufficient informational redundancy exists. However, redundancy in its turn is also related to the build-up of equivalences: The more equivalences that are established the more linguistic redundancy becomes accessible. Clearly, then, there is a three-way interaction. The efficiency both of immediate access and of redundancy are at least partially dependent on previous occurrences of mediated access - the establishment of equivalences. However, immediate access is also directly facilitated by redundancy, and redundancy becomes more accessible the more immediate access is possible.

Growth at the beginning therefore, must be very slow. Not only does the reader need to make considerable use of mediated access but development of the scope of immediate access and the efficiency of redundancy are both at least partially limited by the prior need for a build-up of featural equivalences. Because, as was argued above, the beginning reader does make some use of immediate access and redundancy however, these two will interact and, with the inevitable establishment of featural equivalences, the whole process will begin to accelerate.

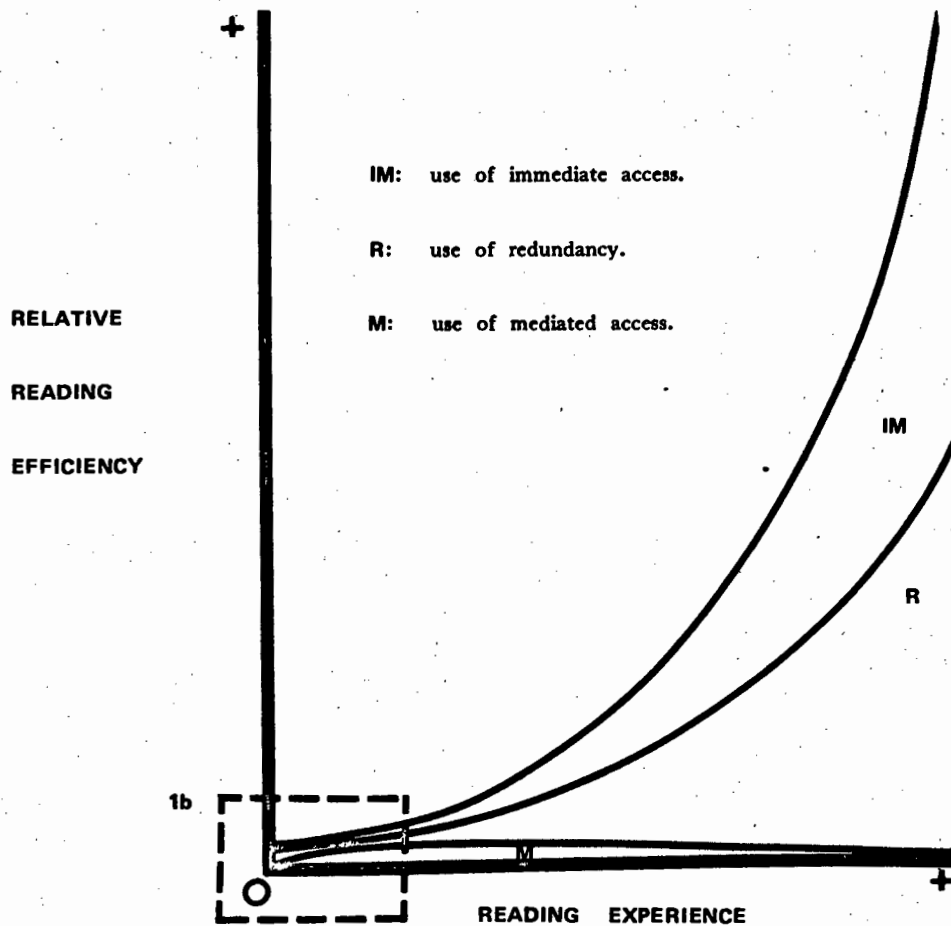
Loosely, the basis for this growth could be expressed as follows:

$$\text{reading efficiency} = \frac{\text{immediate access being used} \times \text{redundancy being used}}{\text{mediated access being used}}$$

Thus, although there is a heavy need for mediated access at the beginning, provided there is some use of immediate access interacting with some use of redundancy, there must be positive, if slow, growth. Progressive acceleration must necessarily follow as more featural equivalences are established, allowing less use of mediated access and facilitating the use of immediate access, of redundancy and of their interaction. Again, without attempting to be quantitatively specific, this could be represented as follows:

FIGURE 1a

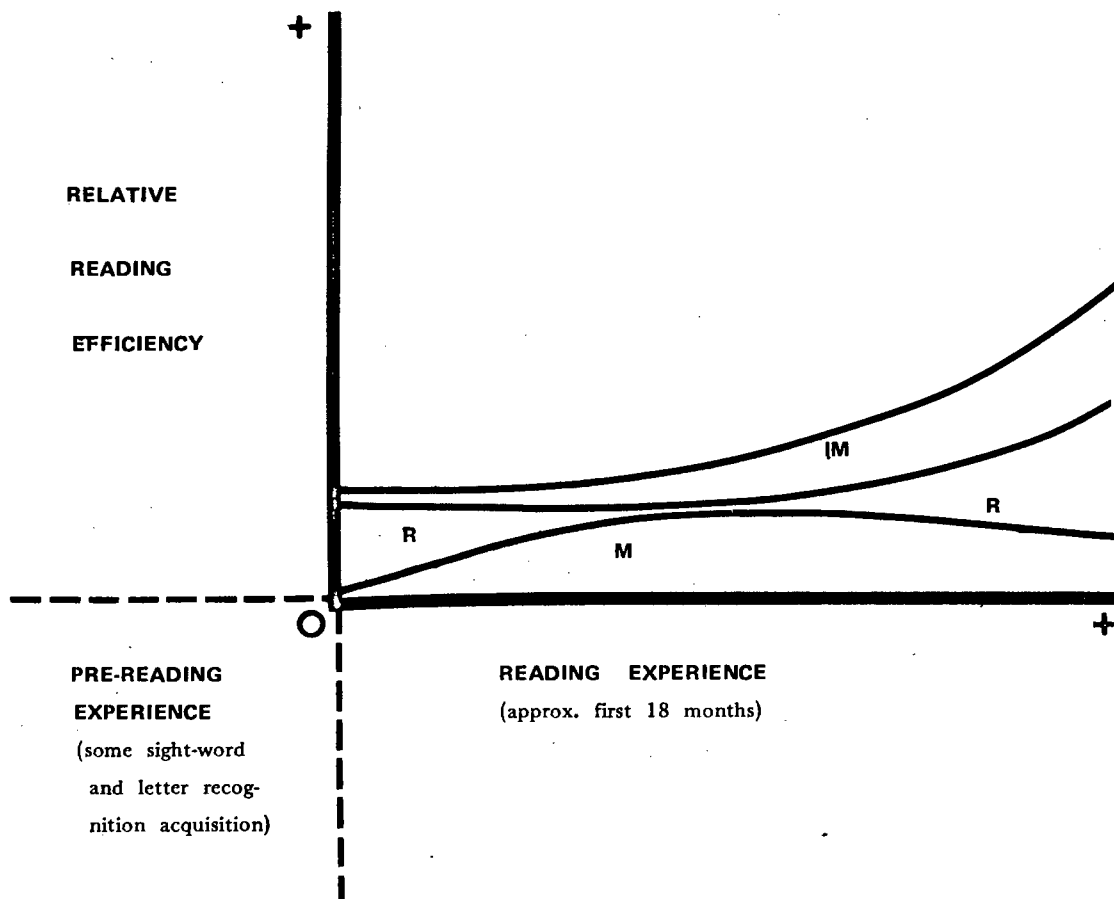
CONCEPTUAL ILLUSTRATION OF THE ACCELERATED
PROGRESSION OF READING DEVELOPMENT



Finally, there is some suggestive evidence (Biemiller, 1970; Donald, 1980) that beginning readers go through a progression from inefficient overdependence on contextual information (use of redundancy) to inefficient overdependence on orthographic information (use of mediated access) to a final and more efficient balance. It is only at the final stage that real acceleration in efficiency can begin. This matches and reinforces the present conception

FIGURE 1b

CONCEPTUAL ILLUSTRATION OF THE INITIAL
STAGES OF READING DEVELOPMENT



In conclusion, the question of how the transition from initial processing to fluent processing occurs is answered in two ways. First the difference between initial and fluent processing is not seen as absolute or as involving totally different means of access. Instead it is seen as a matter of relative balance between the three component means of access; use of mediated access, use of immediate access and use of redundancy. This overcomes the 'bridge' problem in that the beginner can be seen to be crossing in both directions (if one can stretch the metaphor!) to start with and gradually learning the means and strategies for crossing more in the efficient direction. It is neither a sudden occurrence nor does it involve a total change of processing.

Second, the transition is not seen as happening in discrete stages with the attendant problem of how the reader progresses from one stage to the next. Rather, it is seen as an accelerated progression with the component means of access developing cumulatively as well as in facilitative interaction with one another.

4. Implications of the Model

For understanding and evaluating the process of reading development, the model may be seen to have the following implications:

a) The processes involved in learning to read are not essentially different from those involved in fluent reading. Both may be seen to be basically constructive:

the differences are in the efficiency with which message identification is achieved.

Thus: b) The efficiency of the process in the initial stages is particularly handicapped by the beginner's lack of an established featural structure within (and between) the semantic, syntactic and, particularly, the orthographic sources of information. This necessitates the use of a bridging process, mediated access, that involves a protracted and uneconomical build-up of equivalences to, ultimately, access meaning. In the course of using mediated access, however, featural equivalences become established and these facilitate a wider scope for immediate access and a more efficient use of redundancy.

But: c) Mediated access, although necessary within the development of reading efficiency, places too great a load on cognitive processing resources for it to be, alone, effective in reading for meaning. Even from the initial stages, therefore, it must be balanced by some degree of immediate access and use of redundancy. The strategies that the reader develops in order to balance these three means of access, and to make efficient use of the variety of sources of information accessible to him, will determine the efficiency of his processing at any stage of development.

Therefore:

i) In teaching children to read, attention needs to be given to the development of mediated access and the build-up of featural equivalences that this involves.

Phonics, structural analysis and other 'precise identification'

teaching methods do this. However, this should be clearly seen, and taught, as only one means of access to word and meaning identification: not the means. 'Sight' learning that capitalizes on familiar, high frequency words and phrases and emphasises direct meaning identification¹ also has its place not only in fostering a store of words available through immediate access but also, provided it is also emphasised in context, in facilitating a strategy, as well as refinement of the process of immediate access itself. Both of these approaches are generally taught - although the latter with perhaps the wrong emphasis and the former with too much emphasis! What is seldom actively taught, however, is the efficient use of available contextual information and the use of redundancy. As has been indicated, there is sufficient evidence that children naturally make use of redundancy. But do they optimise redundancy? The strategies of efficient information selection, contextual prediction and message confirmation (self-correction) all involve the use of redundancy and it is the author's experience that they are not effectively learned and that they can be effectively taught.

Perhaps most important, however, teaching needs to be concerned with those strategies that optimise the balance between the three interacting processes. This can only occur in contextual reading where the emphasis is placed on the

¹ For example presenting a flashcard and asking a child to react to it as quickly as possible rather than say the word.

ultimate goals of all reading: efficient and accurate message identification and comprehension. A learner who clearly perceives these, and gets adequate feedback on them, as the all important goals of reading will adjust his strategies to optimise these goals.

ii) In evaluating the effectiveness of reading development it is clear that level of performance, in conventional terms, is an important criterion. Since the ultimate outcome of reading is comprehension, this, as one criterion of performance, cannot be questioned. Equally, since accuracy of word identification in context reflects all the component processes and their integration, to that level of processing, it also has relevance. It should now go without saying, however, that isolated word recognition only reflects partial mediated access (to word identification level) and, questionably, immediate access (recalling that the latter interacts with redundancy). It may be taken as a partial measure of the development of a component process - but as no more. Measures of comprehension and of contextual word identification accuracy are thus well accepted as indicators of a child's progress on the ladder of reading acquisition (Neale, 1958; Pumfrey, 1976). What is not as well accepted, but deserves to be, is a measure of the efficiency of the learner's information processing strategies. Performance criteria measure the outcomes of his processing. Whether those outcomes reflect optimal processing is, however, masked. Attempts to gauge the nature, balance and relative efficiency of the reader's strategies becomes a necessity if the very process of

reading development is seen as a progressive refinement and integration of such strategies.

Since these issues and their practical implementation are central to the experimental paradigm they will be developed in the next chapter.

CHAPTER IV

EVALUATING READING DEVELOPMENT:

Theoretical justification of Dependent Variables and their Relevance in Testing the Effects of Illustrations under the Contextual Hypothesis

Where the concern is to establish the effect that any independent variable has on reading development it is clear that the dependent measures of what constitutes reading development must be theoretically justifiable and must demonstrate a reasonable degree of validity and reliability if the experimental effect is to be interpretable with any conviction. The question of the reliability of the measures as used in this study will be dealt with in chapter IX. This chapter is concerned with the theoretical justification for selection of the dependent variables as well as theoretical consideration of their definition and validity.

As has been developed to this point, reading is a complex process and reading development, in particular, cannot be viewed as a simple, unitary process of skill acquisition. The view of the process of reading development that has been put forward encompasses, essentially, two inter-related components. The first is that the process involves the progressive acquisition of featural equivalences across the orthographic, semantic and syntactic sources of information available in any continuous text. The second is

that there is a progressive development of strategies of information processing such that the reader learns to optimise the information that is accessible to him for the purposes of efficient and accurate message identification and comprehension.

Two necessities follow from this view. First, if the process of reading development as it has been defined is to be reflected, then the task on which reading development is gauged must allow that process to act. The reading of continuous text, in so far as it includes those sources of information that the reader is learning to process, must therefore constitute the task. Further constraints of what may be regarded as normal text for the purposes of learning to read have been mentioned in the introduction and will be further developed under the Materials section of chapter IX.

Second, if the process of reading development as it has been defined is to be reflected, then the dependent variables must represent the effectiveness of that process. With this in mind, the key phrase in the definition above is '...for the purposes of efficient and accurate message identification and comprehension'. If these constitute the purposes of the process then the effectiveness of the process must be evaluated against these criteria.

Exactly what is meant by 'efficient and accurate message identification and comprehension' requires some prior clarification. In the first place a distinction is drawn between 'accurate message identification' and 'comprehension'

This refers to the difference between accurate identification of what the author has said on the one hand and interpretation of what he has said on the other. In behavioural terms, this distinction becomes blurred as one can only really judge the former in terms of the comprehension that the reader manifests and thus must confound identification of the message with its interpretation. The closest one can get to accurate message identification is the accuracy with which the reader identifies the author's words. This is a reasonably valid assumption in so far as mis-reading of the author's words necessarily constitutes some distortion of the message. A problem here, however, is that errors vary in the degree to which they alter the meaning of the intended message. Some errors can be seen as virtually synonymous with the textual message while others may be totally meaningless or may only partially distort the message. A solution, therefore, is to measure the degree to which errors are acceptable to the meaning of the textual message and to evaluate accurate message identification on the basis of word identification accuracy as well as on the degree to which errors are acceptable within the semantic framework of the text.

In the second place, the concept of 'efficiency' that is related to 'message identification' refers to how the reader achieves this - as opposed to what he achieves. Again the distinction is not easy to specify in behavioural terms. What a reader achieves, whether in terms of word identification accuracy or of comprehension, is likely to reflect the efficiency of his strategies but it does not

reveal how these have operated. It is the 'how' that is needed in order to judge whether the strategies used to reach a given level of achievement were optimal or not. If the concept of strategy that is so central to the definition of reading development is to be reflected, however, then this needs to be attempted. Given the complexity of the reading process, an almost infinite variety of strategies could theoretically be identified. In practical terms, however, the main factor which serves to limit this choice is the nature of the contextual hypothesis that is being tested. This determines that, at the least, those strategies which reflect the use of contextual information and orthographic information should be revealed. Three strategies, therefore, reflecting the use of semantic information, the use of syntactic information and the use of orthographic information were selected. These are all based on patterns of information usage as revealed in the reader's errors. In addition, a strategy that has been established as related to optimal reading development could, a priori, be regarded as relevant. The strategy of self-correction fits this requirement and was therefore selected.

Finally, as has already been stated, comprehension, as the principal purpose of reading, cannot be questioned in its general relevance and validity as a measure of the effectiveness of reading. Neither is there a problem in finding a range of behavioural measures of comprehension. The problem that does exist is a conceptual one. As a

concept with a defined theoretical framework, comprehension is notoriously ill-understood (Gibson and Levin, 1975). Nevertheless, whatever comprehension is, it is doubtful whether it can be regarded as a unitary concept and whether it can be measured in any one prescribed way (Davis, 1968). The choice therefore becomes one of deciding between alternatives either on an ad hoc basis or on the basis of an, at least, attempted theoretical rationalization of component skills. The latter choice was made in this study and the basis for the selection of dependent variables under the generic label of comprehension will be developed under that heading, below.

The dependent variables selected as representative measures of the effectiveness of reading development as it has been defined, therefore, were as follows:

1. As measures of the effectiveness of accurate message identification:
 - i) Word identification accuracy
 - ii) Semantic acceptability of Errors
2. As measures of the efficiency of message identification or strategy.
 - i) Semantic acceptability of Errors.
 - ii) Syntactic acceptability of Errors.
 - iii) Orthographic acceptability of Errors.
 - iv) Self-correction.
3. As measures of comprehension:
 - i) Literal comprehension
 - ii) Inferential comprehension.

For future reference in the thesis, the principal headings will be reduced to 1. Accuracy; 2. Strategy and 3. Comprehension. Since 'Semantic Acceptability of Errors' is of interpretive relevance under both Accuracy and Strategy it will be considered under both headings where conclusions are being drawn. For convenience, it will otherwise be considered under the heading of Strategy only.

1. Accuracy

i) Word identification accuracy

As a criterion of the effectiveness of oral reading development, word identification accuracy is widely accepted and appears in almost all formal reading tests that are based on contextual reading (Spooner, 1976; Pumfrey, 1976). As a measure it is based on the number of words correctly read taken as a proportion of the total number of words in the relevant section of text. As a reflection of the reading process, therefore, it encompasses all the component skills and strategies that potentially go into identifying the flow of written language in context: in particular, the use of orthographic, semantic and syntactic information as well as the use of redundancy. Presumably, and this is an assumption, the more effective the reader's integration and use of such information, the more effective will his word identification be. This assumption may not be valid, however, since a reader may read very accurately without using more than the orthographic information. It is for this reason that 'accurate message identification' has been argued as the more valid criterion and that word

identification accuracy be taken as a component measure of this but that the semantic acceptability of errors be used as a moderating variable in evaluating the effectiveness of accurate message identification.

A further problem is that what defines word identification accuracy is not as clear as it might appear. Should repetitions be regarded as errors, for instance? Weber (1968) and Goodacre (n.d.) both argue strongly that they should not. Most repetitions, in the actual observation of children reading, may be seen to be a constructive form of pause, where the reader 'takes stock' by repeating a correctly read word or phrase before attempting the next word or section of text (Goodacre, n.d.). As Weber puts it, repetitions are more "an act of confirmation rather than an error" (1968, p.102). Even when repetitions constitute successive, and wrong, attempts at an initial error, they may still be seen as partial corrections of the initial error: it is logically inconsistent to take them as additional errors to the initial error. On the other hand, Ekwall (1974), studying the occurrence of repetitions at the independent, instructional and frustration levels of reading difficulty, maintains that they should be counted as errors since they reflect the level of difficulty that the reader is experiencing. However, in this study at least, the ultimate criterion is taken as accuracy of message identification and in so far as repetitions, of either type, are unlikely to distort the message further, they are not taken as additional

errors in themselves¹.

In testing the effects of illustrations under the contextual hypothesis, therefore, accuracy of message identification is of central relevance. In the sense that information available in an illustration may provide a contextual set or expectancy in the reader and may, furthermore, increase the redundancy of, particularly, semantic information, accuracy of message identification will reflect the degree to which this information is relevant and is effectively utilised in order to generate the actual message of the text.

2. Strategy

As pointed out earlier, in evaluating the process of reading development one needs to be concerned with more than absolute levels of achievement: One needs also to be concerned with how that process is functioning. It is through the learner's actual behaviour, while he is engaged in the process of learning to read, that insight into that process - and the learner's strategies - can be inferred. As Goodman, the principal exponent of this idea, has put it:

"The reader omits a word and some word parts, invents a word, substitutes other words, goes back at times to correct himself, and comes out with a meaningful sentence.

¹ Greater detail of error definition for the purposes of transcription and scoring appear in chapter IX.

We must be concerned with more than his superficial behaviour. We must infer from it the process he has used and his competence with that process... Both his expected responses and his miscues (errors) are produced as he attempts to process the print and get to meaning. If we can understand how his miscues relate to the expected responses we can also begin to understand how he is using the reading process" (Goodman, 1973, p.5 - my emphasis).

Thus, it is the reader's errors that provide the essential raw information on which strategy analysis may be based. In itself the analysis of oral reading errors has a considerable history. However, as Weber (1968), in her extensive review of this, points out, the analysis of oral reading errors has taken place in two distinct contexts. The first and historically oldest context, involves the descriptive categorization of errors. Here, errors have been taken, simply, as indicators of inadequate learning: literally wrong responses. Such categorizations are typified in a number of oral reading tests (inter alia, Durrell, 1955; Neale, 1958; Spache, 1963) where errors are classified as omissions, substitutions, insertions, refusals, repetitions, mispronunciations, reversals, and other variations or extensions of these classes. The difficulty in this context has been in clarifying the classes themselves. The fault, in most cases, has been in a lack of theoretical structure; a lack of recognition that reading involves processing language such that classes by which errors are analysed, to be consistent and meaningful, must necessarily reflect linguistic class structure rather than mere descriptive

characteristics of the errors. The result has been classes that are often ill-defined, overlapping or contradictory. For example 'was' read for 'saw' may be classed as either a reversal or as a substitution: 'us' read for 'use' may be classed as either an omission or as a substitution depending on whether a letter level or word level of classification is being used: Similarly, 'he was running' read for 'he ran' could, at the morpheme level, be taken as two omissions and a substitution; at the word level it could be taken as an omission and a substitution; while at the phrase level it could be taken as a single substitution. The measure of this confusion, as Weber (1968) points out, is that comparisons across research studies in this context are largely meaningless and few general conclusions can be drawn from the results.

However, perhaps the most basic problem with descriptive analyses of errors is that error classes remain linguistically and functionally undifferentiated. (This is particularly true of 'substitutions' which constitute a very large, and in descriptive terms, undifferentiated category. For instance, of 1943 first-year reading errors studied by Weber (1970b), 1674 of them, or 86% were substitutions). For example, there is a wide linguistic and functional difference between a 'nonsense' substitution and one that is appropriate to the context in which it occurs:

e.g., Suddenly they heard a splash (read as 'spalas')

versus

Suddenly they heard a splash (read as 'crash')

This leads into the second context in which oral reading errors have been analysed. In the main it is represented in more recent research (inter alia Clay, 1968; Weber, 1970a; b; Biemiller, 1970; Siler, 1973; Goodman and Burke, 1973; Cohen, 1975; Burke, 1976; 1977; Hood, 1976; Francis, 1977; Donald, 1979a; Potter, 1980). Although some earlier studies (Payne, 1930; Madden and Pratt, 1941) had indicated an awareness, if not explicitly developed, of the linguistic constraints operating on the nature of errors, the more recent studies have been generated within an articulated linguistic framework. This approach to error analysis has its roots in oral language research (Fromkin, 1973) and draws its theoretical underpinning from linguistic (Gurney, 1976) and, more specifically, psycholinguistic (Goodman, 1969) theory. Essentially, since reading errors, alongwith correct responses, must be constrained by the information available, they must reveal the reader's use of the linguistic information that, in itself, constitutes the context and structure of continuous, written text. Thus, as opposed to descriptive error analysis this has ensured a consistent theoretical approach to the linguistic analysis of oral reading errors.

From a practical, methodological point of view, a number of systems of analysis have been developed. Perhaps the most notable and complex system is Goodman's taxonomy (1969). Yetta Goodman and Burke have adapted this to more practical purposes in their commercially published Reading Miscue Inventory Kit (1972). Within the purposes and context of particular research studies (including most of

those mentioned above), a number of other viable systems have been developed. The point to be made is that although variations in methodology exists, this variation nevertheless takes place within clear limits. All systems recognise that an error is not random. All recognise that an error must be constrained by the available information and that this, basically, can be divided into three dimensions: orthographic information (based on visual cues from the graphic display, phonic information - letter-sound correspondences - and information from the broader orthographic structure of written language); semantic information (based on cumulative textual and extra-textual 'meaning' cues interacted with the reader's background of experience); and syntactic information (based on cues from the cumulative syntactic structure of the text interacted with the reader's competence in the syntactic structures of his language). Variations in systems of analysis occur within these boundaries and may, according to the demands of a particular research context, emphasise one or another dimension; analyse at greater or lesser depth; combine dimensions; or define sub-categories within a dimension. The underlying rationale, however, remains constant: an error is the product of the reader's interaction with, and processing of, the three basic sources of information - and the patterns within his errors reveal the competencies and strategies that are characteristic of his processing.

Within this framework, the system of error analysis as developed in this study will be discussed. Decisions

as to how to structure this system were built around four considerations that must be regarded as basic to the structuring of any such system of error analysis (Donald, 1979b). The first concerns the number of errors per analysis that can be regarded as a reliable minimum to reveal a characteristic pattern or strategy of information usage for any one reader. There are no absolute guidelines on this question. Precedent, however, can give some indication. In terms of a minimum, Burke (1976), using the Neale Analysis of Reading Ability (Neale, 1958) at a 'level of difficulty appropriate to age' (Burke, 1976, p.35) - i.e., less than 16 errors per story - must have been working with an average of approximately 10 errors per analysis. Hood and Kendall (1975) in their study of reflective and impulsive readers quote a median of 12,5 errors per subject. Potter (1980), replicating Burke's (1976) study with the Neale material reports an average of 8,6 errors per subject. Donald (1979a: Experiment I) worked with an average of 12 errors per subject. Clearly the more errors that are analysed the more reliable is the analysis. It is doubtful, given the inevitable variance around an average, that an average of less than 10 errors per analysis can give a reliable indication of what is a characteristic pattern of information usage. In Experiment II, therefore, assuming readers to be reading at instructional level (Christenson, 1969) and therefore at approximately 95% accuracy, stories were constructed (average: 323 words per story) to allow an average of approximately 16 errors per analysis.

Since errors followed by self-correction, (not counted as errors on the accuracy criterion) were also to be included in the analyses, this was expected to raise the average to approximately 20 errors per analysis.¹

This was regarded as more than adequate and, in the final event, adds to the confidence with which the results of the analyses can be interpreted.

The second consideration relates to what sorts of error to include in the analyses. One system is to use only substitution errors as, for example, in Cohen (1975) and Weber (1970b). These are the most common sort of error and, unlike omissions and additions, substitutions can clearly be analysed in terms of all three dimensions of information. If this system is used, however, it necessarily distorts the pattern of errors since 10-15% of errors must be excluded from the analyses. The alternative is to include additions and omissions (refusals, a very small proportion of errors if the child understands that he is expected to give a response, must necessarily be excluded since there is no information on which to base an analysis) and to regard them as having, a priori, not been constrained by orthographic information but to have been potentially constrained, as with substitutions, by semantic

¹ In the final results, the average accuracy was 93,54% - slightly lower than expected owing to the lower accuracy on unillustrated stories. With the addition of self-corrected errors in the analyses, the total number of errors analysed was 3140; an average of 26,17 (SD 15,62) errors per analysis.

and syntactic information. This alternative was adopted in both experiments.

The third consideration pertains to selection and definition of the dimensions of information that are to be considered in the analyses. The most clearly documented and accepted are the three primary cue dimensions: orthographic, semantic and syntactic. However, as briefly mentioned above, variation within these can be legitimate. Pumfrey (1977), for instance, suggests an interaction of these with the additional dimensions of 'direction', 'modality' and 'source'; a highly complex system. Equally, the primary dimensions may be sub-divided into more specific sources of information. The orthographic dimension could be broken up into information from both the 'sound' of the stimulus word as well as from the actual graphic features (Yetta Goodman and Burke, 1972). The syntactic dimension may be sub-divided into a variety of complex levels (Goodman, 1969) or, more simply in terms of the error's fit with 'pre-error content', 'sentence context' and 'passage context' (Hood, 1976). The semantic dimension, in turn, may also be divided into a variety of sub-sets which may include an error's link with previous textual meaning (Weber, 1970a) or with specific extra-textual referents (Francis, 1977). Conversely, the semantic and syntactic dimensions may be combined into a single dimension representing 'contextual' information (Biemiller, 1970; Hood and Kendall, 1975; Hood, 1976). The decision in both of the experiments reported here was to retain the three basic dimensions, as most widely accepted, since there was no reason to sub-divide them. There might

have been a case for combining the semantic and syntactic dimensions under the single dimension of 'contextual' information but it was felt that not only would this mask potential differences but that, in relation to illustrations, interest focussed particularly on the semantic dimension and this should therefore be analysed separately. The final definitions and criteria used in analysing errors on the three dimensions will be set out more fully below.

The fourth consideration relates to the criteria by which errors may be judged as reflecting use of information in any one dimension. Such criteria may be discrete (e.g. Biemiller, 1970) or continuous (e.g. Goodman, 1969). The use of a discrete criterion involves a decision - on an all or nothing basis - as to whether an error can be said to have been constrained by the information on any one dimension. The use of continuous criteria, on the other hand, involves a decision as to the degree of constraint evident in the error. The advantage of a discrete criterion is that it is simple to apply and score by and, therefore, tends to be reliable. Continuous criteria, on the other hand, although more complex in application are more sensitive to the relationships involved and are, therefore, more valid. The latter alternative was selected in both experiments and the question of reliability was met through a detailed specification of the scoring procedure (chapters VI and IX).

i) Semantic Acceptability of Errors

For the purposes of both experiments the semantic dimension

of information was defined as that information, textual or extra-textual, which contributes to the actual message of the text being read. This included a contextual build-up of meaning that may apply at the level of word-meaning; at the level of phrase meaning; at the level of clause or sentence meaning; at the level of inter-sentence meaning; or at the level of meaning that encompasses the passage as a whole to the point that is being read. It also includes the level of associated meanings, derived from sources outside the text but relating directly to the message of the text (e.g. illustrations, related texts that might previously have been read; and the reader's own experience as related to the content of the text). In this dimension, therefore, the available information is constituted out of a complex structure of meaning such that semantic cues (cues to the actual meaning of part of a word, a word or a group of words) may derive from any or all of a number of levels of what may be termed accumulated semantic expectancy.

Developing a scale of criteria for evaluating the degree to which an error has been constrained by semantic information was both crucial and difficult. It was crucial in the sense that testing of the effects of illustrations within the contextual hypothesis was seen to be centrally related to effects on the reader's use of accumulated semantic expectancy - as defined. It was difficult in the sense that, by comparison, both the syntactic and the orthographic dimensions have more established sets of criteria that are also easier to specify in terms of clear

hierarchies of relative acceptability. In addition, Goodman (1969) warns against the potential confusion of semantic and syntactic constraints in evaluating the respective acceptability of errors within these two contexts. Although the distinction is not always easy to maintain, and although there is some evidence that syntactically constrained errors are usually semantically constrained as well (Weber, 1970a; Siler, 1973) it is also clear that the degree of semantic and syntactic constraint is not necessarily the same. In other words, an error with, say, a high degree of syntactic acceptability is unlikely to be semantically meaningless¹ but may have only a relatively low degree of semantic acceptability within the textual message. Conversely, an error with high semantic acceptability is unlikely to be syntactically unacceptable but an error with low semantic acceptability may yet be highly constrained syntactically. One of the advantages of using continuous as opposed to discrete criteria is the potential for making such distinctions. Equally, in experiment II at least, the distinction between semantic and syntactic constraint is strengthened by evaluating errors, semantically, against the criterion of acceptability to the meaning of the text. Syntactically, on the other hand, errors are evaluated against the criterion of acceptability to language: where 'language' may be taken

¹ In the actual analyses, even this occurred quite frequently. An error was quite often a 'nonsense' word - e.g. 'fericely' for 'fiercely' - yet displayed obvious syntactic constraint.

as the reader's cumulative version, including a sequence of errors, as distinct from the language of the actual text. Burke (1976, 1977) as well as Potter (1980) used the 'reader's version' for evaluating both the syntactic and the semantic acceptability of errors. It is contended here that this not only clouds the distinction between the two areas of constraint but that it is also theoretically misplaced. Almost any error can be said to be constrained by some meaning related to the reader's idiosyncratic interpretation. Semantic acceptability, however, must have some definable point of reference and the message of the text is the only such reference point. Conversely syntax refers to the structure of language; a structure that can be identified and evaluated independently of the text. An example taken directly from one of the protocols of experiment II demonstrates this convincingly.

Error: pip and starting quickly in this disturbance

Text: Filling his pipe he stared quietly into the distance.

In this case, as in many others encountered in the analyses, the syntax of the language generated through the accumulated errors is totally acceptable while the message that is generated is definitely at variance with that of the text! Thus, in experiment II, the reader's version under specified conditions, may be taken as a legitimate ground for evaluating the operation of syntactic constraints. The message of the text, however, is taken as the only ground for evaluating

the operation of semantic constraints. In experiment I, however, the textual version was used as the criterion on both scales.

In deriving the set of criteria for the semantic scale of acceptability, precedent was less helpful than for the other two scales. The range of specificity has extended from Goodman's (1969) ten-point scale that necessarily involves fine distinctions between shades of meaning that verge on the subjective¹; through Burke's (1976; 1977) five-point scale that uses precisely the same criteria as for the syntactic scale; to the single, if rather insensitive, criterion that the error either does or does not conform to the preceding context (Weber, 1970a; Biemiller, 1970). None of these, or other available alternatives, were regarded as adequate. For experiment II, the following scale of criteria were therefore developed (partially based on Donald (1979a; experiment I) but differing on several points of definition: c.f. chapter VI).

¹ e.g. scale point 4 : 'There is some association between the meaning of the O.R. (observed response) and the E.R. (expected response).'

scale point 6 : 'The O.R. has an associated meaning with the E.R.

(Goodman, 1969, p.26).

Scale for Evaluating Semantic
Acceptability of Errors¹

<u>SCALE</u>	<u>CRITERION</u>
4 :	<p>E (error) is synonymous with S (text) : E is acceptable within the text in terms of both deep and surface structure.</p> <p>e.g.</p> <p>E afraid</p> <p>S ... was not <u>frightened</u> of ...</p>
3 :	<p>E involves a non-substantive change in meaning that does not alter the basic, cumulative message of the text : E is acceptable within the deep structure of the text.</p> <p>e.g.</p> <p>E hand</p> <p>S waved his <u>arm</u></p>
2 :	<p>E is acceptable within the meaning of its clause or sentence <u>but</u> does not fit the cumulative message of the text.</p> <p>e.g.</p> <p>E would</p> <p>S ... he <u>wouldn't</u> allow ...</p>

¹ For this and the two following scales, details of criteria including 'particular cases' and a variety of definitive examples, as used in actual scoring, are provided in chapter IX. What is presented here is only the logical framework.

1 : E is semantically acceptable within neither the general nor the local meaning context, but is a lexically acceptable word.

e.g.

E sat

S He set off through the forest

0 : E is semantically acceptable on none of the above criteria.

e.g.

E 'mencily'

S ... glared menacingly ...

The principal departures from other semantic scales involve the criterion of 'deep structure' acceptability; the concept of the 'message' of the text; and the idea of 'lexical acceptability'.

According to Smith's (1971) argument - and this is expressly related to the body of psycholinguistic theory - one of the ultimate purposes of reading must be to access and identify the deep structure of the text : the message of the text. The fact that the highest semantic scale score involves acceptability to both deep and surface structure merely recognises that synonymity must be closer to the intended message. The distinction between 'substantive' information and 'relational' information is made by Kolers (1970). It is essentially relational information that may be dispensable without altering the deep structure of the text. The manifestation of this concept is made explicit

on the particular sorts of error occurrence that were regarded as acceptable for scale score 3 (see chapter IX).

The concept of the 'message' of the text has already been amplified. In this context it serves merely to crystallize the difference between semantic and syntactic acceptability, and it clearly relates to the concept of deep structure as outlined above.

The idea of 'lexical acceptability' was introduced as it was felt that a distinction needed to be made between nonsense errors that, by definition, exclude semantic constraint and errors that may bear no apparent reference to the text but are semantically constrained at least to the extent that they are meaningful words. In effect this gives credence to the possibility, on the lowest score point of the scale, that the reader may have established some tenuous link with textual or extra-textual semantic information that is not apparent to the scorer. Quite frequently it also allowed a word that was acceptable at phrase level - but not at clause level - to be credited with at least 1.

Apart from these departures, the distinction between errors regarded as acceptable to the local context as opposed to general context (2 vs 3 or 4) appears in most other continuous semantic scales and has clear relevance and definitive potential.

ii) Syntactic Acceptability of Errors

The syntactic dimension may be defined as that information

on the particular sorts of error occurrence that were regarded as acceptable for scale score 3 (see chapter IX).

The concept of the 'message' of the text has already been amplified. In this context it serves merely to crystallize the difference between semantic and syntactic acceptability, and it clearly relates to the concept of deep structure as outlined above.

The idea of 'lexical acceptability' was introduced as it was felt that a distinction needed to be made between nonsense errors that, by definition, exclude semantic constraint and errors that may bear no apparent reference to the text but are semantically constrained at least to the extent that they are meaningful words. In effect this gives credence to the possibility, on the lowest score point of the scale, that the reader may have established some tenuous link with textual or extra-textual semantic information that is not apparent to the scorer. Quite frequently it also allowed a word that was acceptable at phrase level - but not at clause level - to be credited with at least 1.

Apart from these departures, the distinction between errors regarded as acceptable to the local context as opposed to general context (2 vs 3 or 4) appears in most other continuous semantic scales and has clear relevance and definitive potential.

ii) Syntactic Acceptability of Errors

The syntactic dimension may be defined as that information

that may be derived from the cumulative syntactic structure of the text as interacted with the syntactic competence of the reader. It may include information at the level of the morpheme, the word, the phrase, the clause, the sentence or at the inter-sentence, passage level. At whatever level, the information is basically constituted out of the reader's awareness of what he has read - the text or his own sequence of errors - and the constraints that this sets, in terms of his syntactic competence, on what may follow or on the acceptability of what he has read.

Development of a scale of criteria for the syntactic acceptability of errors was not as problematic as in the semantic dimension. As pointed out by Siler (1973), the syntactic dimension is supported by a more established theoretical and taxonomic structure such that criteria are more readily definable. The range of criteria adopted has once again varied from Goodman's (1969) ten-point scale within a complex taxonomic structure of levels of processing; through Burke (1976, 1977) and Potter's (1980) five-point scale; to, again, a single criterion: the acceptability of the error to preceding syntactic structure (Cohen, 1975) or, alternatively, the acceptability of the error within the same grammatical class of word as the stimulus word (Clay, 1968; Kolers, 1969). Weber, (1970a) made separate analyses using each of the single criteria mentioned above as well as a third: the acceptability of the error within the syntactic structure of the sentence as a whole (i.e., preceding and following context). The following scale

as used in experiment II is, with some refinement, the same as that developed by Donald (1979a; experiment I, chapter VI).

Scale for Evaluating Syntactic
Acceptability of Errors

<u>SCALE</u>	<u>CRITERION</u>
4 :	<p>E is syntactically acceptable to both the preceding <u>and</u> the following structure of both its sentence and the passage¹.</p> <p>e.g.</p> <p>E up</p> <p>S They ran <u>down</u> the hill.</p>
3 :	<p>E is syntactically acceptable to both the preceding <u>and</u> the following structure of its sentence <u>or</u> clause but not to the passage.</p> <p>e.g.</p> <p>E run</p> <p>S They <u>ran</u> down the hill.</p>

¹ 'Passage' acceptability usually relates to tense, number or person acceptability at an inter-sentence level.

only to the sentence but to the passage as a whole - is a legitimate extension of this ordering. Burke (1976, 1977) and Potter's (1980) scales of acceptability criteria are essentially similar including, as they do, preceding context, sentence context, and passage context. They omit reference to function, however, and introduce a distinction between an error that is 'acceptable within the sentence but not the passage' (scale point 3) and an error that is 'acceptable with subsequent context' (scale point 2) (Burke, 1976, p.36). This distinction seems somewhat anomalous as extremely few errors are constrained by the subsequent context of the sentence and not by the preceding context as well (i.e., the sentence as a whole). Hood and Kendall (1975) and Hood (1976) also make distinctions (although not on a scale) between preceding context, sentence context and passage context. Precedent, in other words, for the criterion structure and ordering of the present scale is, with minor modifications, regarded as acceptable. The only real departure, apart from the specific ordering, is the allowance that an error may be acceptable to sentence or clause (scale points 2 and 3). Where compound sentences are involved this appears to be justified.

At this point it should be mentioned that Potter (1980) introduces a serious reservation about the validity of measures of syntactic constraint. He points out that the evaluation of many errors as syntactically acceptable could be confounding their supposed syntactic constraint with

what is possibly orthographic constraint. Thus, a word such as 'despairing' misread as 'disappearing' might reflect syntactic constraint in its context or it might reflect orthographic constraint through its beginning, general letter similarity and, particularly, its ending. This possibility was tested through requiring subjects to read a contextual passage as well as a list of words, out of context, containing the same words as in the passage. Recency and sequence effects were controlled. The finding was that although there were significantly fewer errors on the passage than on the list (indicating use of contextual information) there was a non-significant difference between the measure of syntactic acceptability of errors on the passage (where syntactic information was available) and on the list (where no syntactic information was available but where errors had been judged as if it were available). This lack of significant difference applied to both good and poor readers. (On the other hand, the difference was significant for good readers where the semantic acceptability of errors was concerned but not for poor readers). These results are suggestive but on the grounds of several important reservations they cannot be taken as conclusive. On Potter's own admission (op cit), the instructions given might have misled the children into reading for accuracy and therefore to over-reliance on orthographic information in the contextual condition. Equally, the children were reading well into frustration level (86,8% average accuracy). Smith has said, 'the more difficulty a reader has with reading, the more he relies on the visual information'

(Smith, 1971, p.221). The research of Williamson and Young (1974) also suggests that errors are more orthographically constrained and less contextually constrained at the frustration level of reading than at instructional level. It is very probable, therefore, that contextual, and particularly high level syntactic, constraints were not operating optimally under the 'contextual' condition. Not mentioned by Potter, but nevertheless significant are several other considerations. First, it is questionable whether the recognition of a morphemic unit such as as inflexional ending is, in context, orthographically constrained or whether it reflects, rather, orthographic redundancy within the reader's awareness of highly predictable syntactic markers. Syntactic information can operate at an intra-word, morphemic level (Goodman, 1969) and particularly where the higher levels of syntactic constraint are excluded from the 'contextual' condition, it is quite possible that there was apparent similarity in errors on the word list and errors on the passage but for different reasons. Second, Potter's results (op.cit.) do demonstrate a consistent, if not significant, difference in favour of the use of syntactic constraints under the contextual condition. This consistent trend could quite conceivably have reached significance if the instructions to readers and the level of difficulty had been different. Finally, in terms of the present study, Potter specifically mentions that no illustrations were present (usual in the Neale (1958) passages) and it could be argued that this too might have reduced the contextual relevance of the

passages in comparison with the lists.

The conclusion must be that although there might be a degree of overlap between orthographic and syntactic information at the morphemic level - predictable in terms of redundancy anyway - the evidence is hardly sufficient to reject the validity of the syntactic acceptability measure.

iii) Orthographic Acceptability of Errors

The orthographic dimension may be defined as comprising information that may be derived from the total graphic display in interaction with the reader's knowledge of the orthographic structure of the language. This may operate at the sub-letter (distinctive feature), letter, letter-group, word or even phrase level. The information includes graphic featural cues (not only 'letter' cues) together with knowledge of structural regularities that refer not only to equivalent sound or articulatory patterns for particular orthographic features but that also reflect the likelihood of certain orthographic features occurring together or in particular sequences.

The most sophisticated and sensitive scale for measuring the degree of orthographic constraint exhibited in an error was developed by Weber (1970b). She developed a formula which took account of, and weighted, a variety of factors that enter into orthographic constraint:

$$\text{"GS (Graphic Similarity)} = 10 \left(\frac{50F+30V+10C}{A} \right) + 5T + 27B + 18E$$

- where:
- F = the number of pairs of adjacent letters in the same order shared by S and R.
 - V = the number of pairs of adjacent letters in reverse order shared by S and R.
 - C = the number of single letters shared by S and R.
 - A = average number of letters in S and R.
 - T = ratio of number of letters in the shorter word to the number in the longer
 - B = 1 if the first letter in R is the same as the first letter in S; otherwise B = 0.
 - E = 1 if the last letter in R is the same as the last letter in S; otherwise E = 0" (Weber, 1970b, p.155-156).

Apart from the intuitive validity of this formula, Weber (op cit.) quotes a correlation of .93 between rankings by college students of a random list of errors and scores on the basis of the formula. The problem with the formula is that it is extremely tedious and time-consuming to apply and the chances of reliability being affected (the degree of concentration required is intense!) are high.

Other alternatives to this formula have been developed but are either too imprecise or too insensitive by comparison. Goodman (1969), for instance, suggests a ten-point scale but his criteria are not sufficiently explicit - other than referring in a general way to the relative weighting of beginning, end and middle similarities as well as general configuration - and the important

factor of sequence is not mentioned. Cohen (1975) developed two scales, one for nonsense substitutions and another for meaningful substitutions. The reasoning behind this was rather obscure and the two five-point scales somewhat arbitrary. Biemiller (1970) and Hood and Kendall (1975) both judged orthographic acceptability on the sole basis of whether the first letter was in common or not.

On these grounds there seemed to be a case for developing a scale that could include the essential factors from Weber's (1970b) formula, that would be simple and more reliable to apply, and that would not have the arbitrariness of other alternatives. The following scale, therefore, was developed as a refinement of the original attempt to meet these requirements in Donald (1979a; experiment I, chapter VI).

Scale for Evaluating Orthographic
Acceptability of Errors

SCALE

CRITERION

- 5 : E and S contain identical letters, in the same sequential position
- e.g. E wind
S wīnd
- (i.e. homographs; some mispronunciations; or some nonsense words).

4 : E differs from S by one letter, all remaining letters being in the same sequential position.

e.g. E waited

S wanted

OR

E and S contain identical letters, but in differing sequential position.

e.g. E there

S three

3 : E differs from S by two letters, all remaining letters being in the same sequential position

e.g. E open

S over

OR

E differs from S by one letter, with any of the remaining letters in differing sequential position.

e.g. E wild

S wide

2 : E differs from S by two letters, with any of the remaining letters in differing sequential position.

e.g. E his

S the

OR

E and S have first and second letters in common.

e.g. E there

S through

OR

E and S have first and last letters in common.

e.g. E called

S could

1 : E and S have first letter in common

e.g. E but

S because

0 : E fits S on none of the above criteria

e.g. E afraid

S terrified

Note: Sequential position is in all cases taken from the beginning of the word concerned: i.e. forward sequence

This scale includes the widely accepted criterion that beginnings are orthographically more constraining than endings and endings more than middles (Goodman, 1969; Weber, 1968; 1970b; Marchbanks and Levin, 1965; Shankweiler and Liberman, 1972). In addition, however, it includes the important factor of forward sequence (Weber, 1970b) and interacts this with the number of letters in common so that errors that are either scrambled in letter sequence or are of differing length or letter composition to the stimulus word are relatively penalized. The reason why the 'beginning and ending' criterion enters so late in the scale is that all errors that meet the criteria for 5, and most for 4 and 3 are relatively highly constrained so

reader. In so far as illustrations may provide an enriched source of contextual information - particularly of semantic information - the degree to which the semantic and, to a lesser extent, the syntactic strategies of information usage are facilitated by this information will reflect the actual contextual value of the information. Conversely, the degree to which this is balanced by a strategy involving decreased reliance on orthographic information will reflect the predicted effect under the contextual hypothesis. In the final analysis, what matters is the balance between the three strategies of information usage and whether the balance achieved under illustrated reading may be regarded as more efficient in message identification than that achieved under a no-illustration condition.

iv) Self-corrections

Unlike the above three areas of strategy that need to be inferred from a linguistic analysis of the reader's errors, self-corrections represent direct evidence of a strategy: whether the reader characteristically detects his own errors and sets them right. Thus, where the process of reading is regarded as constructive - as involving the reduction of uncertainty through the selection of only minimum cues within the redundancy of available information - it is clear that errors must necessarily occur, more or less often, as part of this process (Smith, 1971). That an efficient reader must develop an effective strategy for recognising when he has made a substantive error and for

that the more obvious constraints, such as having first letter in common, are met anyway. Exceptions to this are often the shorter words (who/how; was/saw) where orthographic constraint is apparent without the first letter criterion being met. It is not absolute; and the scale as it is constructed allows the relevant flexibility.

Using Weber's formula as a criterion of validity, a correlation of ,87 was obtained between formula scores and scores on this scale over 30 errors selected at random from the transcriptions of experiment II¹. Apart from its validity, therefore, the scale meets the additional requirements of being explicit and relatively simple to apply (see inter-judge reliability: chapter IX).

In testing the effects of illustrations under the contextual hypothesis, the three strategies of 'use of semantic information'; 'use of syntactic information' and 'use of orthographic information' as inferred from the relative acceptability of a reader's errors on each of these dimensions, are critically relevant. This is particularly true for the semantic and the orthographic strategies. The contextual hypothesis (Chapter II) states that for optimally efficient reading, the degree of orthographic information required to be selected and processed is in inverse proportion to the degree of contextual information available and accessible to the

¹ Biemiller (1970) claims that the single criterion of first letter in common is, in practice, as valid as Weber's more complex formula. A correlation of ,64 was obtained between this criterion and formula scores on the same 30 errors. This lends some credence to Biemiller's claim but also indicates that considerable information is lost if only the single criterion is used.

correcting that error (re-selecting from the available information) is equally clear. In the context of early reading, where the reader has reduced access to the available information, it may be hypothesised that such a self-correcting strategy is central to optimal development of the reading process.

Clay (1969) made a specific study of self-correction behaviour in early readers over their first year of school.

She observed that a self-monitoring strategy - an awareness that 'something was wrong' when an error had been committed - emerged very early in the process of learning to read.

Under these circumstances children frequently 'stopped, looked puzzled, complained, repeated the line, or ran a finger along a word' (Clay, 1969, p.48). This behaviour, and the comments of the children, served to confirm that self-correcting developed out of an awareness of informational contradiction: an awareness that the error response was not in accord with the total informational structure - semantic, syntactic and orthographic. Thus, 'a reader may become conscious of a difference between what he has said and one of the several messages from the text, and experience feelings of dissonance' (op cit. p.53). Such behaviour often, although not invariably, resulted in a spontaneous correction of the initial error.

Such self-corrections, then formed the basis of an analysis in which Clay (op cit.) demonstrated a significant difference ($p < .01$) in the frequency per error of self-corrections for a high progress group of readers as opposed

to a low progress group. This result supports the notion that the efficient use of a self-correcting strategy (in so far as this is reflected in actual self-corrections) is linked to the optimal development of the early reading process. Both Weber (1970b) and Cohen (1975), although not specifically concerned with the nature of self-correction, nevertheless found the same basic relationship. Weber, focussing on the use of syntactic constraints, noted that good readers in the first grade left uncorrected only 15% of those errors which upset the grammaticality of the sentence, while poor readers left 58% of such errors uncorrected. Cohen noted that self-corrections, as one of his categories of word substitution, increased substantially for good readers as they progressed in their first year of reading instruction while, for poor readers, the increase over the same period was minimal.

Samuels, Begy and Chen (1976) compared adults and fourth grade readers as well as good and poor readers at the fourth grade on a 'word recognition strategy test'. This involved presenting readers with a linguistic context and partial letter cues to a given word and measuring the success with which the target word was predicted. Apart from finding that good readers were consistently and significantly ($p < .001$) better than poor readers at predicting the target, they found 'that more fluent readers (were) ... superior in awareness when a false recognition had been made' (op cit. p.72). Although not a direct reflection of self-correction this supports the idea that detection of 'dissonance' as a component of the strategy of self-

correction is related to optimal reading development.

Finally, Hood and Kendall (1975) were interested in the differences between reflective and impulsive second grade readers. One of the few areas in which a significant difference was demonstrated was that of self-correction where reflective subjects made more self-corrections than impulsive subjects. Although this difference between reflective and impulsive subjects does not parallel the difference between good and poor readers (reflective and impulsive subjects were not significantly different on other criteria of reading competence) it may yet allow that some types of self-correction are effective while others may be redundant: Reflective subjects may make more self corrections, not all of which are necessary to meaningful message identification, while impulsive subjects may make less self-corrections, proportionately more of which are necessary. It may also suggest that where reflective subjects make overt self-corrections, impulsive subjects may be making more covert self-corrections: they may not stop and articulate their correction but they may well modify, covertly, the information from the error in terms of their cumulative interpretation of the text. This would certainly fit with a model of more rapid and impulsive reading.

This possible difference between overt and covert self-correction is important. Clay (1969) herself points out that actual self-corrections may represent only the 'tip of the iceberg'. They reflect the strategy of self-correction

only in so far as this is overtly observable: covert self-correction, on the other hand, may have taken place on correct responses before they are finally articulated - or even on incorrect responses, as above - but has simply not been observably manifested. Moreover, it is likely, although there is no direct evidence for this as yet, that covert self-correction would increase proportionally in relation to overt self-correction as the process of reading develops. The evidence that has been presented, however, clearly indicates that in the early stages of reading development, at least, actual, observable self-corrections reflect a positive strategy that is related to optimal reading development.

At this point it should also be mentioned that Thompson, in a recent article (1981) has criticised the measure of self-correction developed by Clay (1969). He argues that Clay's results on the difference between high and low progress readers on the rate of self-correction is merely a reflection of the difference between these groups on the number of uncorrected errors. In other words, since the rate of self-correction is based on the number of self-corrections expressed as a proportion of the total number of errors (corrected and uncorrected), he shows (on Clay's data) that while the actual number of self-corrections remains more or less constant, the number of

uncorrected errors increases from high to low progress¹. This, he argues, is evidence that the measure of self-correction is merely expressing variation in the number of errors. Two points in this argument are fallacious. First, the actual number of self-corrections is a meaningless measure since a reader who makes only 5 errors, for example, and corrects 2 of these cannot be meaningfully equated with a child who makes 20 errors and corrects only 2 of these. The lack of variation in the actual number of self-corrections over high and low progress readers, therefore, says nothing more, when related to the variation in uncorrected errors, than that the rate of self-correction

¹ Thompson, in fact, argues not in terms of absolute numbers but in terms of proportions of the total number of words read (as, he claims, does Clay in her calculation of self-correction rate). Thus the formula for rate of self-correction is given as:

$$\frac{p(c)}{p(c)+p(b)}$$

where: p(c) is the number of self-corrections as a proportion of the total number of words read.
 : p(b) is the number of uncorrected errors as a proportion of the total number of words read.
 (Thompson, op cit., p.228)

It is readily demonstrated, however, that total number of words read is arithmetically redundant in this formula:

$$\frac{\left(\frac{c}{a}\right)}{\left(\frac{c+b}{a}\right)} \quad \text{where: } a \text{ is the total number of words read}$$

Thus, the simple proportion, $\frac{c}{c+b}$, is a more legitimate formula for calculating rate of self-correction and is used as such in experiment II, chapter IX, (as also in Donald, 1979a; experiment I).

does vary. This is what Clay has claimed.

Second, although variation in the rate of self-correction is related to the relative word identification accuracy of high and low progress readers, this need not imply that self-correction is merely an artifact within this. Since it is expected that low progress readers should make more errors than high progress readers on the same reading material, a null hypothesis on rate of self correction within this would state that the more errors there are the more self-corrections there should be. This was, however, clearly refuted in the significant difference between rate of self-correction for high and low progress readers. Thompson also argues that because self-corrections may be covert the obtained difference is meaningless. However, it is logically likely that the more developed readers would make more covert self-correction than the less developed readers and therefore, if anything, the obtained difference in self-correction rate would be likely to be even greater could covert self-corrections also be recorded.

Finally, findings such as Weber's (1970b) above, that are not based on the contended proportion, also endorse Clay's conclusion on the relationship between self-correction and optimal reading development.

Thus, despite Thompson's criticism, self-correction, as reflected in the proportion of errors that are spontaneously detected and corrected, was recorded as a legitimate measure of reading competence and was selected as a central strategy variable in this study.

In relation to testing the effects of illustrations under the contextual hypothesis, the strategy of self-correction is seen as reflecting the degree to which information in illustrations provides additional contextual information against which 'dissonance' in a particular response may be gauged, detected and observably, at least, set right. If self-corrections are increased under these conditions then this can be taken as evidence that the presence of illustrations in early reading material facilitates a strategy of self-correction that, as has been shown, is related to optimal reading development.

Comprehension

The task of selecting measures of reading comprehension is more complex and less clearly defined than is the case for oral reading performance. Where the latter involves a given, observable and explicit behaviour, the former refers to 'understanding', or the idiosyncratic interpretation that a reader derives from that behaviour. Not only is the concept of understanding, itself, more elusive but the behaviours by which it may be measured are neither given (a wide variety of behaviours may be said to reflect understanding) nor are they easily evaluated. As Gibson and Levin (1975) have said;

"We know far less ... about factors that influence comprehension of sentences and longer passages of discourse in reading than we do about factors that influence recognition of individual words. Why should this be the case? Two reasons stand

out. It is harder to define and segment into units the information contained in a passage of discourse; and second, we have made little progress towards a theory of comprehension" (op cit., p.392).

In the present context the theoretical issue that is of most direct concern is whether comprehension should be regarded as a unitary skill or as a set of separately identifiable sub-skills. As indicated earlier, the validity and relevance of including some measure of comprehension is above question: what is in question is whether there should be one measure or several - and if several, exactly what areas or levels of comprehension are most relevant to testing the effects of illustrations under the contextual hypothesis.

Perhaps the most definitive study to date on the question of sub-skills has been that of Davis (1968). He constructed a range of comprehension items (192 items in all) which were divided into eight sub-tests each designed to measure a specific comprehension sub-skill. Having applied these tests to nearly 1 000 college students, factor analysis of the data led him to conclude that there were four, rather than eight, reliably distinguishable sub-skills:

1. Identifying word meanings;
2. drawing inferences from the content;
3. identifying the writer's purpose, attitude, tone and mood;
4. finding answers to questions asked specifically or in paraphrase.

The last of these is somewhat unhelpful. However, Spearritt (1972) using a more sophisticated procedure on the same data analysed out

similar factors but identified the last as 'the ability to follow the structure of a passage' which, psychologically, makes more sense than 'finding answers to questions'.

Lunzer, Waite and Dolan (1979), however, in a more recent and equally thorough factor analytic study found no evidence of distinct comprehension sub-skills. They conclude;

"... that individual differences in reading comprehension should not be thought of in terms of a multiplicity of specialised aptitudes. To all intents and purposes such differences reflect only one general aptitude: this being the pupil's ability and willingness to reflect on whatever it is he is reading" (op cit., p.64).

At this point one can only conclude that the issue is unresolved. For the purposes of selecting relevant measure(s) of comprehension however, the sub-skill alternative is favoured if only in that favouring the unitary concept would have necessitated the selection of a single variable that would necessarily have excluded all other possibilities. It can also be argued that this study is concerned with the nature of information and its usage and if the sub-skill factors are seen as various orders of task, then it is conceivable that different sorts of information might facilitate different orders of task. There is even some oblique evidence for this in the Lunzer et al (1979) study where performance on the different putative sub-skills differed in some cases quite markedly between the different stories presented for comprehension.

Starting from the four areas identified by Davis (1968), with Spearritt's (1972) modification, it was decided to exclude the third factor (identification of the writer's tone, mood etc.) on the grounds that this was seldom expected as a task at the level of reading being considered. The first factor (identification of word meanings) was felt to be partly tapped in the accuracy of message identification. Apart from this it was felt, along with Lunzer et al (1979), that "there are good psychological grounds for saying that this is prior knowledge which the reader brings to the comprehension task" (p.65) and not effectively an outcome of the task. This left the two areas of 'Inference' and 'Following the structure of the passage'.

If 'following the structure' is interpreted as a task involving, basically, a literal understanding of the content of the text, then the distinction between this and the task of 'inference' is further justified on the grounds of its relevance to the illustration as information issue. The distinction between literal and inferential comprehension tasks has been widely, if somewhat uncritically, accepted (Dechant and Smith, 1977) and is based on an informational distinction where literal comprehension is seen as dependent on the identification of information that is explicitly stated in the text while inferential comprehension is dependent on the identification of logical relationships that are only implicit in the textual information.

As Flood and Lapp (1981) point out, the distinction cannot be regarded as absolute as readers undoubtedly engage in

a degree of inferential processing whatever the task. Nevertheless, a distinction based on relative degrees of dependence on explicit and implicit information is probably still justifiable. It can be argued, then, that contextually relevant illustrations provide a source of information that is complementary to (redundant within) the explicit information in the text. Moreover, the spatial dimension that necessarily dominates illustrative information does not preclude, but certainly limits the sequential, logical form of processing required in an inferential task (Paivio, 1971). There is some experimental evidence for this in that Vernon (1953) found that illustrations facilitated the retention of factual information directly depicted in illustrations but that the illustrations had no effect on the retention of 'relational' information. In testing the effects of illustrations, therefore, the distinction is relevant in that it may operate. In other words, where there is an argument that the effects of illustrations may be different on the two levels of comprehension task, that argument can only be tested on the hypothesised distinction.

i) Literal comprehension (Following the structure of a passage): Experiment II

A wide variety of items encompassing different aspects of this basic task requirement are conceivable. Barrett's taxonomy of the cognitive dimensions of reading comprehension (in Clymer, 1972) was, with some flexibility, used as

the basis for item construction. Constraints that limited the final choice of items included the necessity to have parallel items over three different stories at two levels of reading difficulty (see chapter IX) as well as the practical requirement that the number of items and their level of difficulty should be possible for the youngest subjects (mean C.A.: 6 years 8 months) without too high a level of frustration or loss of concentration developing.

For each story at each level, therefore, four items were constructed within the task requirement of following the structure of the passage.

a) Recall of detail¹

The task requirement, here, was that the reader should produce, from memory, an item of information as explicitly stated in the text.

e.g. Q: Where was the donkey standing when the boys first saw him?²

¹ Since recall is not free but based on a specific question, this is technically 'cued recall'. As Gibson and Levin (1975) emphasise, free recall is problematic as a measure of comprehension while cued recall is theoretically more justified (Guttman, Levin and Pressley, 1977).

² Full criteria for the scoring of answers are set out in chapter IX. The examples given here are from one story at the lower level of reading difficulty. Questions at the higher level required the same information but differed slightly in wording. All questions for all stories are set out in Appendix 3.

b) Recall of cause-effect relationship

The reader was required to produce, from memory, the cause, as explicitly stated in the text, to a specified effect.

e.g. Q: Why did the donkey run away from Roger?

c) Summarizing the main idea

The task here, and for the following item, required more than simple, direct recall. In terms of Barrett's taxonomy (Clymer, 1972), they both fall at his level of 're-organization' - i.e. above 'literal comprehension'.

However, in so far as they both depend on explicit textual information, they were included under 'Literal' in the present context. They also reflect, directly, the task requirement of following the structure of the passage.

For 'Main Idea', the reader was required to summarize a set of ideas or a sequence of events central to the meaning of the passage and as explicitly stated within it.

e.g. Q: How did Roger manage to get onto the donkey to ride him?

d) Grouping related ideas

The reader was required to group and summarize a set of ideas or events as explicitly stated in the text but not necessarily sequentially related or explicitly separated out in the text.

e.g. Q: What happened in the first part of the story?

ii) Inferential comprehension (Drawing inferences from the content): Experiment II

Only one item was constructed at this level of task. The reasons for this were, first, that inference is a more narrowly defined task requirement than following the structure of the text, and second, practical considerations, such as the nature of the stories and the sorts of question that younger subjects could cope with, limited the questions that could be asked.

The task required the reader to draw a conclusion about a character in the story where this conclusion was not explicitly stated in the text but where evidence for the conclusion could be inferred from the actions and motives of the character concerned.

e.g. Q: Who do you think is the cleverest of the two boys? Why?

In conclusion, comprehension as a measure of the effectiveness of the reading process is regarded as essential in testing the effects of illustrations. A distinction between literal and inferential comprehension is upheld on the grounds that the information in illustrations may facilitate the former but not the latter.

For experiment I, the same essential distinction between literal and inferential comprehensions was structured into the relevant tasks. However, for 'literal' comprehension

free recall was used as opposed to cued recall and the measure of literal comprehension was taken as the proportion of ideas, explicitly stated in the text, that the reader recalled (Chapter VI). This system was deliberately altered for experiment II on the grounds that free recall is a less reliable and specific measure of comprehension than cued recall.

CHAPTER V

ILLUSTRATIONS AND READING:

RESEARCH REVIEW

As indicated in the introduction, research on the effects of illustrations on reading can be seen as basically falling into two contexts. The first refers to research that pertains either specifically to the focal attention hypothesis or, less specifically, to the paradigmatic assumptions of that hypothesis. The second refers either to the limited research that pertains specifically to the contextual hypothesis or to the larger body of research that, in one sense or another, fulfills the paradigmatic assumptions of the contextual hypothesis. A third context, however, also exists but is less relevant in that it is not concerned with elucidating the effects of illustrations on reading per se. This refers to the research on word and image interaction and the question of semantic accessing (Pressley, 1977).

1. The Focal Attention Hypothesis

Samuels first formulated what has subsequently been called the focal attention hypothesis in his study of 1967. In his words;

"The purpose of the present study was to test the hypothesis that when pictures and words are presented together, the pictures would function as distracting stimuli and

interfere with the acquisition of reading responses. Pictures may be used as prompts when the reader cannot read a word in the text, but pictures may miscue and divert attention from the critical task of attending to the printed words" (Samuels, 1967, p.337).

In his review of research on the effects of pictures on reading (Samuels, 1980) he goes on to develop a more specific interpretation of findings within the paired-associate paradigm:

"Attentional processes and the principle of least effort explain why pictures interfere with learning to read. The learning task set up for subjects in these studies was essentially a paired-associate task, i.e., learning to associate a common English word with a printed stimulus. The stimulus that was presented to the subjects was complex; it consisted of a picture, which could elicit the correct verbal response by itself, and a printed verbal stimulus, which could not elicit the correct response when first presented. Since the printed stimulus could not elicit the correct response at first, the function of the picture, from a teacher's point of view, was to prompt the correct response. The problem of getting the child to learn to read the word is one of shift in stimulus control, from the picture to the printed stimulus... However, given two stimuli, one which can easily elicit the correct response (the picture) and one which can not (the printed stimulus), the principle of least effort operates.

The principle of least effort is that

when a complex stimulus is presented to a subject, he will select that aspect of the total stimulus which most easily elicits the correct response" (Samuels, 1970, p.400)

Although not specifically related to the effects of illustrations, the view of attention as a limited capacity system is further articulated in La Berge and Samuels (1974). In this model of the reading process limited attention capacity is seen as deployed in specific, sequentially constrained shifts, the effectiveness of the whole process being dependent on sub-attentional (automatic) steps in the sequence that do not, by definition compete for the limited capacity of attention. In this model, the process of reading development is seen as essentially mechanistic, involving the acquisition of a hierarchy of associational responses that, progressively, can be processed at an automatic level. It is within this framework that the focal attention interpretation of the effects of illustrations has its theoretical justification. In other words, reading development is effectively the acquisition of specific - particularly grapho-phonetic - associations to the degree that these become capable of automatic processing. In the sense that attention is a limited capacity system and is seen as deployed in a complex and competitive sequence of processing steps, anything that distracts attention under this conception must necessarily be seen as interfering with the very process of reading development.

The 'classical' paradigm that has been used to test the hypothesis has consisted of the following steps and conditions (Arlin et al., 1978-79; Singer, 1980).

1. The reader is presented with reading material (single words, sentences or a full passage) with or without a picture.

2. The reader is instructed to attempt the reading.

3. "Students either get feedback of 'right' for correct responses or the experimenter correctly identifies the word for incorrect responses" (Singer, 1980, p.293). In addition, if no response is forthcoming after a period (usually 4 to 7 seconds) then the correct response is given by the experimenter.

4. Once the acquisition trial on all the material to be read is complete, the reader is tested on his recognition of single words (without the picture) either as in the original list or as selected from the sentences or passage.

5. Since a trials-to-criterion learning paradigm is used, acquisition trials are usually interspersed with test trials. The typical dependent variables are time-to-criterion, number of test trials to criterion, number of words correct on acquisition and test trials, and post- or retention-tests.

Samuel's original study (1967) consisted of two experiments. The first was conducted on a group of pre-first grade children who were required to learn to read four

words not known to them before the experiment. The procedure was as outlined above. Three treatment conditions were introduced: no-picture; simple-picture (outline drawing of referent); and complex-picture (colour representation of more than the referent). On acquisition trials, the mean accuracy for no-picture was 25,30; for simple-picture 39,40; and for complex-picture 36,90. The differences of both simple- and complex-picture from no-picture were significant. On test trials the reverse was true: mean accuracy for no-picture 19,20; for simple-picture 11,30; for complex-picture 11,60. Again both comparisons were significant. This was taken as direct confirmation of the hypothesis.

In the second experiment, an attempt was made to approximate 'classroom learning conditions'. This was conducted on mid-first grade children divided into good and poor readers on the basis of a pre-test. The reading material was a story called Fun at Blue Lake, 106 words long, containing 50 different words, and accompanied, in the picture condition, by a single picture depicting a lake, woods, a cabin and a family at the lake shore. The post-test was the same as the pre-test and contained a list of the 50 different single words in the story. The learning procedure consisted basically of preparation (background), silent reading and oral reading. During this procedure any words that were not known were verbally prompted. For the good readers, the mean post-test accuracies were 42,08 (no-picture) and 43,15 (picture); non-significantly different. For the poor readers, on the other hand, the accuracies were 26,23

(no-picture) and 23,69 (picture); significant $p < .01$. The results were interpreted as confirming the hypothesis, under more natural reading conditions, for poor readers who were seen as more subject to distraction than good readers.

Apart from more general observations about the paradigm that will be introduced below, several observations particular to the second experiment need to be made. First, one picture per 106 words of text is far from representative of normal reading conditions in the first grade. Even conservative counts based on 'old-fashioned' first grade basal readers average at approximately 1 : 50 (see chapter IX). The amount of information that one picture can carry is limited and the ratio must therefore be crucial in determining its information value over its distraction value. Second, the nature of the illustration sounds, from its description, to be particularly 'static'. There is no norm for this, but a picture that relates to actions and events (Fun at Blue Lake?) is likely to carry more meaningful information than a 'scenic' one as described. Both of these observations relate to the third of the factors outlined in the introduction as influencing the effects of illustrations on reading; the relationship of illustration to text. Finally, it is anomalous, given the experimental condition and the purpose of the experiment, that the post-test should only have consisted of a word list. The final accuracy, after the learning trials under different treatment conditions, of story reading would surely be the most natural classroom criterion - not

to mention comprehension.

Harzem, Lee and Miles (1976) replicated Samuel's first experiment on slightly older children (mean C.A. 6,5) but introduced four treatment conditions: (i) a picture of an object named by the word, (ii) a picture of an object unrelated to the word, (iii) a nonsense picture not resembling any object, and (iv) no picture. They also contrasted massed and distributed practice during the alternating acquisition/test trials, and gave a retention test 28 days after the last test. Apart from this the experiment followed the paradigm.

The results showed a similar pattern to the Samuel's (1967) experiment for the acquisition trials with the mean accuracy for the appropriate picture being higher than for the other three conditions. On test trials, the no-picture condition was significantly higher than the nonsense and appropriate picture. On the retention test, the data were evidently too little to statistically analyse but a similar pattern as on the test trials was evident. Differences between massed and distributed practice were of no real consequence. The conclusion was that an appropriate picture facilitates acquisition most but 'learning to read' least while the reverse is true of no-picture.

Willows (1978a) carried the investigation further within the same paradigm by looking at the effects of illustrations on the speed and accuracy with which words that had previously been learned were recognised. He also worked with slightly older children, 2nd and 3rd graders, and included the

variable of reading ability. Three treatment conditions were investigated: (i) no-picture; (ii) related-picture (the picture was within the same semantic category but not directly representing the object named - e.g. word, 'cat'; picture, 'dog'); and (iii) unrelated-picture (the picture represented a different semantic category - e.g. word, 'cat'; picture, 'lemon'). He also investigated the relative positioning of the picture, behind or adjacent to the word, but this produced no variation of any consequence. Apart from a prior training period where the experimental words were learned, the procedure followed the paradigm, but without, obviously, interspersed acquisition/test trials. The dependent variables were recognition latency (time) and accuracy. In summary, he found that latency was significantly affected with unrelated-pictures interfering more than related-pictures and these more than no-picture. This pattern held over second and third-grade but was more marked for second graders. A correlation between reading ability (as previously measured on a standard test of accuracy and comprehension) and latency revealed a significant inverse relationship on both related- and unrelated-picture conditions. Again this was found over both grades but was more marked for the second graders. On accuracy of word recognition, it was found that compared with no-picture, related-pictures improved accuracy while unrelated-pictures interfered with it. This was significant at the second grade level but only partially significant (pictures behind) at the third grade level. The correlations between reading ability and accuracy under the various

treatments were low but revealed a similar pattern with the poorer readers being more affected than the better readers by the semantic distraction of unrelated pictures. The conclusion was that

"... pictures in the periphery do affect children's speed and accuracy of reading; the size of the interfering effect of the pictures depends on their relevance to the words printed near them; younger, less skilled readers are more susceptible to those influences" (Willows, 1978, p.261).

Again, a number of observations are called for. First, the finding that latency of recognition was affected under these experimental conditions need not be interpreted as interference with learning to read. The nature of the pictures was such that both related- and unrelated-pictures were directly at variance with the stimulus word. It is hardly surprising, therefore, that the dissonance thus created interfered with recognition time. The fact that better readers were less affected can simply be interpreted in their quicker development of an ad hoc strategy to ignore this dissonance as irrelevant to the task at hand. Second, on accuracy it was found that the related-picture facilitated performance relative to both no-picture and unrelated-picture. And this was with a semantically related but nevertheless different representation to the stimulus word! How much more might one expect a matching picture to facilitate accuracy? Ostensibly this contradicts the findings of Samuel's experiment I

(1967) and those of Harzem et al (1976). There is a difference, however. Willows was looking at the recognition of known words while the others were looking at the learning of unknown words. Even more important, Willow's measure of accuracy was taken with the pictures present and is thus most comparable with the measures of acquisition accuracy taken by the others. In Willows case, therefore, the related-picture could be acting as a contextual pointer (the semantic category) and, since the word was already known, a source of check for the response alternatives; i.e. redundant information of, admittedly, a low order. In the other situations, under acquisition conditions, the picture also facilitated recognition (acting, presumably, as a contextual pointer). It was only where the reader had nothing but the orthographic information to go on (test trials) that the picture was found to have distracted from the learning of these cues.

In other words, in all the results discussed so far, where contextual information is present in even a reasonably related form, the contextual hypothesis is supported. Contextual information in the form of illustration does facilitate word identification. In so far as the focalattention hypothesis sets its criterion at the effective learning of orthographic cues to isolated word recognition, it is also supported; although not by

Willow's results¹.

Briefly, since the pattern is now clear, several other studies that relate to the focal attention hypothesis, directly or indirectly, can now be mentioned.

In another experiment by Willows (1978b) on third grade children only, he found similar results to his previous study (1978a). In the second experiment the three experimental conditions were; i) no-picture; ii) identifying-picture (i.e. same as stimulus word) and iii) unrelated-picture. He also divided the words to be identified into easy, moderate and difficult (although all known words) and compared good, normal and poor readers. In the results for latency, the identifying picture facilitated performance for moderate and difficult words but not for easy words. The unrelated pictures, however, interfered with performance but significantly more for the poor readers than the normal or good readers. On errors, it was found that identifying pictures made no significant difference for normal and good readers when compared with no picture. The poor

¹ It is interesting, and somewhat disturbing, to find how results such as those of Willows become distorted and reified in support of a particular point of view. In Willows (1978b), he summarizes these findings as; '.. when children in second and third grade read a set of familiar words, their performance was less efficient in terms of both time and errors if there were pictures in peripheral vision' (p.837). This is simply not true. It was only the unrelated and semantically dissonant picture that increased errors: the reverse was true for the 'related' picture. The importance of not generalising on this issue without reference to such factors as the relationship of illustration to text (chapter I) is thus emphasised.

readers, however, made fewer errors with identifying pictures on the difficult words. As previously concluded, this demonstrates that pictures may improve word identification accuracy - at least for poor readers on more difficult words. Interpretation of the latency effect is obscure and, particularly with unrelated pictures, of doubtful relevance to normal reading conditions. The greater susceptibility of poor readers to this effect, however, could be interpreted as evidence of distraction.

Braun (1969), using the basic paradigm, tested a large sample of pre-school children on the effectiveness of sight-word acquisition under picture and no-picture conditions where learning and test trials were alternated. He found that children in the no-picture condition acquired the sight words in significantly fewer trials than children in the picture condition.

Kier (1970) presents the case, based on her observations of 189 poor readers aged from 7-11 years, that pictures as an aid to phonic recognition (the picture dictionary idea) are subject to various types of misinterpretation ranging from non-recognition of the object drawn to recognition of the object but inability to name it. She concludes that as specific phonic cues, pictures paired with words are of dubious value.

Other studies have shown no significant effect. Miller (1938) presented first to third graders with appropriate basal readers that were either illustrated or not. On a composite test that had a heavy emphasis on single word

or phrase recognition as taken from the basal readers, he found no significant difference between those who had been under the illustrated/unillustrated learning conditions.

Similarly, lack of significant differences between picture-word and word only instructional treatments were found by both Ollila and Olson (1972) and Kiraly and Furlong (1974) on pre-school childrens' learning of sight words.

Hartley (1970) working with beginning first-graders found a significant interaction between list similarity (similar or dissimilar phonically structured words) and treatment condition (word alone, word + picture, word + oral context) such that pictures facilitated the learning of dissimilar words but not similar words.

King and Muehl (1965), however, had found the opposite effect with similar aged children; that pictures have no significant effect when the words to be learned are dissimilar but that they facilitate sight word learning when the words are similar.

Finally, two studies (Singer, Samuels and Spiroff, 1973-1974; Denberg, 1976-77), will be discussed in detail in the next section as they both bear specific reference to the contextual hypothesis. However, they are both, also, of relevance to the focal attention hypothesis. In brief Singer et al's results support the focal attention hypothesis while Denberg's results are at variance with it.

At this point, it can be said that within the paradigm there is considerable evidence in favour of the focal

attention hypothesis in so far as acquisition of the orthographic cues to sight words is the criterion of effective learning to read. Even within this, however, there is a considerable measure of disagreement.

Further, several studies have challenged the hypothesis on the grounds of assumptions and conditions in the 'classical' paradigm. Montare, Elman and Cohen (1977) found, in the course of a replication of Samuel's experiment I, that there was a confounding of oral feedback with treatment condition which seriously brings the results into question. Because the picture successfully provides information for making a correct response in the acquisition trials, they found that the no-picture condition was getting 88% of the oral feedback. In other words, the 'no-picture' readers were having their responses corrected (or provided) with specific oral feedback (see the paradigm) that could be seen as specific re-inforcement of attention to the orthographic cues. By contrast the 'picture' readers were having the alternative contextual strategy re-inforced to a far greater extent.

Samuels (1977) has offered some cogent criticism of the Montare et al (1977) study, and the fact that test trials followed acquisition trials for each word, on its own, invalidates their results so that these will not even be considered. Nevertheless, although Samuels challenges the argument, there can be no doubting the specific evidence quoted above, and its implications.

Denberg's (1976-77) study, to be discussed fully in the

next section, also specifically questions the paradigm on the condition of oral feedback. Her results on test trials, where no oral feedback other than general encouragement during acquisition trials had been given, were significantly at variance with the predicted effects under the focal attention hypothesis.

Arlin, Scott and Webster (1978-79) have offered perhaps the most serious challenge. They identified two problems in the paradigm as well as the unresolved list similarity problem mentioned earlier (Hartley, 1970; King and Muehl, 1965). The two paradigm problems are the one mentioned above as well as the temporal sequence of picture cue first then verbal cue. Their study, therefore, aimed "to clarify the role of pictures in learning sight words by testing 3 dimensions: 1) the relative effectiveness of the simultaneous presentation of word with picture and word with voice, and the presentation of the word alone; 2) the relative effectiveness of word with picture, word with voice, and word alone as types of feedback; and 3) the effects of list similarity" (p.651). Using an appropriate design and procedure to disentangle these effects they found that on a measure of learning rate (number of correct responses on a retention test divided by the time it took to learn the words), there was a significant main effect for presentation medium such that word and picture was superior to word + voice and this, in turn, was superior to word alone. The main effect for feedback medium was not significant although the tendency was for word + voice to be most effective. Finally, there was a significant

effect for similarity, with dissimilar words learned more effectively than similar words but no interaction. Rank ordering of the effectiveness of the various treatment combinations clearly showed the most effective as picture + word presentation combined with voice feedback (i.e., the correct oral version of the word being provided whether or not the initial response was correct) - mean learning rate : 70,4 - while the least effective (at the bottom of 9 ranked alternatives) was word only presentation with no feedback whatever - mean learning rate : 1,0. They conclude:

"Some researchers might be tempted to explain the best condition within the framework of the focal attention hypothesis, because attention is focussed on the word without picture during voice feedback. But if voice feedback is so powerful, then voice accompanying the word both on presentation and feedback should be the most facilitative. However, this was not the case, as this condition produced almost the worst results (rank order 8; mean learning rate, 16,2). The focal attention hypothesis does not seem to explain these data" (Arlin et al, 1978-9, p.656).

Singer (1980) has replied to this study with a criticism of what he maintains are methodological and analytical flaws. It seems doubtful that the criticisms can override the extent and import of the findings, however. Nevertheless it would probably be wise to await a neutral replication - if this is ever possible! - before final judgement of the issue can be made.

In conclusion a number of summary observations need to be made:

i) In its pure form, on the 'classical' paradigm, the focal attention hypothesis has considerable support. Under the conditions of the paradigm, it would appear that illustrations have a distracting effect on the learning of orthographic cues to word recognition.

ii) This effect has been clearly demonstrated for pre-school and first grade children only. The effect at subsequent levels of reading development is as yet unclear.

iii) There is some indication that the effect is more marked for poor readers than for good readers.

iv) Serious criticism of the interpretation is based on the finding that the effect may not be due to pictorial distraction but to the differential oral re-inforcement that is provided under a no-picture condition.

v) Evidence of the facilitative effect of pictures on word identification during acquisition trials and less analytic attention to orthographic cues (as evidenced on test trials) is completely in line with the contextual hypothesis.

The next section will consider the evidence for the contextual hypothesis. It is important to stress, again, that the contextual hypothesis does not exclude the interpretation of the focal attention hypothesis. It merely makes it redundant within a wider and more representative view of what is involved in the process of reading development.

2. The Contextual Hypothesis

The theoretical basis for the contextual hypothesis has been fully developed (Chapter 2). It remains, however, to define it more precisely with specific reference to the effects of illustrations. In this specific context, therefore, the contextual hypothesis may be stated as follows

Where illustrations are sufficiently and relevantly related to continuous text, they constitute a source of contextual information that, in particular, complements and is redundant within the semantic dimension of information. In illustrated reading, therefore, the efficiency and accuracy of message identification and the effectiveness of comprehension will be facilitated through reducing the reader's reliance on orthographic information while increasing his use of the available contextual information¹.

As mentioned in the introduction, few studies have been specifically directed at testing this hypothesis. The first to be considered is a study by Singer, Samuels and Spiroff (1973-74) that set out with the specific aim of testing the focal attention hypothesis against the contextual hypothesis:

"On one side of the controversy, Samuels (1967)

¹ From this point on in the thesis, 'the contextual hypothesis' will refer to this version which now encompasses, specifically, the effects of illustrations.

found that in comparison with words alone, presenting pictures in association with words apparently interfered with acquisition of reading responses ...

In contrast to Samuel's focal attention hypothesis, Goodman (1965) formulated a linguistic or contextual hypothesis based on his demonstration that contextual constraints facilitated identification of words children could not recognise when the words were presented in isolation ...

This study was an attempt to resolve the focal attention versus context controversy" (Singer et al, 1973-74, p.557-8).

The subjects were 80 grade 1 and 84 grade 2 children. A 2 (grades) x 4 (treatments) factorial design was used where the treatment conditions were; i) word + picture; ii) word + no-picture; iii) sentence + picture; iv) sentence + no-picture. The task consisted of learning to identify the words 'cup, cat, bat, and bed' as written in an artificial alphabet. The procedure was otherwise according to the 'classical' focal attention paradigm with alternating acquisition trials and tests. All tests consisted of the target words, isolated, on separate cards. The dependent variables were; trials to criterion and accuracy on test trials. For both variables significant main effects were apparent for both 'grade' and 'treatment' with no significant interaction. The data were therefore combined for both grades. Appropriate testing indicated that on trials to criterion, the word + no-picture treatment required significantly fewer trials to criterion compared with each of the other treatments, and the word + picture

treatment required significantly fewer trials than sentence + picture. Other comparisons were non-significant. On test trial accuracy, the word + no picture treatment had significantly more correct responses than each of the other treatments. No other comparisons were significant.

The conclusion drawn was that;

"... the evidence does not support Goodman's (1965) contention that sentence context facilitates acquisition of correct responses ..."

and, in a footnote,

"... Goodman's results can be attributed to the process of reading, in which semantic and syntactic constraints can be used for predicting unknown words. With a minimum of sampling of the unknown words, a high per cent of accuracy in identifying them can be attained. The present study emphasises the process of learning in which focal attention on the unknown words is a prerequisite to such subsequent processes as discrimination, hooking-up responses to graphemes, and re-inforcement of correct responses. Thus, the Goodman-Samuels controversy can be at least partially resolved by realising that they are referring to 2 different, but inter-related processes: the reading process and the learning process" (Singer et al, 1973-4, p.564 : their emphasis).

Neither on procedure and interpretation nor on conclusion can this pass without comment.

First, on procedure and interpretation: the criticism levelled at the paradigm in the previous section, that the differential oral feedback to the word + no picture treatment

could account for the effect, holds here as firmly:
 the interpretation that word + no picture is least distracting from orthographic information is not necessarily valid: it may be because it gets the most specific oral information and feedback.

Second, a three word sentence, e.g. 'the cat sleeps' (p.559) is linguistic context at its most minimal. Even if it were more representative, the procedure of testing the target word out of context, precludes the very effect that context might have. To test the contextual effect with any validity, the least that could have been done would have been to reveal the acquisition trial data. On previous results it could be predicted that the contextual treatments would have facilitated word identification but it would have been interesting to see if sentence + picture was the most facilitative as predicted in this study. This would at least have been some test of the contextual hypothesis¹.

On the conclusion: the data as presented do not support the contextual hypothesis, true. The experiment, however, is irrelevant to the contextual hypothesis. As the authors go on to say in their conclusion, the contextual hypothesis refers to the 'reading process'. But their contrast, that the focal attention hypothesis refers to the 'learning process',

¹ Examination of the means on both D.V's reveals that sentence + picture was the least effective of the four treatments. Since an inverse relationship has existed, in previous studies, between tests and acquisition trials, this is indirect evidence that sentence + picture contained the most contextual constraint allowing the most ready and accurate identification of the target word in acquisition trials.

has bizarre implications. Must it be assumed that the reading process comprises no learning? Must it also be assumed that the 'learning process' that the focal attention hypothesis refers to ('hooking up responses to graphemes') is all that learning to read and the process of reading development is about?

The simple answer to all this is that it is irrelevant to try to test the one hypothesis against the other. The focal attention hypothesis refers to a sub-skill aspect of reading acquisition. The contextual hypothesis refers to reading as it normally occurs and therefore to the complex of skill interaction and strategy that goes into the process of reading development. What is true is that, if the focal attention hypothesis is valid, then the contextual hypothesis must be able to accommodate it.

The second study to specifically invoke the contextual hypothesis is that of Denberg (1976-77). Where this study differs from Singer et al's (1973-74) study is in its testing of both hypotheses - not one against the other. The test of the focal attention hypothesis differs from the 'classical' paradigm on the crucial condition of oral feedback. In Denberg's procedure, "verbal feedback was limited to words of encouragement" (p. 180) on both the acquisition trial and the, words-only, test trial.

48 children from the first-grade participated in the study, all of whom could understand the meaning of the test words but could not 'visually identify' them. Four treatment conditions existed: i) No-picture (a sentence with no

Even under the picture treatment, the greater number of successful identifications facilitated a relatively greater acquisition of graphemic information than under the no-picture treatment where fewer words were correctly identified and therefore relatively less graphemic information was acquired. In other words, the presence of pictures as additional contextual information facilitates accurate word identification but does not exclude attention to graphemic information. This contrasts with the situation where there is no picture information and fewer words are accurately identified. The focal attention hypothesis predicts that, in the latter situation, attention to graphemic cues is not distracted so that, even where fewer words are correctly identified during acquisition, they will nevertheless be better identified during test trials. This was not demonstrated in these results. The only factor that can account for the difference between these and previous results is that no oral feedback was given at all. It appears to be the differential oral feedback, therefore, and not the distracting effects of pictures that has generated previous results.

Finally, through the analysis of errors in context, it was found that the presence of pictures (full or partial) increased the contextual appropriateness of responses as well as the degree of 'integrated graphemic and contextual information' (p.186). This has a direct bearing on the evaluation of strategies of information usage that constitutes one of the aims of the present study. In so far as the present scales and criteria for measuring acceptability are more developed (Denberg used Biemiller's (1970) discrete

criteria), it will be interesting to see whether Denberg's findings are confirmed under more rigorous analysis.

The results of Denberg's study, therefore, not only cast further doubt on the validity of the focal attention hypothesis per se: they also, and independently, provide direct and convincing evidence of the validity of the contextual hypothesis. As evidence under the contextual hypothesis their only limitations are the nature of the task (reading a sentence with an illustration is relatively less representative of a normal reading task than reading a full passage with illustrations); and the dependent variables (lack of comprehension as a variable, and the relatively simple measures of error acceptability). These are not criticisms of Denberg's study which, with a remarkably effective design, aimed to test both the focal attention hypothesis and the contextual hypothesis. They refer to the criteria that are necessary for a full and exclusive test of the contextual hypothesis.

The third study that has focussed specifically on the contextual hypothesis, attempted to meet these criteria (Donald, 1979a; experiment I, Chapter VI). Since this is a direct precursor of the major experiment to be reported in this thesis, it is important to set out its aims, methodology and results in some detail. This will be undertaken in the next chapter. At this point in the research review it is sufficient to indicate, briefly, that the paradigm was intentionally different to the focal attention one (see chapter I for the reasoning behind

this); that the task involved the reading of a continuous passage with and without illustration; that the dependent variables included measures of contextual word identification accuracy, strategies and comprehension, and that, for average, second-year readers, the results were significantly in favour of the facilitative effects of illustrations. Within its limitations, therefore, the study provided further, direct support for the contextual hypothesis.

As stated in the introduction there are a number of studies that have been concerned, basically, with the effects of illustrations on reading comprehension. Although these studies have not been specifically concerned with the contextual hypothesis, and although the effects have been generally limited to measures of comprehension, the task, within treatment variations, has consisted of reading continuous text with and without illustrations. To this extent, studies in this area can be adduced as, at least, indirect evidence of the validity of the contextual hypothesis.

In Samuel's (1970) review of the effects of illustrations on reading, he concludes, inter-alia, that:

"There was almost unanimous agreement that pictures, when used as adjuncts to the printed text, do not facilitate comprehension" (p.405).

In the light of both the interpretation that Samuels puts on some of the studies quoted as evidence, as well as evidence from other studies (not quoted or of more recent origin), it would appear that this general conclusion requires

considerable qualification. Three major factors appear to be relevant in qualifying the effects of illustrations on reading comprehension: i) the first is peculiar to comprehension itself and concerns the nature of comprehension (see chapter IV) and the criteria by which comprehension is evaluated; ii) the second factor concerns the relationship of illustration to text (see chapter I) or how particular sorts of illustration affect the comprehension of particular sorts of text; iii) the third factor concerns the level of reading development under consideration (see chapter I) and the relative effects of illustration on comprehension at such different levels.

i) First, on the nature of comprehension criteria: Miller (1938) constructed a composite test where "children were required to choose, from a group of words, a word spoken by the teacher; to select a phrase from two phrases when one phrase was spoken by the teacher; to cross out an extraneous word from a group of three words; to complete sentences after reading a paragraph; and to put in proper sequence the happenings recorded in a paragraph to be read" (Miller, 1938, p.678). All the test material was re-constructed from material in a basal reader, the reading of which, with and without pictures, had constituted the experimental treatments. The results demonstrated no significant difference between the picture and no picture treatments. Whether such a composite score can be taken as 'comprehension' at all is doubtful. The majority of the criteria (three out of the five) reflect word or phrase recognition more than comprehension and none of the criteria

reflect comprehension of the text itself: only of material based on the text. Vernon (1953) found that although there was no overall significant difference in points remembered with and without illustration, "it did appear that certain major points in the text directly illustrated by pictures were remembered better than those same points unillustrated by pictures" $p < .01$ (p.182). Vernon concluded that illustrations may facilitate the retention of factual information but do not appear to have any effect on the comprehension or retention of relational information. Both of these studies are mentioned as specific evidence in support of Samuels' (1970) general conclusion. Whether they can be taken as such is, however, questionable.

Magne and Parknas (1963), focussing on the interaction of verbally and pictorially presented information with tested information found that "a higher degree of retention was recorded when tests adequate to the mode of presentation were used than when the tests differed from the information type" (p.270). Dwyer (1970) also concluded, after a variety of experiments involving different visual aids combined with textual material, that "the effectiveness of a particular visual aid in facilitating student achievement of a specific objective depends on the type of information needed by the student to achieve that objective" (p.247). Although both of these statements appear self-evident they underline the point that illustrations may facilitate the comprehension of certain types of information more than other types. Clearly, therefore, some reservation is called for in concluding that illustrations have no

effect on reading comprehension: it depends essentially on what is being comprehended.

ii) Second, on the relationship of illustration to text: Vernon, in her 1954 study, structured several illustration-text variations and noted that in one situation where illustrations were ordered in a logical sequence and were closely relevant to the text, the amount recalled was consistently, although not significantly, more than in two other situations where illustrations were more randomly ordered or of less obvious relevance to the text. Ketcham and Heath (1962), in an oral comprehension task, re-inforce the relevance question in their finding that retention of factual information was significantly greater for sound + pictures condition than a sound only condition, and that sound+ irrelevant pictures was significantly worse than for either of these. On the sequence question, Bransford and Johnson (1972), again on an oral comprehension task, demonstrated that providing an illustration before subjects heard a complex passage was significantly more effective on a measure of recall than providing the illustration after the subject had heard the passage or than providing no illustration at all. Perhaps the most convincing findings with specific regard to reading comprehension have been those of Peeck (1974). He systematically varied the informational relationship between illustrations and text. It was found that on three, spaced multiple-choice retention tests, the picture condition produced significantly higher scores than the no-picture

condition for specifically and congruously illustrated text content. This was particularly evident over the longer retention periods. Even for text content that was not specifically illustrated there was a consistent although non-significant trend in favour of the picture condition. For information in the text that had been non-congruously illustrated, however, subjects in the no-picture condition were, at least on short-term retention, significantly superior to subjects who had been presented with the conflicting verbal and pictorial information. This finding emphasises, in particular, how crucial the factor of informational congruity between text and illustration is.

Another interesting aspect to Peeck's findings is the trend in favour of the picture condition for retention of non-illustrated information. This, it will be recalled, was also found by Denberg (1976-77) on accuracy of word identification where words that had not been specifically illustrated were identified with greater accuracy under the partial-picture condition than under the no-picture condition. In two studies, Rusted and Coltheart (1979a; b) confirmed this finding for reading comprehension. They found that simple illustrations presented together with a short descriptive passage about an unusual plant or creature (including living/eating habits some of which were illustrated and some not) significantly facilitated not only the recall of illustrated features but also the recall of non-illustrated features¹.

¹ In the 1979a study, it was only for good readers that non-illustrated features were significantly better recalled under the illustration condition.

This evidence would seem to indicate that for both word identification and literal recall, the relevance an illustration has to the textual content is not necessarily limited to directly illustrated features: an illustration would appear to be providing a semantic framework within which the processing and retention of the general semantic content is enhanced. This interpretation is consistent with another of Bransford and Johnson's (1972) findings. They not only presented appropriate illustrations before and after the text presentation but they also presented, before the text, an illustration that contained all the elements of the appropriate illustration but in an in-appropriate relationship to one another. The recall under this condition was significantly different to that under the appropriate illustration condition and not significantly different to the no-illustration condition. In other words, the relevance of the illustration appeared to be not in its reference to specific elements referred to in the text but to the relationship of elements one to another and to the general semantic content of the text. Guttman, Levin and Pressley (1977) in another oral comprehension task found, over a series of experiments, that partial illustrations (containing all relational features but naturalistically 'blocking out' a crucial element) were as effective in facilitating answers to literal, cued recall questions as were full illustrations at the third grade level but not at a kindergarten level. Moreover, 'dynamic' partial pictures (omitting a crucial element but strongly suggesting the activity expressed in the text) were found

to be significantly more effective on answers to propositional questions than 'static' partial pictures¹. These findings suggest that the relevance of an illustration to its text is not only dependent on the relational qualities in the illustration but, in particular, it is dependent on the 'dynamic' (action related) qualities of the illustration.

Finally, Ruch and Levin (1977), on a similar task with third grade children, were able to show that information in an illustration does not facilitate comprehension only in the sense that it affords opportunity for rehearsal. They compared a partial picture condition with a repetition condition (when children heard the story twice) and a control condition (no-picture; no-repetition). On the ability to answer paraphrase questions, the partial picture condition was significantly superior to the repetition and control conditions. Taken all together, these results suggest that a facilitative relationship between illustration and text certainly may exist and that how this is structured appears to be crucial. It seems likely, provided an illustration is congruous, relevant (particularly in its dynamic qualities) and sequentially related to the content of the text, that it can provide a semantic framework which goes beyond cueing or rehearsal of particular illustrated features of the text: it can provide a basis

¹ Again this is consistent with Denberg's (1976-77) findings on partial pictures (where the relation between actor and object nouns was suggested even when one had been excluded) for word identification accuracy.

for the semantic processing and retention of the general semantic content of the text. Although several of the studies quoted refer to the task of oral comprehension, the consistency in the findings across oral and reading comprehension suggest that the comprehension processes involved are similar enough to justify at least a tentative, common interpretation. Certainly, Samuels' (1970) conclusion, although it may be true in particular instances, can no longer be held as generally valid. As Guttman et al (1977) conclude;

"... there is no longer a need to continue accumulating evidence against those who are skeptical of the positive effects of illustrations on children's prose learning performance (e.g. Samuel's, 1970). Enough has accumulated without having to belabour the issue.. What is sorely needed, however, is an assessment of the nature and extent of picture benefits ..."

(p. 480. Their emphasis)

iii) Third, on the relative effects of illustrations on reading comprehension at different levels of reading development, Guttman et al (1977) found evidence at the kindergarten level that complete pictures significantly facilitated answers to cued recall questions following an oral presentation. This was also true of second and third grade children but there was a progressive improvement in the facilitating effect of the partial-picture to the point, in third grade, where partial and complete picture conditions were not significantly different. In the Ruch

and Levin (1977) study, partial pictures were also found, at the third grade level, to be significantly more facilitative than repetitions and no-picture. This indicated that there might be a shift from illustrations having a more direct semantic reference function in kindergarten children to providing a more integrative semantic framework by third grade. This was somewhat modified by the relative effectiveness of dynamic partial-pictures as compared to static partial-pictures as found with first grade children in a subsequent experiment (Guttman et al, 1972). These results, however, do refer to oral comprehension and need to be warily interpreted in terms of reading comprehension, particularly at these very early stages of reading development. It is clear that more information is required on the effects of illustrations on reading comprehension, specifically, at these stages.

At the fourth grade level of reading development, Peeck (1974), as mentioned, demonstrated the facilitative effects of congruous illustrations on prose retention. Rusted and Coltheart, in both their studies (1979a; b) mentioned earlier, also demonstrated the facilitative effects of illustrations on recall at the 9 year old level. Rasco, Tennyson and Boutwell (1975) presented fourth and fifth grade children with a lengthy passage about mathematical concepts. The children had to read the text under 4 conditions, two of which involved the provision of illustrations depicting important points in the passage. In both illustration conditions (illustrations on their own; illustrations plus instruction to use mental images)

retention of text content was significantly superior to the control condition (no illustration and no image instruction). However, in the third condition (no illustration but instruction to use mental images) retention was also superior to the control. This could indicate an element of redundancy in the value of overt illustrations at this stage of reading development.

This finding, as a developmental trend, is confirmed in another study (Kulhavy and Swenson, 1975). Fifth and sixth grade children read a 20 paragraph text under an induced imagery condition (instruction to image the material) and a control condition. Although there was only a non-significant trend in favour of imagery on an immediate cued recall post-test, the difference emerged as significant on delayed recall a week later.

Rowher and Matz (1975), on orally presented material, provided either appropriate illustrations or a printed version of the text to fourth grade children as additional information. Yes/no answers to assertions about the text were facilitated significantly more by illustrations than by the printed text, but black, low status children gave as many correct answers under the illustrated condition as white, high status children gave under the printed text condition with the difference in performance of the two groups under the text condition being greatly reduced under the illustration condition. This suggests that the illustration effect may be more effective for at least certain types of poor readers than for good readers at this level of reading development.

The results at this level are also qualified in an unpublished dissertation by Haring (1978) who found that illustrations at 4th and 6th grade level facilitate recall of 'important' detail but not of 'unimportant' detail. Bianco (1980), in another unpublished dissertation found that illustrations in fourth grade basal readers had no significant effect on reading comprehension. Taken in conjunction with Vernon's (1953; 1954) findings on 12 and 17 year old children, it would appear that the effects of illustrations on comprehension of basic reading material begin to diminish or become more uncertain from the fourth or fifth grade upward. This is obviously only a tentative assertion and must be qualified by the nature of the task. Dwyer's (1970) findings on college students indicate that under specific task requirements (mainly of certain types of technical reading) illustrations may be beneficial. Bransford and Johnson's (1972) results also show that where the reading required is obscure, illustrations may provide a contextual framework for comprehension at late high school level. However, the tentatively identified differences in the role that an illustration might play in reading comprehension from the initial stages through to adult fluent reading are intuitively suggestive and warrant further, more controlled investigation. Whether conclusions about the effects of illustrations can be drawn, generally, for all developmental levels is more certainly brought into question.

Taken over all, there is some evidence that illustrations may facilitate aspects of reading comprehension for readers

in the first four or five years of reading development. That, with development, illustration may provide a semantic framework that facilitates the processing of the general semantic content of the text, as opposed to specific illustrated features, also appears likely. Interpretation of these findings within the theoretical framework of the contextual hypothesis therefore, lends some support to the validity of the hypothesis.

Following these findings on comprehension and the results of the few studies that have focussed specifically on the contextual hypothesis, it appears that the hypothesis has more than a small measure of support. Nevertheless, it is also clear that a more extensive and specific testing of the hypothesis (meeting all the assumptions and testing its relevance at various levels of reading development and competence) is required if valid interpretations are to be made. This is the purpose of the main experiment to be reported in this thesis.

Before this is undertaken, brief mention needs to be made of research and theoretical evidence that does not relate specifically to reading development but that has relevance within the theoretical framework of cognitive processing that is being developed.

3. Semantic Access : Word and Image Interaction

Since this area is only of indirect relevance, the evidence will be dealt with mainly in terms of general conclusions as opposed to the results of particular research studies.

Research and theory within the area is extensive (Paivio, 1969; 1971; Kosslyn, Holyoak and Huffman, 1976; Pressley, 1977) and it would be both irrelevant and impractical to attempt to be exhaustive.

Within this context, two experimental effects that have each generated considerable research are considered to be of relatively most direct relevance. The first is the verbal/pictorial elaboration effect, and the second is the semantic priming effect.

The first of these effects has been studied mainly through the paired-associate paradigm. Within this paradigm the task involves presenting the subject with a list of pairs, usually pairs of words, pairs of pictures, mixed word-picture pairs, etc. There are two parts to each pair, the stimulus and response terms and the task is to learn the appropriate response to the given stimulus.

Where the elaboration effect is concerned, there are two basic possibilities; the pair can be elaborated in which case the stimulus and response terms are interacted in some form; or the pair can be unelaborated in which case the stimulus and response terms are presented as separate items. The main issues in the research have been whether picture pairs or word pairs, in either elaborated or unelaborated form, are learned more easily. Within this there has been concern to establish whether there is a developmental effect and perhaps, most basically, whether the cognitive processes involved in imagery processing are of the same or different order to the processes involved in verbal processing (Kosslyn and Pomerantz, 1977).

In the present context, the conclusion that has most relevance is the following, as expressed by Pressley (1977) after his review of the evidence:

"In children's learning it seems to matter little whether an elaboration is presented pictorially, verbally or both visually and verbally. One thing is certain: if there is any difference between the potency of pictorial and verbal elaborations, the difference is miniscule compared with the effect of any elaboration over none" (p.590).

One explanation of why this should be so comes from Kowher (1973) who maintains that an elaboration is effective because it provides a common referent for both the stimulus and response terms. In other words, meaning is more than the particular form in which it is presented (words, pictures, etc.): the generation of meaning between the pairs, in the nature of their interaction, is what makes elaborated learning more effective than unelaborated learning. Although the relationship of illustrations to continuous text cannot be equated with this, there are some similarities that might allow some conjecture about common processing. Although illustrations and text are presented separately, there is, particularly when illustrations are presented sequentially within the text, a degree of implicit 'elaboration' in that illustrations and text become cumulatively more related within a common semantic framework. In so far as this may be regarded as a form of 'elaboration', it could be argued that the cumulatively constrained semantic relationship between the components

would be likely to facilitate access to, and therefore learning of, appropriate textual responses. Although this conjecture interprets the technical meaning of 'elaboration' with considerable freedom, it is possible that the process of extracting meaning, or of establishing a common semantic framework, between the components is more constrained - i.e., more 'elaborated' in effect - than, for instance, in a picture-word task where the common semantic referent is far less constrained (see chapter 1).

The second effect has been studied mainly through lexical decision tasks where a subject must decide, as quickly as possible, whether a given letter string is a common English word or not. The semantic priming effect has been verified where subjects have been shown to identify a word more quickly when its immediate predecessor is a semantically related word rather than an unrelated word (e.g., 'leaf' preceded by 'tree' rather than, say, 'road'). Effectively, two models have evolved as attempts to explain the effect. The first is the 'spreading excitation' model which maintains that when a word activates its logogen, the excitation spreads to semantically related logogens that, in their turn, are more readily activated as the result of the original excitation. The second is the 'location shifting' model which is based on the conception of a limited attention capacity mechanism that requires to be shifted for each stimulus to the appropriate logogen. Where logogens are semantically related, the shift is held to be less radical and therefore quicker

(Neely, 1976). A model which accommodates both conceptions as operating in conjunction with each other has also been developed (Posner and Snyder, 1975).

Within this framework, Sperber, McCauley, Ragain and Weil (1979) tested the effects of pictures as stimuli in the semantic priming task. They constructed prime and target pairs that were both pictures, both words or mixed. Of particular interest was their finding that significant semantic priming effects occurred with both mixed as well as unmixed pairs. This supported the assumption that pictures and words access semantic information from a common semantic store.

Again, the relationship of illustration to continuous text cannot be equated with the picture-word relationship in which this semantic priming effect was found. However, it can be directly equated with the acquisition task in the focal attention paradigm. Under these conditions the semantic priming effect may partly account for the easier and more accurate access to the target word that has consistently been found under the picture condition on acquisition trials. As has been pointed out, however (chapter 1), the degree of semantic constraint between a picture and an isolated word is far less than where a picture is related to continuous text. In the latter situation the cumulative semantic constraints in the text itself must constrain the semantic alternatives in the illustration. These, in turn, must constrain access to the semantic alternatives in the text-to-be-read. In

other words it is quite reasonable to expect a strong semantic priming effect in the illustration-continuous text relationship. As with the elaboration effect - and the two are at least conceptually related - a situation where a series of illustrations are cumulatively and sequentially related to the text would appear to be optimal for the semantic priming effect.

Since the conception of attention as a limited capacity system is common to the theory behind the focal attention hypothesis as well as the semantic priming effect, it is perhaps relevant to consider how it might be accommodated within the contextual hypothesis. If illustrations are seen as primes, and not necessarily as specific cues competing for attention to orthographic cues¹, then, on the 'location shifting' model, illustrations may be seen as facilitating the deployment of attention to the target logogen and, in the process, requiring less use of limited attention resources on the accumulation of orthographic - acoustic - semantic equivalences as the alternative means of access to the target word. This would result in a far more efficient use of limited attention and could result in more attention being given to identification of the message and to comprehension of the text. This is precisely what is predicted under the contextual hypothesis.

¹ The evidence, quoted above, that illustrations facilitate both identification (Denberg, 1976-77) and recall (Peeck, 1974; Rusted and Coltheart, 1979a; b) of words not specifically illustrated supports this assumption.

4. Theoretical Derivations

Apart from the central theoretical issues already raised in the research review, there are several other theoretical points of view derived from research into the processing of verbal and imagery information that, although not specifically related to the development of reading, have relevance in possibly explaining an illustration effect in reading. These will, briefly, be developed.

Differential effects for imagery and verbal information have been theoretically accommodated by Paivio (1969; 1971) in the so-called dual-coding hypothesis. Basically Paivio distinguishes between the sorts of cognitive processing that are possible within the two systems. The imagery system, he maintains, is essentially a parallel processing system functioning in a spatial dimension. This limits its organizational, logical properties but enhances its associational properties - i.e., the speed and directness with which imagery associations may be linked and elaborated in parallel. The verbal system, on the other hand, is essentially a sequential processing system and may function in an abstract dimension. This enhances its organizational properties, giving it the capacity to order, categorize and logically relate associations while it also limits the freedom and speed with which it can operate. These distinctions are not absolute, as Paivio emphasises:

"They must be assumed to interact continually in any but the simplest of tasks"

However,

".. one important hypothesis concerning the interaction of the processes is that images are particularly effective in promoting rapid associations while verbal processes give them direction" (Paivio, 1971, p.38).

Reading is certainly not 'the simplest of tasks' and to that extent the relationship between illustrations and reading may be seen as just such a complex interaction between images and language; including language as aroused by images (illustrations) and images as aroused by language (text). The interaction is necessarily complex, but given the essential properties of the imagery system it is likely that illustrations will give rise to a loosely structured, but associationally rich, informational structure. Because of the parallel properties of the imagery system this could be derived from a rapid and brief focus of attention¹. Given the essential properties of the verbal system, on the other hand, it is likely that the imagery information will be progressively interacted with, and given direction within, the cumulative message of the text

¹ This was borne out in observations of subjects reading the illustrated version in the experiments to be reported. Very seldom did subjects 'ponder' the illustrations. More commonly an initial, brief glance at the illustration with, occasionally, a return glance during the course of reading the relevant passage was all that was observably manifested.

as it is read. The nett result would be a loosely structured contextual framework initially provided by the illustration that would become semantically more specific as logical integration with the text progresses. Once again, a series of sequentially related illustrations would cumulatively reach a higher level of initial, semantic specificity.

Whether such an interaction of the information systems would be facilitative would depend on the three factors identified in chapter I: the level of reading development (the degree to which imposed imagery as opposed to self-generated imagery is relevant to the reader's information needs); the level of reading competence (the degree to which strategies for integration of information are characteristic within the reader's processing); and the relationship of illustration to text (the degree to which information is congruous and relevant to the content of the text). The last of these factors has already been discussed. The former two factors, as independent variables in the effects of illustrations on reading, will be developed in chapter VII.

Mention should also be made of an alternative theoretical interpretation that relates particularly to the observed differential in the recall potential for 'imaged' information as opposed to verbal rehearsal. According to Kosslyn (Kosslyn et al, 1976; Kosslyn and Pomerantz, 1977) the differential can be accounted for in a semantic elaboration hypothesis. According to this, the active generation of images requires a greater degree of semantic processing, with resultant semantic elaboration, than mere verbal rehearsal. This, in turn, creates a more durable - more deeply processed - memory structure

that is more available for retrieval (Craik and Lockhart, 1972). This interpretation may account for findings such as those of Peeck (1974) and Rusted and Coltheart (1979b) that the difference between illustrated and unillustrated reading is greatest on delayed as opposed to immediate recall. Although the image is not strictly being 'generated' under these circumstances, it has already been argued that a degree of elaboration is inevitable in the interaction of illustration and text. To this extent the resultant, common, semantic information has been more than rehearsed: it has been actively integrated into a meaning - and possibly 'imaged' - framework.

Finally, a theoretical concept as developed by Neisser (1967) is particularly relevant in considering the process of reading and the possible role of illustrations within this process. Based on an extensive body of research, his theory of analysis-by-synthesis develops the functional concept of 'focal attention'. Analysis-by-synthesis is essentially a constructivist theory of perception in which the perceiver selects, from the information in the signal itself and in the contextual constraints in which it is embedded, certain minimal information on which basis he constructs (synthesises) a message that, in turn, is tested for authenticity against the original signal and its constraints. Neisser maintains that it is the mechanism of focal attention itself that directs what will be selected, what constructed and what tested. In its turn, focal attention is determined by the individual's 'set'; his expectancy, based on the total contextual build-up

and experience of the perceiver as active at that moment. Since the process of reading can be viewed as a complex case of analysis-by-synthesis (chapter II; Smith, 1971; Goodman, 1967), the role of focal attention - and the reader's set or expectancy that directs this - are of central importance in determining its efficacy. Illustrations, as legitimate contextual information, can therefore be expected to exert an influence on the individual reader's set. Whether this set will be facilitatory in directing the reader's focal attention is, again, likely to depend on the reader's level of development and the relevance of illustrations as information within his need for contextual information; the reader's level of competence and his ability to modify his set in terms of cumulative semantic information; and the relevance of the particular illustration as contextual information in that particular text.

Of further interest in this context, Gibson and Levin (1975) in discussing comprehension, refer to the concept of 'foregrounding';

"Certain concepts are at the moment in the mind of the reader, in 'sharp focus', which puts the new information in a semantic framework that permits comprehension. ... Foregrounding is not the same thing as fitting new information into a whole system of knowledge, but it is a necessary and more immediate preparation for comprehension ..." (p.394),

This concept is, in all essential respects, parallel to

the concept of 'set'; and both constitute a semantic expectancy that may be influenced by illustrations and may exert considerable influence on the ensuing process of message identification and comprehension.

CHAPTER VIEXPERIMENT I : The Effects of Illustrations
on the Oral Reading Accuracy, Strategies
and Comprehension of Second-Year Average
Readers : A Test of the Contextual Hypothesis¹

As was pointed out in the previous chapter, although there is some support for the contextual hypothesis, there was (previous to Donald, 1979a, which is reported here) no study that had attempted to test the hypothesis in strict adherence to its assumptions. More specifically there was a need to test the hypothesis on the reading of relevantly illustrated, continuous text passages (i.e. more than short sentences) and for the efficacy of, at least, word identification accuracy, strategies of information use and comprehension to be assessed in this context. The first experiment was undertaken, therefore, in a specific attempt to meet these requirements. Concern was with average readers at an early stage of reading development where illustration is traditionally regarded as important (Gurney, 1976). The study focussed therefore, on average, second-year (7 year old) readers who were required to read continuous text passages with and without illustration. Within this the dependent variables

¹ This research constituted the basis of a thesis submitted in partial fulfillment of the requirements for the degree of Master of Education in the Faculty of Education, University of Exeter (Donald, 1978). It is reported in this context in so far as it has indispensable relevance to the follow-up experimental study that constitutes the major focus of this thesis.

- were:
- i) Word identification accuracy (the proportion of words correctly read in context)
 - ii) Semantic acceptability of errors (the degree to which errors were acceptable within the semantic framework of the text)
 - iii) Syntactic acceptability of errors (the degree to which errors were acceptable within the syntactic structure of the text)
 - iv) Orthographic acceptability of errors (the degree to which errors were acceptable within the orthographic structure of their stimulus equivalents in the text).
 - v) Self-correction (the proportion of errors spontaneously self-corrected)
 - vi) Literal idea recall (the proportion of ideas, as stated in the text, that were freely recalled by the reader)
 - vii) Inferential comprehension (the degree to which the reader could make appropriate, cued inferences beyond the information explicitly stated in the text).

1. Method

i) Design: Owing to the variety of dependent variables under consideration as well as the nature of the independent variable, a repeated measures design was selected as the most appropriate and powerful for testing the illustration effect. This involved two randomly

selected groups each reading two matched but different stories, A and B; the first group reading story A with illustration and story B without; the second group reading story B with illustration and story A without. The results for each dependent variable were tested on a two-factor analysis of variance with repeated measures. The within subjects main effect, over illustration, no illustration was of principal interest. The between subjects main effect over the two randomly selected groups was an artefact of the design. Nevertheless, the interaction, in so far as it might reflect variance specific to the relationship of a particular story and its illustration was of equal relevance to interpretation.

ii) Sample: A sample of 20 children was randomly selected from an urban population of average, second-year readers, ranging in age from 6,10 to 7,4 years across three schools in Exeter. Average was defined, in the first place, on the rough criterion of falling within the middle 50% of reading achievement for that year group. Since variations between schools could be expected on this criterion, a further screen in which only those scoring, on accuracy, within one standard deviation of the mean for the whole pre-selected group on a parallel form of the stories to be read were included. The sample of 20 was then randomly divided into the two design groups.

iii) Materials: The materials consisted of two stories and two illustrations mounted on white card,

24 x 15 cm (Appendix 2). The stories were structurally matched in so far as each was 66 words long; the content of each consisted of an every-day narrative account; the readability index (Spache, 1953) for each was 1,9; simple, active syntactic structure alone was used in each; and one sentence per line, plus uniform print-size, etc., was used in each. The illustrations were clear line drawings in pen and ink and each illustration was required, in terms of relevance, to depict the central event in the story, as well as to represent the central agents and physical setting of the story. Relevance was also constrained in the ratio of length of text to illustrations which, in each case, was 66:1.

iv) Procedure

The order of presentation, story A first or story B, was randomly determined for each child. Tape recordings of oral reading and comprehension responses were made and subsequently transcribed for scoring and analysis. Scoring of the seven dependent variables was as follows¹:

Word Identification Accuracy. Each misinterpretation of the text, at the whole word level, conventionally excluding

¹ Full details of the scoring system for this experiment are not included as, although there are minor differences, the system was essentially the same as that used in experiment II. Since the latter constitutes the major study in the thesis, and since the scoring system for that experiment is reported in full detail in chapter IX the principle criteria only are reported here.

self-corrected responses and repetitions, was counted as one error. The remaining correct responses constituted the percentage accuracy.

Syntactic Acceptability. Each error (including responses subsequently self-corrected as well as repetitions - all of which were taken as legitimate indicators of strategy) was scored on a five-point scale for its degree of acceptability within the syntactic structure of the text.

Criteria and scores for the relative syntactic acceptability of errors were:

- 4: Acceptable within both the full sentence and the passage as a whole.

hurt
e.g. She was lost and lonely

- 3: Acceptable within the full sentence, but not the passage as a whole

is
e.g. She was lost and lonely.

- 2: Acceptable in the sentence, to the point of the error only.

came
e.g. She was lost and lonely.

- 1: Acceptable only in that it has the same grammatical function as the stimulus word.

were
e.g. She was lost and lonely.

- 0: None of the above.

The total syntactic acceptability score, based on all the subject's errors in any one condition, was expressed as a percentage of the total possible score for those errors. The final percentage reflected, therefore, the 'average' syntactic acceptability of a particular set of errors. The same applied to semantic and graphic acceptability. Over all conditions and all subjects the total number of errors analysed on each acceptability scale was 479: a mean of 12 errors per subject per analysis with a range limited by the initial screening.

Semantic Acceptability. Each error was scored on a five-point scale for its degree of acceptability within the semantic framework of the text. Criteria and scores for the relative semantic acceptability of errors were:

- 4: Fully synonymous with textual meaning.

e.g. Father ^{doesn't} does not tell her

- 3: Marginal but non-substantive change in meaning.

e.g. Now she stays in our ^{house} home

- 2: Meaning acceptable to the cumulative sequence of ideas

e.g. There is pretty ^{ribbon} paper round the box. (Story of a party and getting a present)

- 1: Meaning having a loose association with general textual content.

e.g. There is pretty ^{cakes} paper round the box

0: None of the above.

Orthographic Acceptability. Each error was scored on a five-point scale for its degree of acceptability to the orthographic structure of the stimulus word. Criteria and scores for the relative orthographic acceptability of errors were:

- 4: Total correspondence
e.g. heard/heard (as in beard)
- 3: Differing by only one letter (including reversals)
e.g. stays/says or was/saw
- 2: Having first and second or first and last letter in common.
e.g. until/under or stands/slips
- 1: Having only first letter in common.
e.g. table/top
- 0: None of the above.

Note: In all three of the above scales specific criteria for scoring 'exceptional' errors including omissions, insertions, nonsense words, errors at the beginning of a sentence, etc., were also specified in detail in order to maximise reliability of scoring.

Self-Correction. The number of spontaneous self-corrections of whole word errors was expressed as a percentage of the total number of errors.

Literal Idea Recall. The structure of the text allowed each sentence to represent one idea. The child's spontaneous recall of the story was analysed into corresponding units each of which was scored on a four-point scale, the criteria of which were:

- 3: Identical or synonymous with text idea.
e.g. I gave her something to drink. (text)
'He brought the cat a drink of something'. (response)
- 2: All essential elements of text idea present.
e.g. 'He put a bowl of milk down for her'.
- 1: Some element(s) of text idea present.
e.g. 'He gave her some food'.
- 0: Irrelevant or incorrect.
e.g. 'He gave the cat a pat'

Note: Examples given are deliberately 'borderline'; most responses were, in fact, more easily classifiable.

The total score was expressed as a percentage of the total possible score.

Inferential Comprehension. Three questions of the order 'Why - - -' or 'How do you know ---' followed each story. Responses to these questions were scored as 3, 1 or 0 depending on the relevance of the response. Total score was again expressed as a percentage of total possible score.

2. Results¹

i) Word Identification Accuracy: Over combined groups the mean word identification accuracy for the illustration condition was 86,67 per cent and, for the no-illustration condition, 82,73 per cent. This difference was statistically significant ($F = 6,14$, $P < 0,05$). The difference between groups was non-significant ($F = 0,01$). An interaction ($F = 5,30$, $P < 0,05$), however, was present. Examination of the component means suggests two alternative interpretations for the interaction (Table 1).

TABLE ICELL MEANS FOR WORD IDENTIFICATIONACCURACY

	No-illustration	Illustration
Group 1	Story B 84,39	Story A 84,70
Group 2	Story A 81,06	Story B 88,64

One possibility is that Group 2 was affected by illustration while Group 1 was not. This seems unlikely in terms of

¹ The relevant means, standard deviations and analysis of variance summary tables for each variable are presented in full in appendix 6.

both the random selection of groups and the other results. The other possibility is that accuracy was as much affected by the particular story as by the presence of illustration. Story B, in other words, was read with greater accuracy than story A; but the reading of both was facilitated through the presence of illustration. This is the most likely interpretation. It also preserves the significance of the main effect.

ii) Syntactic Acceptability. Over combined groups the mean syntactic acceptability for the illustration condition was 51,49 per cent and, for the no-illustration condition, 37,46 per cent. This difference was statistically significant ($F = 7,33$, $P < ,05$). Neither the difference between groups nor the interaction were significant, thus leaving the interpretation of the main effect clear.

iii) Semantic Acceptability. Over combined groups the mean semantic acceptability for the illustration condition was 44,24 per cent and, for the no-illustration condition, 27,33 per cent. This difference was statistically significant ($F = 16,02$, $P < ,001$). Neither the difference between groups nor the interaction were significant.

iv) Orthographic Acceptability. Over combined groups the mean orthographic acceptability for the illustration condition was 35,25 per cent and, for the no-illustration condition, 39,95 per cent. This difference was statistically significant at the $P < ,05$ level ($F = 4,46$). The other effects were non-significant.

v) Self-Corrections. Over combined groups the mean self correction rate for the illustration condition was 22,22 per cent and, for the no-illustration condition, 10,49 per cent. This difference was statistically significant ($F = 13,07$, $P < ,01$). The other effects were non-significant.

vi) Literal Idea Recall. Over combined groups the mean idea recall for the illustration condition was 26,39 per cent and, for the no-illustration condition, 18,75 per cent. This difference was statistically significant ($F = 9,85$, $P < ,01$). The other effects were non-significant.

vii) Inferential Comprehension. Over combined groups the mean inference score for the illustration condition was 68,33 per cent and, for the no-illustration condition, 53,33 per cent. This difference was not statistically significant ($F = 3,28$), neither were the other two effects.

3. Discussion

In the light of previous results regarding the effects of illustration on reading these results presented an interesting array. First, in terms of the contextual hypothesis, the results confirm what was found by Denberg (1976-77). Word identification accuracy, in context, is significantly facilitated in the illustration condition. This also endorses the finding in the focal attention studies that word identification accuracy in acquisition trials, when the picture is present, is better than in the no-picture acquisition trials (Samuels, 1967; Harzem et al, 1976; Willows, 1978a; b). On the question of accuracy of message

identification, both the semantic and the syntactic acceptabilities of errors - particularly the former - were significantly greater in the illustration condition.

Taken together with accuracy, this means that, under the illustration condition, subjects made significantly fewer errors and those errors that they did make were significantly more appropriate to the semantic and syntactic content of the text than under the no-illustration condition. Message identification, therefore, was clearly facilitated by the presence of an illustration.

On the question of the efficiency of message identification, the relevant section of the contextual hypothesis states that '... the efficiency and accuracy of message identification and the effectiveness of comprehension will be facilitated through reducing the reader's reliance on orthographic information while increasing his use of the available contextual information.' In so far as error acceptability can be taken as evidence of the reader's relative use of information in any one dimension (chapter IV), the significantly higher acceptability of both semantic and syntactic errors in the illustration condition endorses the predicted increase in use of contextual information.

Equally the significantly lower acceptability of orthographic errors in the illustration condition endorses the prediction of reduced reliance on orthographic information. That these two effects in combination lead to greater efficiency of message identification is suggested in the significantly higher accuracy and literal comprehension achieved in the illustration condition. Of further relevance is the much

smaller reduction in orthographic acceptability when compared to the corresponding increases in semantic and syntactic acceptabilities (mean differences, illustration - no-illustration: - 4,70 (orthographic) vs. 16,91 (semantic); 14,03 (syntactic). This accords with the idea that orthographic information is not 'sacrificed' on an all or nothing basis: reliance on orthographic information is reduced only to the extent that it is redundant within the additional contextual information. This is consistent with Denberg's (1976-77) finding that there is a partial rather than a total 'trade-off' of orthographic information under conditions of enriched contextual (illustration) information.

The argument (Samuels, 1970) that attention to orthographic cues is distracted, on an 'either-or' basis, by the alternative pictorial stimulus may operate in the paired-associate task (although this itself is subject to question: chapter V). Under the more normal reading task conditions reported here, however, this interpretation seems most unlikely. On the contrary, the evidence suggests that attention is given to orthographic cues but that deployment of attention is selective such that only those orthographic cues that are necessary, within the redundancy of additional contextual information, will be selected.

Theoretically, at this point, it is important to return to the concepts of mediated access, immediate access and redundancy. It could be argued, as effectively it is in the focal attention hypothesis, that reduced use of orthographic cues will lead to inadequate development of the equivalences that make up mediated access. Three

points counter this argument. First, strategies that make use of the available contextual information to reduce uncertainty are also, of necessity, facilitating the establishment of equivalences in so far as words are correctly identified. The more words that are correctly identified, the more equivalences must become established. The evidence is quite clear that word identification is facilitated under illustrated conditions and development of the necessary equivalences must follow. Second, self-correction was also found to be significantly improved under illustration conditions. Not only has this strategy been shown to be related to optimal reading development in the early stages (chapter IV), but the very act of detecting an error and re-selecting information in order to correct it must facilitate the establishment of appropriate equivalences. Third, as argued in chapter III, the development of mediated access is seen as necessary within the accelerating model of reading development but over-reliance on its use is inefficient at any stage. The orthographic acceptability of errors reflects the use of mediated access, not necessarily its development. As already pointed out, strategies which facilitate accurate reading and self-correction are sufficiently effective vehicles for the development of the necessary equivalences. Finally, with regard to immediate access and the use of redundancy, it was argued in chapter III that these two means of access are essential, at any point in reading development, if efficient message identification and comprehension are to be achieved. Evidence from the results is that strategies reflecting the use of

contextual information (semantic and syntactic acceptabilities) are facilitated under illustration conditions: it necessarily follows that use of redundancy is also facilitated and that under these conditions a strategy of efficient use of redundancy will develop. Equally with immediate access: where more words are being correctly identified through only partial orthographic analysis it follows that a strategy of immediate identification is being, at least partially, developed and must lead ultimately to more efficient reading development.

The results on comprehension confirm that recall of literal information is facilitated by illustration. This is consistent with the majority of studies reported in chapter V where comprehension - particularly on measures of retention - was concerned. As pointed out in that chapter, however, very little information is available on the effects of illustrations on comprehension - particularly reading comprehension - at the very early stages of reading development. These results, therefore, fill a gap and confirm that the effects on free recall in reading comprehension are similar to those on cued recall in oral comprehension (Guttman et al, 1977) at this level. The result on inferential comprehension, as distinct from literal recall, is apparently a new finding. The distinction has not been explicitly drawn in previous studies in this area although it is possible that previous measures of comprehension may have included degrees of inferential comprehension. The finding that there was no significant difference between

the illustration and no-illustration condition on this measure is, however, consistent with Vernon's (1953) conclusion that illustrations may facilitate retention of literal or factual information but not relational information. It is also theoretically consistent with the idea that inferential comprehension involves, essentially, verbal sequential thinking that is unlikely to be directly affected by the spatial imagery of an illustration (Paivio, 1969; 1971). This is probably a relative effect and the trend in favour of the illustration condition could well reflect this. Whatever the actual nature of the effect, it is certainly not adverse.

4. Limitations and Research Implications

A number of limitations in this experiment indicated that replication and extension would be appropriate.

i) The materials. In order to fully meet the assumptions underlying the contextual hypothesis (chapter I), some modifications in the experimental material as they affect the reader's task were indicated. The main consideration is that a series of illustrations should be sequentially related to the text. As indicated at various points in the theoretical argument, a series of related illustrations are not only more common in children's reading material but the semantic specificity or contextual constraint operative in a sequential series must necessarily be greater than in a single illustration. This was also suggested in Vernon's (1954) findings. Second, in order to develop a narrative

theme and to test whether illustrations have a cumulative effect or only an initial effect when the content of the text is still relatively unpredictable, stories need, ideally, to be longer than the 66-word passages used in this experiment. Clearly this also relates to the need for a series of illustrations.

ii) Error Analysis. The number of errors per analysis must affect the reliability of the measures of acceptability. Although the mean of 12 errors per analysis was regarded as within a permissible minimum on the basis of precedent (chapter IV) it was felt that replication on the basis of larger error samples would provide more reliable evidence. Some dissatisfaction was also felt in relation to the error acceptability scales themselves. On the semantic acceptability scale, in particular, which is so central to testing the hypothesis, it was felt that more explicit criteria needed to be developed in order to increase the reliability of scoring (see chapter IX below).

iii) Comprehension. Free recall as a measure of comprehension is neither as reliable nor as valid as cued recall (Gibson and Levin, 1975). Although the results on free idea recall in this experiment were consistent with findings on cued recall, it was felt that replication of the experiment using more carefully selected measures of literal comprehension and the cued recall technique would be more appropriate.

iv) Design. Although the repeated measures design was selected for a good reason, it produced some minor problems in interpretation of the interaction effect. In this particular experiment the problem did not emerge as serious but the design was not felt to be appropriate for any further study of the illustration effect, where more than one independent variable was to be investigated. It was also felt that more than 10 subjects per cell would be desirable in terms of the fairly high variance found on some of the dependent variables.

v) Independent Variables Most important, and not as a criticism of the experiment since these were not within its aims, it was felt that the experiment needed extension on at least two crucial independent variables. As repeatedly emphasised, the question of whether the effects of illustrations on reading are facilitatory or not must depend on three factors: the relationship of illustration to text; the developmental level of the reader; and the relative competence of the reader. The first factor has received considerable research attention, albeit mostly in the area of comprehension effects (chapter V). The other two factors, however, have received very little specific attention and no studies have attempted to clarify their possible interaction.

In conclusion, within its limitations, the results of this experiment endorse the contextual hypothesis as a valid interpretation of the effects of illustrations on the reading of average, beginning readers. In this sense the findings

are consistent with those of Denberg (1976-77) who also focussed on a similar although somewhat younger group. A need exists, however, not only to replicate the study under more rigorous conditions but to extend the test of the contextual hypothesis to different developmental and progress levels. The next chapter will consider the theoretical grounds and basis for selection of these independent variables.

CHAPTER VIITHEORETICAL DERIVATION OF THE INDEPENDENT
VARIABLES; DEVELOPMENTAL LEVEL AND PROGRESS LEVEL

In their recent review of research into the effects of illustrations on reading, Goldstein and Underwood (1981) conclude:

"In general, the evidence reviewed above implies that the less competent a reader is with print, or the younger he is, the greater will be the relative influence of pictorial information. This influence may or may not be of benefit, depending upon whether the picture is representative of the text, and to what extent it corresponds to the intended meaning" (p.14).

As reviewed in this study as well, there can be no argument with the second part of their conclusion that the relationship between illustration and text is crucial in the effect. The first part of the conclusion is less empirical. With only one or two exceptions in the relatively peripheral areas of research into semantic access and oral comprehension, all the studies reviewed by Goldstein and Underwood are also reviewed here. In addition, several central and recent studies (Singer et al, 1973-74; Denberg, 1976-77; Arlin et al, 1978-79; Willows, 1978b; Bransford and Johnson, 1972; Rasco et al, 1975), apart from a number of more minor studies, were not reviewed. Within this comparative framework, it is hard to see how they reach the conclusion

that they do. Certainly there are isolated indications in some studies that the conclusion may be justified. However, differentiation of the two basic paradigmatic contexts in which results have been obtained, which Goldstein and Underwood do not do, is also a prerequisite to drawing conclusions. The point is that the conclusion may be justified but the evidence as it stands at the moment is simply not sufficient to draw such a conclusion.

If Goldstein and Underwood's conclusion is asked as a question requiring an empirical answer then it has far more justification. The question is: are the effects of illustrations on reading influenced by the relative developmental and progress (competence) levels of readers? And more specifically: is there an interaction between developmental level and progress level in the extent to which readers are influenced by illustration? The evidence for maintaining that these questions are worth asking, and in what form they need to be asked, will now be reviewed.

1. Developmental Level

Within the focal attention paradigm almost all studies, with only three exceptions, have focussed on the very early stages of reading development - pre-school to first-grade - and none of the results within this narrow developmental range have shown developmentally related variation.

With considerable consistency, what they have shown is a clear contextual effect on acquisition trials and an equally clear 'distracting' effect on test trials. As

pointed out, however, the validity of the interpretation of the distracting effect has been seriously questioned. What remains without question, therefore, is the clear contextual effect of illustrations at this early stage of reading development. The distracting effect must remain at the level of a possibility only. Singer et al (1973-74) worked with second graders but their results were totally consistent with those for younger children. Willows (1978a; b) worked with second and third graders and one of his specific conclusions was that second graders were more susceptible to the distracting effects of pictures than third graders. It will be recalled, however, that the principal evidence for this was a non-significant but more marked response latency effect with related and unrelated pictures compared to no-picture in the second grade than in the third grade. As argued, this can hardly be taken as evidence of the influence of illustrations on reading as response latency in that situation probably only reflected legitimate puzzlement on the part of the children at the bizarre picture-word pairs they were being presented with. At both levels accuracy was increased by the related picture and decreased by the unrelated picture relative to no-picture. This tended to be more marked for the second graders, however, and may suggest greater susceptibility at lower developmental levels to the direct influence of pictures. The finding, in 1978b, on third graders only, that the effect was non-significant for normal and good readers and only significant for poor readers on more difficult words could be taken as further support for this.

Specifically under the contextual hypothesis, Denberg (1976-77) worked with first graders and Donald (1979a) worked with second year readers. On their criteria of accuracy and strategy, in so far as they can be directly compared, there was no detectable developmental difference. However, both studies indicate a clear contextual illustration effect at this level and Donald's study, in particular, demonstrated that the effect is specific not only to accuracy but also to comprehension as well as to the balance of strategies.

On this evidence there would appear to be a clear contextual effect up to second grade that becomes less certain by third grade. A distracting effect for picture-word pairs is also possible at this level although subject to some interpretive doubt.

In the area of comprehension, the developmental trends have already been reviewed (chapter V). These are tenuous but suggestive. Perhaps the most interesting and well supported indication is that there appears to be a developmental progression from first grade through to third grade and beyond in the ability of children to benefit from full pictures to partial pictures to induced imagery (Guttman et al, 1977; Rasco et al, 1975; Kulhavy and Swenson, 1975). This would seem to imply that, on comprehension at least, external, explicit illustration of text may be more influential at lower developmental levels, simply because it becomes potentially more redundant at higher levels. The greater inconsistency of results on

the effects of illustration on comprehension at the higher developmental levels would also seem to indicate that the contextual effect becomes progressively less certain with development.

Evidence from outside the specific relationship of illustration to reading may also be adduced in so far as it may theoretically predict a developmental trend in the relationship.

The most compelling argument here is that readers develop in their ability to make use of the linguistic constraints in the text itself. They learn to optimise the cumulative semantic and syntactic constraints in the text to the extent that these constitute sufficient contextual information and redundancy for efficient prediction. Clearly, the more contextual information the reader can extract from the text itself the less relevant is an illustration as a source of contextual information. Put alternatively, the less skill the reader has developed in extracting contextual information from the text the more he is likely to need the information in illustrations to supplement the limited contextual information that he can extract.

Considerable evidence does exist to show that readers develop in their use of text based contextual information. Evidence from eye-voice-span (EVS) studies is some of the most convincing. EVS refers to the number of words a reader, who is reading aloud, can predict after the text he is reading is removed. Gibson and Levin (1975) report:

"From the fourth grade on the EVS is longer for sentences than for word lists, indicating that children start taking advantage of grammatical structure sometime between the second and fourth grades. There is a tendency starting with the fourth grade for reported EVS to end at phrase boundaries, especially for more skilled readers. Second graders do not yet chunk their reading in phrase units" (p.370-1; my emphasis)

From the point of view of selecting developmental levels to test a predicted change in the contextual value of illustrations this evidence is useful. It suggests quite clearly that a significant change in the ability to use text-based contextual information (particularly syntactic) occurs between second and fourth grade (approximate reading ages: 7 and 9). The emphasis in the quote refers to a probable interaction of progress level with developmental level that will be picked up in the next section.

Development in the use of semantic information appears to parallel the use of syntactic information. Felzen and Anisfeld (1970) asked third and sixth grade children to judge, in a list of words, whether a word had occurred before. The evidence was that semantic similarity caused more errors for sixth graders than for third graders and vice versa for phonetic similarity. Burke (1976), using an error analysis technique, demonstrated a significant development in the use of semantic constraints from a seven year level to a nine year level with the increase being particularly marked between the eight and nine year levels.

Particularly relevant evidence for this effect comes from a study by Mackworth (1972b). In an experiment with good and poor readers over grades two, four and six, he required children to read an incomplete sentence (e.g., 'He could not carry the ..') and then search two separated lists of words to find the most appropriate word to fit the sentence. The two lists differed in that one consistently contained nouns (e.g., books, year, sky) and the other consistently contained verbs (e.g., ran, threw, makes). The missing words in the sentences were either verbs or nouns. The dependent variable was a measure of how long the children spent searching the wrong list. In other words, it was a test of how clearly they had picked up the syntactic (as well as, to some extent, the semantic) constraints of the sentence. The results showed that there was a significant decline in wrong searching between second and sixth grade, but particularly between second and fourth grade. Of significance to the next section on progress level differences, and related to my emphasis in the Gibson and Levin quote above, it was also found that although both groups of readers improved, this was significantly more marked for the good readers than for the poor readers who showed very little improvement beyond fourth grade. This would seem to indicate that for poor readers the contextual constraints in an illustration may continue to have relevance from the fourth grade onwards while, for the good readers, this would be minimal by the fourth grade and reduced even further as they progress.

Further unexpected evidence that the linguistically based contextual turning point is around the developmental level of 9 years comes from findings on deaf children. Furth (1966) quotes evidence that before the age of ten deaf children advance in reading development to almost a third grade level. Between the age of 10 and 16 there is less than a grade's development (i.e., median grade equivalent is still 3,4 by 16 years). In other words where normal children advance in reading development beyond third grade level on the basis of linguistic competence and the use of text based linguistic context, deaf children are unable to do so because of their lack of linguistic competence. Apart from anything, this suggests the possible relevance of illustrations in reading material for deaf children.

On this evidence, and the more tenuous evidence surrounding illustrations and reading itself, it was decided to investigate developmental level as an independent variable and to select the developmental levels of seven and nine years for specific investigation. In this way it was hoped to bridge what appears to be a crucial period in the sorts of information that a reader uses.

2. Progress Level

Evidence from the research on illustrations and reading on differences between good and poor readers is, again, incomplete but suggestive.

Within the focal attention paradigm, Samuels (1967) in his

experiment II divided his first graders into above and below the median groups on pre-test scores and designated the respective groups as better and poorer readers. Unfortunately the results for the acquisition trial (the reading of 'Fun at Blue Lake') are not given as these would have made an interesting comparison with Donald (1979a) on the criterion of accuracy. On the test trial (words only), the poorer readers who had read the original with illustration were significantly worse at isolated word identification than the poorer readers who had read without illustration.

The difference between better readers, however, was negligible and non-significant. Samuels interpreted this as evidence of the poorer readers' greater susceptibility to distraction by the illustration. It could also be that the better readers were more able to integrate the contextual (illustration) information and therefore less susceptible to the distracting effects. This dual interpretation is consistent with Denberg's (1976-77) finding that first-grade children with high orthographic skill (who are usually 'better' readers at this level) made use of a 'partial trade-off' of information and were better able to benefit from the contextual information in the illustrations. Children with low orthographic skill, however, ('poorer' readers), were more inclined to use a 'total trade-off' strategy and although they benefitted from the illustrations this was not as marked as it was for the 'better' readers. Willows' (1978a) findings that poorer readers in both second and

third grade, made more word identification errors when words were paired with unrelated pictures than with related pictures, while better readers were not differentially affected, is also consistent with this pattern as is his finding (1978b) that poor readers identified more difficult words with identifying pictures but showed themselves more susceptible to the distracting effects of unrelated pictures (on latency).

At this developmental level, therefore, (grades 1, 2 and 3) there is the suggestion, coming most clearly from Denberg's results, that good readers are able to integrate contextual information from illustrations such that their word identification accuracy is more clearly facilitated than it is for poor readers. Poor readers, particularly from Samuels and Willow's findings, appear less able to disregard the distractions and irrelevancies that may be presented in illustrations while this does not appear to affect good readers. The net effect is that good readers experience a substantial gain (context) and no loss (distraction) with illustrations; while the poor readers experience only a partial gain (context) as well as a likely loss (distraction).

At older developmental levels, the pattern appears to change somewhat although the evidence is now from comprehension criteria and a different paradigm. Rusted and Coltheart (1979a) found with 9-year olds (4th grade) divided into good readers and poor readers that both groups benefited significantly on recall after having read with illustrations. The mean recall scores were as follows:

	<u>Good</u>	<u>Poor</u>
<u>Picture</u>	22	21
<u>No-Picture</u>	21	18

Although the interaction was non-significant, a trend is clearly apparent and it would appear that good readers might benefit relatively less than poor readers from the presence of an illustration. In addition there was a significant interaction for type of feature recalled in the picture condition such that good readers recalled more non-illustrated features than illustrated features and the reverse held true for the poor readers. It would appear from this that the benefit gained by good readers from illustrations was marginal yet of an integrated nature while that gained by poor readers was more substantial but less integrated. In Rusted and Coltheart (1979b), with similar groups, it was found that there was no significant relationship between amount recalled and reading ability on any of the illustration conditions (all relevant illustrations). However, on the no-illustration condition, significant correlations ($r = .73$; $p < .001$ immediate recall; $r = .45$; $p < .05$ delayed recall) were obtained between recall and reading ability. This suggests that poor readers recalled as well as good readers in the illustration conditions but were much worse than the good readers in the no-illustration condition. This lends support to the previous tendency for poor readers to benefit more substantially from the presence of illustrations at this older developmental level.

Indirectly the Rowher and Matz (1975) study described earlier also lends support to this observation. They worked with fourth graders as well. The two groups were middle class white and lower class black. Given the established relationship between social class and reading ability (Kellmer-Pringle, Butler and Davie, 1966), these two groups could be regarded as better and poorer readers. What was demonstrated was that although both groups benefitted more from a pictures than from a text accompaniment to an oral story, the lower class black group benefitted substantially more than the middle-class white group.

Another indirect source of evidence comes from a study by Levin (1963). From fourth graders, he selected three groups: a 'deficit' poor reader group; a 'difference' poor reader group (defined as poor on comprehension and not basic skills); and a good reader group. Three conditions also existed: a text-only condition; a pictures-only condition (a sequence of pictures with no accompanying text); and a text-plus-image condition (text plus instruction to image the text). On comprehension questions answered, text-plus-image was superior to text-only and pictures-only was inferior. What was interesting, however, was that the 'difference' group benefitted more under the text-plus-image condition than the 'deficit' group and in fact, under this condition, they achieved at the same level as the good readers achieved under the text-only condition. Again, the suggestion is that poor readers ('difference' poor readers at any rate) benefit substantially from imagery.

And as Guttman et al (1977) have shown, this is as effective as overt illustration by third grade level.

The indications from all this, therefore, appear to be that good and poor readers benefit differentially from illustrations but that there is a reverse in the differential with developmental level. This leads on to the next section.

3. Interaction

The evidence presented thus far would seem to indicate an interaction between developmental level and progress level where the effects of illustration are concerned. In order to clarify and summarise the indications that have been discussed they will be weighted and expressed in tabular form.

TABLE 2:

WEIGHTINGS FOR THE THEORETICALLY EXPECTED EFFECTS OF ILLUSTRATIONS BY DEVELOPMENTAL AND PROGRESS LEVELS.

		<u>DEVELOPMENTAL LEVEL</u>	
		7	9
<u>PROGRESS LEVEL</u>	HIGH (good)	+2	0
	LOW (poor)	+2	+1

+2

0

+1

+2

0

+2

-2 strongly negative

-1 marginally negative

0 nil effect

+1 marginally positive

+2 strongly positive

Thus, on developmental level, it would appear from the evidence that by age 9 children are able to make substantial use of linguistic constraints in the text itself as contextual information and are able to use induced imagery, both of which make the value of overt illustrations redundant. Neither of these factors operate for children at the lower developmental level, however, so that illustrations are more likely to constitute a valuable source of contextual information at age 7. The evidence from Mackworth (1972^b) modifies this somewhat in that the redundancy of illustration as contextual information is unlikely to be as great for low progress readers¹ as it is for high progress readers. (On this evidence, therefore, the lower developmental level is weighted as +2 but, at the higher level, the high progress readers are weighted as 0 while the low progress readers are weighted as +1).

With evidence from studies on illustrations and reading, the situation appears more complex. For the lower developmental range there appears to be a clear contextual gain but no loss for the high progress readers with illustration (+2), but only partial contextual gain and a likely loss (distraction) for the low progress readers (0). For the higher developmental range there appears to be a marginal, integrated gain for the high progress readers (+1) and for

¹ The terms 'high progress' and 'low progress' are more accurate descriptions of the sample groups selected in the study (see chapter IX) and will therefore be used consistently rather than the alternative, 'good and poor' description.

the low progress readers a more substantial, if feature specific, gain (+2).¹

The evidence for giving these weightings is admittedly circumstantial and, with the complexity of factors and processes involved, it is doubtful if the basis for assigning the weightings is firm enough to justify formulating specific hypotheses.

The intention, therefore, is to investigate the effects of illustrations at the two developmental levels and the two progress levels in the knowledge that the effect is unlikely to be equivalent over all groups. In effect, the argument in this chapter has been at the level of justifying investigation of these variables rather than predicting precise effects. In so far as the weightings have relevance, they serve as a loose prediction of what, on indirect evidence, may constitute the form of an interaction effect.

¹ It is assumed, in giving the weightings, that a relevant relationship exists between illustration and text (i.e., it is congruous, sequentially related and potentially meaningful).

PART B

EXPERIMENT II

The effects of illustrations on the accuracy and efficiency of oral reading message identification and on the comprehension of high and low progress readers at seven and nine year reading age levels.

CHAPTER VIII

AIMS AND EXPERIMENTAL DESIGN

1. Aims

The aims of the experiment have already been set out in the introduction (chapter I) within the elaborated aims of the thesis as a whole. At this point, therefore, all that is required in an explicit, reiterated statement of specific aims:

- i) To test the contextual hypothesis under conditions that fully meet its assumptions.
- ii) Within this, to test the effects of illustrations on;
 - a) the accuracy of oral reading message identification,
 - b) the efficiency of oral reading message identification as evaluated in strategies of information usage and self-correction,

and c) the effectiveness of literal and inferential comprehension.

iii) Within this, to compare the effects of illustrations over;

- a) seven and nine year reading developmental levels;
- and b) high and low reading progress levels.

2. Design

Given that the experiment aimed to compare the effects of reading with and without illustrations over two developmental levels as well as over two progress levels, a number of difficult design problems were generated. The most obvious design might have compared high and low progress readers at one grade level (say, grade I) with high and low progress readers at a higher grade level (say, grade III).¹ The reasons why such a design would not have been appropriate will serve to illustrate why the final design was selected.

Let the hypothetical design look as follows:

¹ The standard North American system of 'grades' is used in preference to the South African equivalent terminology as this can be confusing to a non-initiate.

TABLE 3:

HYPOTHETICAL 2 x 2 x 2 FACTOR DESIGN WITH
ILLUSTRATION CONDITION, PROGRESS LEVEL AND
DEVELOPMENTAL LEVEL AS TREATMENTS

		P1(ILLUSTRATION)	PO(NO-ILLUSTRATION)
GRADE I (CA: 6)	HIGH PROGRESS (RA: 7)	} Story content and readability at average grade I level	
	LOW PROGRESS (RA: 5)		
GRADE III (CA: 8)	HIGH PROGRESS (RA: 9)	} Story content and readability at average grade III level	
	LOW PROGRESS (RA: 7)		

Two problems would have arisen in such a design. The first problem would have been across developmental levels. It would not have been possible to compare results at grade I with grade III if the reading at the two levels had taken place on stories with different content, readability level (Gilliland, 1972) and illustration since each of these factors, let alone their effects in combination, would be likely to cause differences on the dependent variables that would obscure the experimental effects.

Conversely, it would not have been possible to have grade I and grade III children reading precisely the same material because the readability of the material would have been either too difficult for grade I or too easy for grade III. This factor, in particular, constituted a crucial constraint in the design. Williamson and Young (1974), and Kibby (1979), for instance, have demonstrated that readers' strategies, as gauged from error acceptabilities, vary with the readability level of the material; especially between 'instructional' and 'frustration' level (Betts, 1950). In terms of the centrality of strategy as a dependent variable, therefore, it became important to control the relative difficulty of the reading material for each experimental group.

The only possible solution was to equate content and illustration across the two developmental levels but to vary the readability level of the material to fit, as nearly as possible, an instructional level of reading difficulty for each group. With only the readability level of the material varying - and this systematically controlled since both groups would be reading at their instructional level - comparisons across groups would become possible.

The second problem is clearly related to the first. If high and low progress levels had been selected from the same grade level, then the readability problem would have been encountered once more. Even with appropriate grade level adjustment of readability, the material would still have been either too easy for the high progress readers

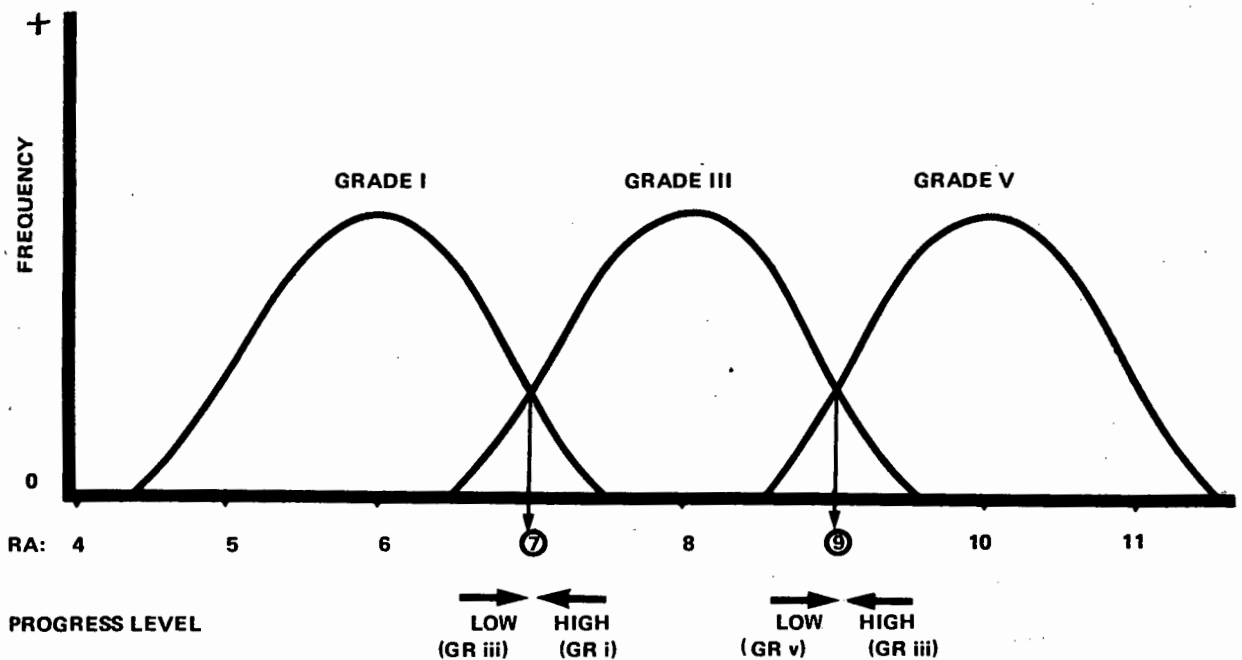
or too difficult for the low progress readers (approximately two years reading age difference). Thompson (1981b) has specifically criticised the majority of studies on good and poor readers for precisely this. As he argues, obtained differences between good and poor readers - particularly on such factors as strategy and comprehension - might simply reflect the difference in difficulty that the two groups experience in reading the same material. These differences might not emerge were the two groups reading at a comparable level of difficulty.

Given the necessity, therefore, to have material at, as near as possible, an instructional level of reading difficulty for each progress level within each developmental level there were only two possibilities. The most obvious would have been to have developed material at the four levels of reading difficulty with parallel content and illustrations. Apart from the practical problem of developing three stories (see below) at each of four different difficulty levels, the main problem with this solution would have been developing 300 word stories (minimum necessary for sufficient errors to be generated - see chapter IX) for grade I low progress readers (RA: 5): they would simply not have coped. Since the particular span of reading development, between approximately seven and nine years, was theoretically important (chapter VII) the alternative solution was to define developmental level in terms of reading age (RA: 7 & 9 respectively) and to select younger children (high progress relative to their

own grade level) who were reading at reading age x, together with older children (low progress relative to their own grade level) who were reading at the same reading age x. Thus:

FIGURE 2

SELECTION OF HIGH AND LOW PROGRESS GROUPS
AT TWO READING AGE LEVELS ACROSS HYPOTHETICAL
FREQUENCY DISTRIBUTIONS OF READING AGE
FOR GRADES I, III and V



The advantages in adopting this solution were that the theoretically important span was defined specifically in terms of reading development; and the readability problem between high and low progress groups was eliminated

through selecting them at the same reading ages. The readability level of the material, therefore, only had to vary between the two developmental levels of reading age seven and nine. The final design, therefore, was as follows:

TABLE 4

ACTUAL 2 x 2 x 2 FACTOR DESIGN WITH
ILLUSTRATION CONDITION, PROGRESS LEVEL AND
READING AGE DEVELOPMENTAL LEVEL AS TREATMENTS

		P1 (ILLUSTRATION)	PO (NO-ILLUSTRATION)
READING AGE 7:	HIGH PROGRESS: (GRADE I)	Stories: content 1,2,3; readability RA:7 (instructional) Illustrations: 1',2',3'	
	LOW PROGRESS: (GRADE III)		
READING AGE 9:	HIGH PROGRESS: (GRADE III)	Stories: content 1,2,3; readability RA:9 (instructional) Illustrations: 1',2',3'	
	LOW PROGRESS (GRADE V)		

It will be noted that three stories (1,2 and 3: see chapter IX), each with the same content and illustrations (1',2' and 3') across reading age levels, were read within each cell of the design. This served as a form of replication such that the particular content, or relationship of illustration to text, in any one story would not unduly influence the results.

This would allow conclusions to be generalized to, at least, other narrative reading material and illustrations of the general type represented across the three stories rather than limited to a particular story and its illustration - a weakness in some previous designs (Samuels, 1967, in particular).

In order to obtain data under the illustration conditions, two alternatives were available; a repeated measures design such as was used in experiment I, or a matched samples design. The repeated measures design used in experiment I, although theoretically more appropriate, had created minor interpretation problems where the two stories (A and B), although they had been matched for readability, still generated their own variance that threatened to confound the illustration effect. In that context, the problem was not serious as Groups I and II were design artefacts and not true independent variables. The story based variance could therefore be logically accounted for (chapter VI). Within the present design, however, with the existence of reading age and progress levels as true independent variables, such a problem would have been uninterpretable. With a matched samples design, on the other hand, it would not be necessary to vary the story across the two illustration conditions. Since different subjects would be reading under illustration and no-illustration conditions, the stories themselves could remain the same under the two conditions. For this reason, and on the evidence from experiment I that story related bias was a real possibility, the matched samples alternative was selected.

Within this final design, two problems remained both of which were unavoidable in terms of the readability requirement discussed above. The first related to the difference in readability level of the material for reading ages seven and nine. As was pointed out, however, this was controlled in the sense that the material was effectively at an instructional level of reading difficulty for each reading age. Moreover, since the actual differences in the material involved systematic variation in semantic and syntactic structure, it was felt that the influence of this variation would be logically interpretable in differences between reading age levels.

The second problem was the variation in grade level against the constant of reading age. In other words, in so far as a high progress reader at reading age seven, for example, is two grades younger than a low progress reader at reading age seven, this would be bound to introduce variance specific to the chronological difference. As pointed out, however, this choice was inevitable. The alternative would have been even more problematic. It was felt, nevertheless, that differences that could be attributable to grade level would be logically interpretable in terms of the known and systematic variation on this factor between the respective groups.

Experiment I had used 10 subjects in each cell but it was felt that the degree of variance on some of the dependent variables warranted more subjects. A target of 15 subjects within each cell of the design was therefore set.

Given the final design, three way analysis of variance was computed to test the significance of data for each dependent variable. Throughout the experimental report, factors and levels will be consistently represented as follows:

<u>Factor</u>		<u>Level</u>	
<u>Description</u>	<u>Representation</u>	<u>Description</u>	<u>Representation</u>
A. Illustration	P	1. Illustrations	P1
		2. No-illustrations	P0
B. Progress Level	Pr	1. High Progress	H
		2. Low Progress	L
C. Reading age Developmental Level	RA	1. Reading Age 7,3	7
		2. Reading Age 9,4	9

Unless otherwise indicated, the 5% level of significance was adopted for all analyses, and probabilities are consistently listed in the tables and the text as: + = $p < ,05$; ++ = $p < ,01$; +++ = $p < ,001$. Tables of individual variates for each dependent variable are provided in appendix 6.

CHAPTER IX

METHOD

1. Subjects and Sampling Procedure

According to the design the target samples were to consist of 30 high progress and 30 low progress readers at, or near, the seven year reading age level as well as 30 high progress and 30 low progress readers at, or near, the nine year reading age level. In order to select such samples, a population of 1868 children in grade I, grade III and grade V were tested on the D. Young Group Reading Test (Young, 1968) to obtain their reading ages.

i) Population considerations

The population was selected on the following counts:

a) Ethnic group

Only children from 'white' schools were used for the following reasons. In the South African context, considerable schooling differences exist between the 'white', 'coloured', 'indian' and 'black' education systems. Factors such as size of class, availability of reading materials, level of teacher training and competence and methods of teaching reading vary between the four systems creating, in effect, four different reading populations. Such differences would be accentuated by not only the language factor (see below) but also by the cultural background of the various groups

that would affect the relevance of both illustrations and story content. In addition, the study was designed around a problem where previous investigations had focussed on children in a 'first world', westernized context. Children in the 'white' education system in South Africa can be compared directly with others in this context, but the same would not hold true for children from the other educational systems much as this might be politically and educationally desirable.

b) Language

Only children with English as their first language were included. This meant, in fact, that only English medium schools were selected (as opposed to Afrikaans medium) and any children in the relevant classes whose home language was not English were also excluded. Since the screening test (Young, 1968) and the experimental reading material were in English the reasons for this are obvious.

c) Chronological age

Children whose chronological age deviated by more than one year from the median age for the relevant grades (as at 1st October: 7 years 1 month, grade I; 9 years 1 month, Grade III; 11 years 1 month, grade V. Cape Education Department, 1977) were also excluded. The reason for this was that, in selecting high and low progress readers on the basis of reading age, children could fall into a high progress category simply by virtue of being considerably older than their class peers and vice versa for the low

progress category. Since this would constitute an atypical reason for being classed as high or low progress, it was decided to exclude such cases to preserve a reasonable degree of homogeneity in the respective high and low progress groups.

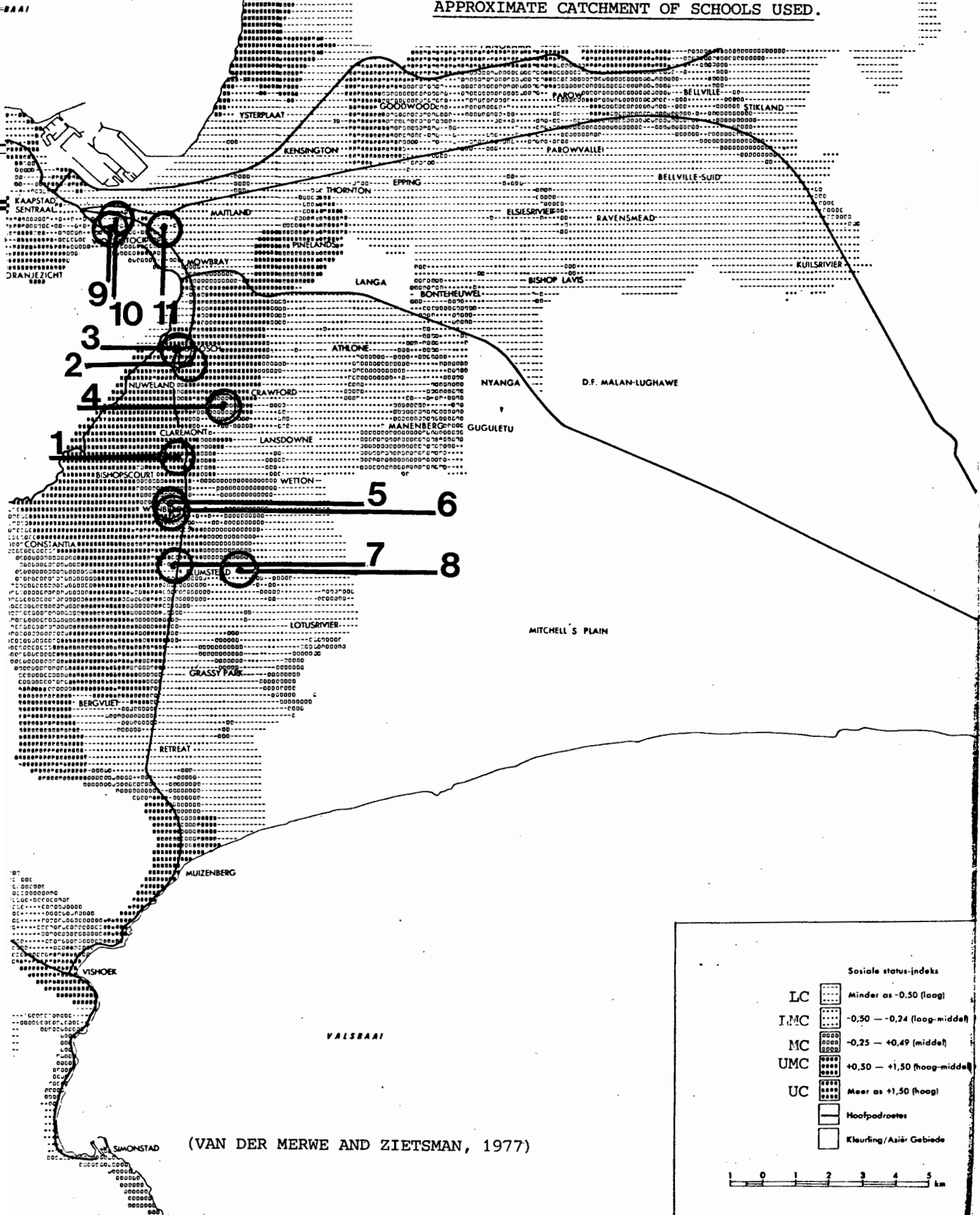
d) Socio-Economic status

By contrast, socio-economic factors can be considered as 'typical' or central in determining high or low progress in reading (Kellmer-Pringle et al, 1966). For this reason the population was deliberately selected from a variety of schools in the Cape Town Metropolitan area whose catchment would generally represent the typical range and distribution of socio-economic status in the South African 'white' group: Table 5; Figure 3.

e) Selection of grades I, III and V as target grades for screening

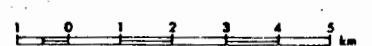
The decision to screen over these three grade levels was taken on the basis of a pilot application of the D. Young Group Reading Test (Young, 1968) to the first five grade levels in an average middle class school. The results of this pilot survey indicated that better than average readers in grade I were reading at much the same reading age level (\pm RA: 7) as poorer than average readers in grade III. Similarly the better than average readers in grade III were reading at much the same level as the poorer than average readers in grade V (\pm RA: 9). Since this spanned, exactly, the crucial developmental period identified earlier (chapter VII)

SOCIO-ECONOMIC DISTRIBUTION BY GEOGRAPHICAL AREA
OF THE POPULATION OF CAPE TOWN: POSITION AND
APPROXIMATE CATCHMENT OF SCHOOLS USED.



Sosiale status-inks

LC	Minder as -0,50 (laag)
IMC	-0,50 - -0,24 (laag-middel)
MC	-0,25 - +0,49 (middel)
UMC	+0,50 - +1,50 (hoog-middel)
UC	Meer as +1,50 (hoog)
[Symbol]	Hoofpadroetes
[Symbol]	Kleurling/Asiër Gebiede



(VAN DER MERWE AND ZIETSMAN, 1977)

TABLE 5

SOCIO-ECONOMIC CATCHMENT OF SCHOOLS USED
IN SAMPLE SELECTION

SCHOOL	NO.OF PUPILS SCREENED	SOCIO-ECONOMIC CATCHMENT
1	125	UC
2	181	UMC
3	180	UMC
4	91	MC
5	228	MC
6	223	MC
7	166	MC-LMC
8	340	MC-LMC
9	154	LMC-LC
10	30	LMC-LC
11	150	LMC-LC
	1868	

UC: upper class

UMC: upper middle class

MC: middle class

LMC: lower middle class

LC: lower class

Note: In this context these are intended as general descriptive labels only. A more precise definition of each child's socio-economic status was used in a later stage of the sampling procedure (p 210)

it was decided to limit the screening to grades I, III and V. (A similar overlap was clearly present at RA: 8 between grades II and IV but this was not regarded as relevant to the requirements of the design).

ii) Screening

a) The test

The D. Young Group Reading Test (Young, 1968) was selected as an appropriate screening test for the following reasons. First, it covers the range of reading ages under consideration (RA: 6 - RA: 10). Second, since the experimental situation was to require contextual reading and comprehension it was important that the screening test should at least approximate this task requirement. In other words, selection of the experimental groups in terms of reading age and progress level on the screening test would need to be reasonably predictive of those same groups' performance on the experimental reading material - at the least on word identification accuracy. In so far as the D. Young incorporates the reading of sentences the meaning of which need to be comprehended for the correct word alternative to be selected, it was felt that this requirement was at least partially satisfied. Satisfactory concurrent validity is also quoted with the Neale Analysis of Reading Ability (Neale, 1958) which involves full contextual reading and comprehension (Young-Neale accuracy; ,884; Young-Neale

comprehension; ,735. Young, 1968). Since only individual tests (e.g., the Neale) could provide a total task correspondence with the experimental situation, and since such tests would have been impracticable for screening purposes the D. Young was considered to be a reasonable compromise. That the reading age norms for the test are British was of no significance as the objective was to locate subjects at homogeneous reading age levels, not to compare them with their peers. (This fact nevertheless accounts for the apparent inconsistency where the median chronological age in grade I, at the time of testing was given as 7,1 years and high progress readers were selected at 7,3 years! This was, however, merely an artefact of the norms).

b) Application

Standard instructions were given to all 66 teachers of the classes involved. The tests were then applied over a period of one week in October, 1979, and were subsequently all scored by the experimenter. For each child, information from school records was also obtained concerning the father or caregiver's occupation. The distribution by sex over the three grade levels was as shown in Table 6.

TABLE 6

DISTRIBUTION BY SEX AND GRADE LEVEL FOR
THE TOTAL SCREENED POPULATION

	GRADE I	GRADE III	GRADE IV	TOTAL
GIRLS	256	293	319	868
BOYS	335	325	340	1000
TOTAL	591	618	659	1868
NO. OF SCHOOLS	11	11	11	11
NO. OF CLASSES	22	22	22	22

The slightly higher proportion of boys is due to school 1 being all boys while all other schools were either mixed or balanced by their respective brother/sister schools. It was decided not to balance school 1 as this would have overloaded the UC socio-economic factor which was regarded as more important than a perfect sex balance.

iii) Sample Selection

On the basis of the screening test results (appendix 7), the task was to select, at approximately the 7 year reading age level, a pool of high progress readers from grade I to correspond with a pool of low progress readers from grade III; and, at approximately the 9 year reading age level, a pool of high progress readers from grade III to correspond with a pool of low progress readers from

grade V. Additional constraints on this task were that the pools should be as homogeneous (in terms of reading age) as possible, and that there should be sufficient numbers in the pools to allow selection of the relevant sample numbers.

Using one standard deviation from the respective means as the basis for selection, the closest that this could be approximated, within the other constraints, was selection of those readers falling on score points 23, 24 and 25 for high progress readers in grade I and for low progress readers in grade III. This yielded two corresponding groups of 38 and 39 individuals respectively at RA: 7,27. Since mid-test scores are less sharply discriminative of reading age than peripheral scores (appendix 7) selection over three score points was still regarded as meeting the homogeneity criterion. Since the standard error of measurement of the test is given as 2,18 (Young, 1968 : p.22), and since the high progress grade I group was 8,5 points (1,29 SD) above the respective mean and the low progress grade III group was 10,8 points (1,57 SD) below the respective mean, this selection was regarded as entirely satisfactory and representative of true high and low progress groups.

Selection of the corresponding groups at approximately reading age 9 was more problematic. Not only are scores at the upper periphery of the test more sharply discriminatory of reading age (requiring selection at a single score point) but for the grade V results it was clear that since the median score was in fact the ceiling of the test (appendix 7) that the criterion of one standard deviation from the respective

mean would not be a true reflection of low progress for this group. A ceiling effect was also partially true for the grade III group. Given these problems, selection was made largely in terms of inspection and to meet the other criteria of homogeneity and sufficient numbers. This yielded score point 41 (RA: 9,4) as the most appropriate correspondence point. For grade III this was 6,2 points (0,90 SD) above the respective mean and well outside the error of measurement. There were 39 individuals at this point for both grades. For grade V however, score point 41 was only 1,27 points (0,35 SD) below the respective mean. Given the marked ceiling effect for this group, however, and the expectation that the mean should be reasonably close to the median chronological age of 11,1 years, it can safely be assumed that, had the upper levels of the normal distribution been available for this group, the mean would have been equivalent to \pm RA: 10,5 years. With a reading age of 9,4 selected as representing low progress, this matches the difference of approximately one year's reading age from the mean as selected for the corresponding high progress (grade III) group. On this evidence, and given that the other criteria were adequately met, this was regarded as a representative selection of high and low progress readers at this reading age level.

For each of the children in the selected sample pools, the father or caregiver's occupation was rated on the following socio-economic scale: Table 7:

TABLE 7

SOCIO-ECONOMIC SCALE USED FOR SAMPLE
MATCHING ACROSS ILLUSTRATION CONDITIONS

RATING	A	DESCRIPTION
4	UPPER MIDDLE CLASS:	Professional; semi-professional; higher administrative; executive
3	MIDDLE CLASS:	Lower administrative; clerical; sales personnel
2	LOWER MIDDLE CLASS:	Skilled artisans; trained field workers
1	LOWER CLASS:	Semi-skilled and unskilled workers.

Note: This scale is a condensed version of the scale used by Roos (1970, p.37) in an extensive educational survey of South African 'whites' conducted by the South African Human Sciences Research Council.

In each pool individual subjects were then paired off according to their socio-economic rating (4 with 4; 3 with 3, etc.) In random order, each pair was divided between the illustration (P1) condition and the no-illustration (P0) condition until 15 pairs in each pool had been allocated. The remainder constituted either unpaired individuals or unallocated pairs. The resulting socio-economic, grade, chronological age and sex distribution within the final design samples was as follows: Table 8.

TABLE 8

FINAL CELL SAMPLES BY SOCIO-ECONOMIC RATING,
GRADE LEVEL, CHRONOLOGICAL AGE AND SEX

CELL	MEAN SOCIO- ECONOMIC RATING	GRADE	MEAN CA		SEX		TOTAL
			YRS	MTHS	M	F	
7,H,PI	3,6	I	6	8	4	11	15
7,H,PO	3,6	I	6	9	8	7	15
7,L,PI	2,2	III	9	2	8	7	15
7,L,PO	2,2	III	9	2	13	2	15
9,H,PI	3,3	III	8	10	6	9	15
9,H,PO	3,3	III	8	9	6	9	15
9,L,PI	2,1	V	11	4	7	8	15
9,L,PO	2,1	V	11	3	5	10	15
TOTAL					57	63	120

A number of observations from Table 8 are of interest. First, the consistent difference between socio-economic rating for the high and low progress groups re-inforces the well-established finding that reading progress is directly affected by the socio-economic factor (Kellmer-Pringle et al, 1966). This will necessarily, therefore, have some bearing on interpretation of the results.

Second, through random allocation, the mean chronological ages emerge as sufficiently matched between the PI and PO

groups and that this could be assumed in interpretation. Certainly sufficient evidence exists on the relationship between intelligence and reading progress (Dechant and Smith, 1977). On the question of matching the PI with the PO groups, it was argued that this would be sufficiently controlled through random allocation together with the matched socio-economic ratings.

2. Materials¹

Since the aim was to compare, directly, the oral reading and comprehension of children not only at two different reading age levels, but also at two different progress levels (or seen alternatively at three different grade levels), this generated a number of problematic constraints on the development of appropriate experimental reading material.

i) Content

The first requirement was that the stories to be read should have a basic content that would be meaningful to all children participating in the experiment. Given the grade level variation, two possibilities existed; either to find three levels of story with age-appropriate content for each grade level or to find the sort of content that

¹ All six stories are reproduced, together with black and white reproduction of the relevant illustrations, in appendix 2. A single example of the actual colour illustrations used is provided on p.221.

As far as possible, the content was held constant across the two readability levels such that vocabulary changes involved essentially synonymous meanings and syntactic changes involved simpler or more complex forms, respectively, of the basic sentence material.

Since three different stories were developed at each of the two readability levels (reasons for this are given in v. below) it was desirable to have as close a readability match between the three stories at any one readability level as possible. Readability formulae, such as the one used, are useful but can be relatively insensitive to degrees of difficulty in vocabulary, syntax as well as logic (Gilliland, 1972). For this reason, stories were not only matched at the appropriate readability level as given by the Spache formula but additional criteria were also used in establishing the match: viz. number of 'very difficult words' (based on an estimate of relative frequency of usage and difficulty of meaning); number of repeated 'difficult' and 'very difficult' words; and number of embedded sentences.

In addition it was essential that the stories should be pitched at, or as near as possible, an instructional level of reading difficulty for the respective reading age groups (chapter VIII). Constructing the stories to an appropriate readability index was a necessary first stage in order to achieve this. The second, and equally necessary stage, involved pre-testing the stories on a representative sample of children in order to check and, if necessary, refine the readability levels. This was achieved through

would be acceptable to all three grade levels. The first alternative, although more natural, would have complicated - and very possibly invalidated - comparisons (chapter VIII). Although such stories could have been controlled for the appropriate level of reading difficulty, differences in content, as any reading teacher knows, exert considerable influence on oral reading and comprehension. The second alternative was therefore favoured. Simple adventure stories in which the characters, activities, and settings would, as far as possible, allow identification from grade I to grade V children were selected. The basis for these stories was taken from Shiela McCullagh's reading series, One, Two, Three and Away (McCullagh, 1965).

ii) Readability

Having selected in terms of uniform content, the next requirement, within this, was to set the stories at the two readability levels that would allow readers at the reading ages of seven and nine to read, respectively, at the instructional level of difficulty (chapter VIII). In order to achieve this, the stories were partially re-written making use of the Dale list of 769 'easy words' (Dale, 1931) to provide a basis for vocabulary difficulty and the Spache readability formula for primary grade reading material (Spache, 1953) to provide an overall basis for measuring syntactic and semantic difficulty¹.

¹ The Spache formula, as opposed to others (Gilliland, 1972), was chosen as it relates specifically to primary level material and was, therefore, most relevant.

using those children who had participated in the pilot run of the screening test and who were therefore not included in the final sampling. The groups were necessarily small and a slightly wider reading age range than was selected in the final sampling had to be used to locate children at the respective reading age and progress levels: Table 9.

TABLE 9

SAMPLE USED FOR PRE-TESTING
EXPERIMENTAL READING MATERIAL

READING LEVEL	NUMBER IN SAMPLE	MEAN RA	RANGE (MTHS)
7,H	3	7 - 3	2
7,L	3	7 - 4	2
9,H	2	9 - 2	3
9,L	3	9 - 3	0

The socio-economic catchment of the school concerned was average middle-class. The deficiency in numbers was largely compensated for by requiring each child to read all three of the stories designed for his reading age level. The nett result of this was a pool of 18 stories read at reading age seven level and a pool of 15 stories read at reading age nine level. Since the primary objective was to check the readability of the material on readers at these two levels,

this was regarded as an adequate data base for the purpose. The stories were scored for word identification accuracy and total comprehension and only the illustrated versions were used: Table 10.

TABLE 10

RESULTS OF PRE-TESTING EXPERIMENTAL READING
MATERIAL FOR ACCURACY AND COMPREHENSION

READING LEVEL	MEAN ACCURACY	MEAN COMPREHENSION
7H	95,73	47,54
7L	94,62	63,71
7: COMBINED	95,18	55,63
9H	98,01	65,63
9L	96,77	77,21
9: COMBINED	97,39	71,42

The definition for instructional level reading as used by Christenson (1969) is that reading level at which the reader achieves between 95% and 98% accuracy and/or 75% to 90% comprehension. On the basis of these results it appeared that stories were being read at instructional level - at least on the criterion of accuracy. Comprehension was somewhat low particularly in the grade I group, but it was felt that this was to be expected in terms of chronological age. Informal reading inventories usually rely only on

literal recall as the basis for judging comprehension (Betts, 1950) so that the lower level of comprehension on the stories under consideration was regarded as adequate given the more difficult comprehension questions being asked. (In the final results, the average cued recall of main ideas for the P1 condition - most comparable to the IRI criterion - was 72% which is nearly enough adequate).

It was nevertheless felt that the reading difficulty at the reading age 9 level was marginally too easy, even although it fell within the instructional range. Since much of the data was to depend on errors and their analysis it was also desirable that not too few errors were made. In other words a very delicate balance had to be struck between subjects making enough errors to allow reliable analyses as well as their reading at instructional, not frustration level. On these grounds, the stories at reading age 9 were each altered to include fifteen more words classed as 'very difficult' (the pre-test had indicated the approximately 3 out of 5 such words were being correctly identified).

In the final results, the average word identification accuracy for all groups under all conditions was 93,54%. This was somewhat lower than the planned optimum of 95% owing to the effects of reading unillustrated text. For the illustrated condition, however, - the basis on which the pre-testing had been done - the average was, in fact, 95,17%; a very close approximation to the optimum. In effect this meant that children reading the illustrated

text were reading at instructional level while those reading unillustrated text were marginally into frustration level reading.

iii) Illustration

Two factors were crucial in determining and defining the nature of illustration as used in his experiment; the relevance of illustration to the content of the text, and the frequency of illustrations within the text.

a) Relevance

As has already been pointed out (chapter V), several studies have shown that the effects of illustration on reading - and particularly on comprehension - are in part determined by the relationship of relevance that exists between the illustrations and that text. The aim of this experiment was not to explore this relationship further but, by specifying and controlling it, to investigate the effects of relevant illustrations, per se.

To specify relevance, therefore, the following criteria were established as minimum requirements for the relationship between illustration and text:

1. Each illustration to be depicted in clear and unambiguous line with a minimum of decorative or non-essential detail. This follows Dwyer's (1970) finding that simple line drawings were more effective than more complex treatments.

2. Each illustration to depict central substantive information in that section of the text to which it refers; and to be sequentially related to that section of the text. This does not imply that an illustration should be exhaustive in its depiction of textual information but that it should at least refer to what is currently in the text and that what it depicts should be central, and not incidental, to the information in that section of the text. This follows various findings (chapter V) that directly and congruously illustrated textual information is retained better than incongruously, indirectly or non-sequentially related information.

On the basis of these two minimal criteria a good proportion of children's early reading material can be regarded as relevantly illustrated. Nevertheless it is also true that there are many single texts and reading schemes that could also be regarded as not having relevant illustrations. Another reason for selecting stories from the Shiela McCullagh (1965) series for use in this experiment was that the illustrations do fit these criteria: Figure 4 and Appendix 2.

FIGURE 4

COLOUR EXAMPLE OF ILLUSTRATIONS USEDb) Frequency

The frequency of illustrations in a given text, or the ratio of the number of illustrations to the number of words has not been investigated as a factor in the role of illustrations in reading. Nevertheless, its importance is obvious. One illustration per 200 words, for instance, is likely to have a different effect to one illustration per 50 words if only because any one illustration can only carry a limited display of relevant and current information. This is more likely to match the text for the smaller ratio than for the larger. Samuels' (1967) ratio

of 1:106, for first grade readers, for example, would appear to have been rather unrealistic and, despite his claim, hardly representative of normal classroom reading conditions. A brief survey of books from three well-established basal reading series (One, Two, Three and Away, Shiela McCullagh, 1965; Janet and John, O'Donnell and Munro, 1950; and Happy Venture, Schonell, 1939) reveals that for a reading age of between 6 - 6 and 7 - 6 the average ratio is 1:50. This is deliberately conservative. On more contemporary, although less well established, reading material the ratio would be considerably less. It was clear, therefore, that a ratio of at least 1:50 should be used for the seven year reading age level in this experiment. The question of what ratio to use at the nine year reading age level, however, was more problematic. On the one hand the ratio is usually larger for older reading ages but, on the other hand, comparisons between groups would have been largely invalidated without an equal basis for the illustration effect (chapter VIII). It was therefore decided to retain the same basic ratio for both reading age levels.

Once again, the McCullagh series was found to be appropriate. The re-writing of the original stories changed the ratios slightly and, in effect, because the stories at reading age 9 were slightly longer than those at reading age 7, the ratios for the former group were marginally larger: Table 11.

TABLE 11

RATIOS OF ILLUSTRATION TO NUMBER OF WORDS
IN EXPERIMENTAL READING MATERIAL

STORY	RA: 7	RA: 9
1. The Donkey	1 : 34	1 : 37
2. The Island	1 : 39	1 : 43
3. The Giant	1 : 34	1 : 37

Apart from the ratio, the absolute number of illustrations was also considered important. As has been emphasised, giving readers several sequentially related illustrations within any one story is not only likely to increase the cumulative contextual relevance of the illustrations (chapter V) but it also relates to normal reading conditions and ensures that any effects are not due to the characteristics of any one particular illustration. (This was one of the specific deficiencies in experiment I). An average of nine illustrations therefore appeared in each story: Table 12.

iv) Story Length

The decision as to how long the stories should be was dictated by several considerations. First, there had to be sufficient length (number of words) for children, reading at instructional level, to make enough errors to

allow for error analyses (chapter IV). An aim was therefore set to generate an average of not less than 15 errors per analysis. With children reading at instructional level ($\pm 95\%$ accuracy) therefore, a minimum of 300 words per story was indicated.

Second, how long a story the grade I children could be expected to cope with in one sitting was also a consideration. Again, it was felt that a story of approximately 350 words would be a maximum: that to make it any longer would create a very abnormal oral reading task for children at this level. Given this constraint, and the additional constraint that all stories (at reading ages 7 and 9) had to cover the same basic content (p. 213), this meant that all stories would have to approximate the maximum of 350 words.

Third, given McCullagh's original stories (± 600 words), it was desirable to end the experimental section at a point that made some sense; a plausible termination point, if not actual ending, to the story.

Taking all these considerations into account, the final story lengths, and number of illustrations, were as follows: Table 12.

TABLE 12

NUMBER OF WORDS AND NUMBER OF ILLUSTRATIONS
PER STORY IN EXPERIMENTAL READING MATERIAL

STORY	<u>RA7:</u>	WORDS	ILLUS- TRATIONS	<u>RA9:</u>	WORDS	ILLUS- TRATIONS
1. The Donkey		309	9		336	9
2. The Island		311	8		341	8
3. The Giant		308	9		334	9

v) Replication

Since the relationship between any one story and its set of illustrations must, to some extent, be unique it was decided, despite the difficulties of matching for readability, to deliberately vary this within the design (chapter VIII). In other words, to have selected one story to be read with and without illustration would have opened the experiment to criticism on the grounds that any effects could be due to the relationship between that particular story and its illustrations and that the same effects might not appear given a different story and set of illustrations. Replication was therefore built into the design such that three different stories, each with their own set of illustrations, were to be read within each cell of the design. Given that the text was 'narrative'

in all three cases, this would allow any illustration effects to be generalized to, at least, other situations involving the reading of narrative text in general.

vi) Production and Format

For the illustrated versions, the format followed the original McCullagh (1965) stories. In other words, that section of text to which an illustration referred was located immediately below the relevant illustration. The cover illustration was retained but both text and illustrations following the termination point of the experimental section of the story were covered with blank white sheets.

For the no-illustration version, all illustrations were covered. Since the same format as used in the illustrated version would have appeared unnatural - large blank pages with a small section of text - the text in the no-illustration versions was presented continuously. The same section divisions as dictated in the illustrated versions were nevertheless retained as clear paragraph breaks so that the text, in both versions, was broken up into the same 'chunks'.

For each reading age level, print size and word spacing was matched with that used in the McCullagh series for the same approximate reading age levels.

All the above is evident in the reproductions of stories and illustrations as set out in appendix 2.

3. Experimental Procedure

Each of the 120 subjects was given the experimental task by the experimenter. This ensured maximum consistency of procedure over all subjects. Subjects were tested randomly according to the school they attended and the entire procedure was completed over the first three weeks of November, 1979 in order to minimize developmental changes beyond the reading levels as determined on the screening test (October, 1979).

Testing took place in a quiet room in each of the schools concerned. This was specifically requested to minimise distraction or interruption during the reading of the stories. Subjects were set at ease and the following task instruction was given to all, irrespective of experimental condition.

"I'm trying some stories out to see if they are right for children of your age. I'd like you to read one for me, and then I'll ask you some questions about it when you're finished. I'd like you to read it without help, so if you come to a word that is difficult, just try to work it out or take a guess"

It should be noted that under this 'reading directed' task expectation, none of the subjects spent time glancing through the illustrated versions before beginning to read. Under normal reading conditions, where a child is more relaxed, this would usually occur so that one could probably assume a stronger illustration effect under such conditions.

Specifically drawing the reader's attention to the illustrations in the illustration condition was considered but it was felt that this would be equally artificial and that if the effect emerged without doing this then the test of the hypothesis could only be strengthened.

The instruction to "... read it without help ... work it out or take a guess" was felt to be necessary for two reasons. First, it would minimise experimenter interaction with the reader during the course of reading and apart from maintaining standard procedure for all subjects it would, specifically, avoid any differential in oral feedback (see criticisms of the focal attention paradigm; chapter V) to subjects under the two experimental conditions. Second, since error analysis was to constitute the basis for several dependent variables, it was important that subjects make their own attempts - and have an expectation to do so - rather than looking to the experimenter to provide answers. Once reading was started, therefore, the only interaction permitted was general encouragement.

Following the oral reading, comprehension questions were asked in the form in which they appear in Appendix 3. No deviation from the form of questions was permitted but questions were repeated if the subject appeared uncertain. The procedure was tape-recorded continuously from the beginning of oral reading to the last comprehension response.

In order to minimize distraction, no written observations were made at all during the test procedure.

The three stories, The Donkey, The Island and The Giant were assigned, prior to testing, in rotation order to the list of subjects in each cell of the design. This ensured that each story was read, randomly, five times within the fifteen subjects in each cell.

4. Transcriptions of Oral Reading and Comprehension Responses

Triple-spaced type-written transcripts of all six stories were prepared on which all relevant data from the tape-recorded test sessions could be entered (appendix 4). The following data were recorded.

i) Oral Reading Errors

Deviations from the text, at the whole word level, were written out or coded according to the following system (based on Donald, 1979b): Table 13.

Several points need to be made about this system.

a) As Weber (1968) has pointed out, one of the most frequent confusions between studies of oral reading errors has been over the level of categorization. Taking the same examples used in chapter IV, an error such as 'he ^{ran} was running', taken at the morpheme level, would involve three errors; at the whole word level, two errors; and at the phrase level, only one error. The main problem with both the morpheme and the phrase levels of

TABLE 13

TRANSCRIPTION AND CODING SYSTEM
FOR ORAL READING ERRORS

ERROR TYPE	CODE
Substitution	The river rushed ^{over} <u>past</u> threateningly underneath ...
Non-word substitution	The river rushed past <u>'thragily'</u> <u>threateningly</u> underneath ...
Non-word substitution with clearly meaningful stem	The river rushed past <u>"threatenly"</u> <u>threateningly</u> underneath ...
Addition	The river rushed past threateningly ^{and} underneath ... ^
Omission	The river rushed (<u>past</u>) threateningly underneath ...
Self-Correction	The river ^{C was} <u> </u> rushed past threateningly underneath ... ^

categorization, although they might each be linguistically justified, is that they do not correspond to the conventional measures of 'accuracy' which are based on the whole word level (e.g., Neale, 1958). In addition, they both generate difficult decision problems in scoring that would

inevitably reduce reliability. For these reasons, and to maintain consistency, all errors in this study were taken at the whole word level.

b) The system as set out in Table 13 differs somewhat from other error categorization systems that have been developed for essentially diagnostic purposes (Donald, 1979b; Pumfrey, 1976). For example, reversals have clear diagnostic significance but, within the purposes of the present experiment, it would have contributed little to categorize these errors separately: they were simply treated as substitutions. The same holds true for successive attempts and for sounding out. In the case of successive attempts - a form of repetition - only the final attempt was taken as the error. This avoided undue inflation of error scores. Where repetitions took the form of 'running starts' - i.e., repetition of a correctly read word or phrase - (Goodman and Burke, 1973) they were not counted as errors. This follows Weber's (1968) and Goodacre's (n.d.) opinion that such repetitions are not in the nature of errors at all and, if taken as such, serve only to distort the measure of word identification accuracy. Punctuation errors were also ignored. Although the degree to which such errors reflect the use, particularly, of syntactic information would have been of relevance in the experiment, the reliability of detecting them is so low (Weber, 1968) that it was regarded as better to ignore them with consistency than to include them with doubtful reliability. Finally, refusals as a category of error did not occur

simply because readers expected, and were encouraged, to make some response (see Procedure).

c) In order not to distort the measure of orthographic error acceptability, non-word substitutions were transcribed, not phonically, but in a form that retained the spelling of the stimulus word as closely as possible (as recommended by Goodman and Burke, 1973). For example, the word, menacingly, pronounced as the non-word (*mEnə^{tr}I*) would have been transcribed as 'menacily', not as 'menakly'. (Since comparisons with the standard orthography of the text were at stake, phonetic transcription would also have been inappropriate).

d) Proper nouns occurred quite frequently throughout the stories used in the experiment. Once again, to avoid as artificial distortion of both accuracy and the measures of average semantic/syntactic/orthographic error acceptability, substitution of a proper name (e.g., John for Johnny) was taken as an error on its first occurrence only. Thereafter, provided the substituted name was used consistently, it was ignored. Any variation, however, was taken as a new error.

ii. Comprehension Responses.

Subjects' answers to the comprehension questions were transcribed verbatim.

5. Scoring of Dependent Variables

Using the data transcribed according to the system above, the scoring of each dependent variable for each subject was as follows:

i) Word Identification Accuracy

The number of uncorrected errors was subtracted from the total number of words in the story concerned. The resulting total, expressed as a percentage of total words in the story, represented the accuracy with which words had been identified in the story.

ii) Semantic Acceptability of Errors

Each error including those that were corrected¹, was scored

¹ The inclusion of self-corrected errors in the semantic/syntactic/orthographic error acceptability analyses was justified on the grounds that an error, before it is corrected, is not qualitatively different from an uncorrected error: both must necessarily reflect the process of information selection as used by the reader. Significantly, their inclusion also served to boost the number of errors per analysis. As pointed out earlier, the total number of words per story was narrowly planned to allow, at an instructional level of reading, a minimum number of errors for the purpose of error analysis without, at the same time, making the stories unrealistically long for children in grade I. The inclusion of self-corrected errors (not calculated into the word identification accuracy score) in the analyses was partly instrumental in the final average number of errors per analysis being as high as 26 (total number of errors analysed was 3140: N = 120). Since the purpose of each error acceptability analysis was to determine a characteristic or representative measure of the reader's use of semantic, syntactic and orthographic information, the relatively high number of errors analysed per analyses (cf precedents, chapter IV) could only increase the reliability of these measures.

SCALE	CRITERIA
4 (cont.)	<p>iv) <u>Contractions or Expansions</u></p> <p>e.g. E He's</p> <p>S <u>He</u> is ...</p> <p>E cannot</p> <p>3 ..<u>can't</u>..</p> <p>v) <u>Pronoun/proper name substitution</u> (sex and number appropriate)</p> <p>e.g. E Roger</p> <p>S <u>He</u> pointed ...</p> <p>vi) <u>a for the substitution</u> (where referent has not appeared before)</p> <p>e.g. E a</p> <p>S Roger stopped near <u>the</u> gate ...</p>

- 3 E INVOLVES A NON-SUBSTANTIVE CHANGE IN MEANING THAT DOES NOT ALTER THE BASIC, CUMULATIVE MESSAGE OF THE TEXT : E IS ACCEPTABLE WITHIN THE DEEP STRUCTURE OF THE TEXT.

Particular Cases:

- i) Verb tense or verb form changes

e.g. E will

S you ^ climb up ...

E Holding

S Hold on tightly ...

on the following scale for its degree of semantic acceptability within the meaning of the text.

TABLE 14

SCALE AND CRITERIA FOR SCORING
SEMANTIC ACCEPTABILITY OF ERRORS

SCALE	CRITERIA
4	<p>E (ERROR) IS SYNONYMOUS WITH S (STIMULUS) : E IS ACCEPTABLE WITHIN THE TEXT IN TERMS OF BOTH DEEP AND SURFACE STRUCTURE.</p> <p><u>Particular Cases:</u></p> <p>i) <u>Substantively synonymous substitutions:</u></p> <p>e.g. E Among gentle</p> <p>S <u>Amongst gently</u> rolling hills ...</p> <p>..(both score 4)</p> <p>E afraid</p> <p>S ... was not <u>frightened</u></p> <p>ii) <u>Additions or omissions that preserve meaning and acceptable syntax:</u></p> <p>e.g. E you</p> <p>S Shall I help ^ ?</p> <p>E/S ... help me (to) pull it in.</p> <p>iii) <u>First occurrence of name change:</u></p> <p>e.g. E John</p> <p>S <u>Johnny</u> ...</p>

SCALE	CRITERIA
1	E IS SEMANTICALLY ACCEPTABLE WITHIN NEITHER GENERAL NOR LOCAL MEANING CONTEXT BUT IS A LEXICALLY ACCEPTABLE WORD
	e.g. E sat S He <u>set</u> off through the forest E first S ..raised his <u>fist</u> and...
	<u>NOTE:</u> Only substitutions and additions can score 1: Omissions can score only 4,3,2 or 0.
0	E IS SEMANTICALLY ACCEPTABLE ON NONE OF THE ABOVE CRITERIA
	e.g. E 'mencily' S ..glared <u>menacingly</u> .. E/S .. the <u>river</u> rushed past..

The sum of all semantic scores was expressed as a percentage of the total possible score for all errors (i.e., total errors x 4). This represented the average semantic acceptability of that subject's errors: the relative tendency of that subject to use available semantic information in his word identification strategy.

iii) Syntactic Acceptability of Errors

Each error, including those that were corrected, was scored

on the following scale for its degree of acceptability within the syntactic structure of the text (or of the subjects own continuous errors.)

TABLE 15.

SCALE AND CRITERIA FOR SCORING
SYNTACTIC ACCEPTABILITY OF ERRORS

SCALE	CRITERIA
4	<p>E (ERROR) IS SYNTACTICALLY ACCEPTABLE TO BOTH THE PRECEDING AND THE FOLLOWING STRUCTURE OF BOTH ITS SENTENCE AND THE PASSAGE</p> <p><u>Particular Cases:</u></p> <p>i) <u>All name change errors.</u></p> <p>ii) <u>Errors where E may not necessarily be acceptable to the meaning context but is syntactically interchangeable with S.</u></p> <p>e.g. E cat</p> <p> S ..the <u>kite</u> was swept higher..</p> <p> E up</p> <p> S They ran <u>down</u> the hill</p> <p>iii) <u>Non-word substitutions where E clearly has the same syntactic suffix/inflection as S:</u></p> <p>e.g. E 'hesi<u>section</u>' (noun)</p> <p> S <u>hesitation</u></p> <p> E <u>fasted</u> (verb)</p> <p> S <u>fastened</u></p>

SCALE	CRITERIA
4 (contin.)	<p>E 'ferasome' (adjective)</p> <p>S <u>fearsome</u></p> <p>E 'agrily' (adverb)</p> <p>S <u>angrily</u></p>
	<p>iv) <u>Errors that change the syntax of the text but still result in a syntax that is acceptable within sentence and passage context.</u></p> <p>e.g. E steal</p> <p>S You have no right to <u>be</u> <u>stealing</u> my donkey. (both score 4)</p>

3 E IS SYNTACTICALLY ACCEPTABLE TO BOTH THE PRECEDING AND THE FOLLOWING STRUCTURE OF ITS SENTENCE OR CLAUSE BUT NOT TO THE PASSAGE

Particular Cases:

i) Verb tense changes (in most cases; not all):

e.g. E runs

S She ran down the hill... (passage in past tense)

ii) Number changes (in most cases: not all):

e.g. E islands

S Alongside the island (single island previously mentioned)

iii) Pronoun changes (in most cases: not all)

e.g. E she

S .. he explained

	4	4	4	
e.g. E	forward and took			
S	The donkey came	<u>forwards</u>	<u>to</u>	<u>take</u> the apple.
		2	2	2
E	and starting quickly			
S	Filling his pipe	<u>he</u>	<u>stared</u>	<u>quietly</u> into the distance.
	0	4	2	
E	'persad'	you	under	
S	Let's <u>persuade</u>	<u>him</u>	<u>underneath</u>	with ...

(not a continuous string: corrected error intervenes)

The sum of all syntactic scores was expressed as a percentage of the total possible score for all errors (i.e., total errors x 4). This represented the average syntactic acceptability of that subject's errors: the relative tendency of that subject to use available syntactic information in his word identification strategy.

Note: The scoring of non-word substitutions, with suffixes/inflections matching the stimulus word, on the maximum of the syntactic scale (4 : iii above) should, on Potter's (1980) argument (chapter IV), have confounded further the relationship between the orthographic and syntactic acceptability ratings. In fact, the correlation between scores on the two scales (chapter X) was $-.103$; a non-significant relationship. Further, since syntactically constrained non-word substitutions comprised a significant proportion of all errors, and since most of these scored 0

semantically, another effect was to reduce the inevitable relationship between scores on the syntactic and semantic scales (the relevant correlation being ,440 which, although significant, was considerably lower than that obtained in experiment I (where these errors had been scored as syntactically 0).

iv) Orthographic Acceptability of Errors

Each error, including those that were corrected, was scored on the following scale for its degree of orthographic acceptability to the equivalent stimulus word in the text.

TABLE 16

SCALE AND CRITERIA FOR SCORING
ORTHOGRAPHIC ACCEPTABILITY OF ERRORS

SCALE	CRITERIA
5	E (ERROR) AND S (STIMULUS) CONTAIN IDENTICAL LETTERS, IN THE SAME SEQUENTIAL POSITION e.g. E 'island' 'giant' (pr.'geeant') wind S island giant wind
4	E DIFFERS FROM S BY <u>ONE</u> LETTER, ALL REMAINING LETTERS BEING IN THE SAME SEQUENTIAL POSITION e.g. E strong was fire a waited S string has fir at wanted

SCALE	CRITERIA
4 (cont.)	<p><u>OR</u></p> <p>E AND S CONTAIN IDENTICAL LETTERS, BUT IN DIFFERING SEQUENTIAL POSITION</p> <p>e.g. E there saw who 'ferice' no S three was how fierce on</p>
3	<p>E DIFFERS FROM S BY <u>TWO</u> LETTERS, ALL REMAINING LETTERS BEING IN THE SAME SEQUENTIAL POSITION</p> <p>e.g. E bright open the was near S brightly over that had nearer</p> <p><u>OR</u></p> <p>E DIFFERS FROM S BY <u>ONE</u> LETTER, WITH ANY OF THE REMAINING LETTERS IN DIFFERING SEQUENTIAL POSITION</p> <p>e.g. E from what tell had wild S for want let and wide</p>
2	<p>E DIFFERS FROM S BY TWO LETTERS, WITH ANY OF THE REMAINING LETTERS IN DIFFERING SEQUENTIAL POSITION</p> <p>e.g. E far went opened his will S after wanted pointed the while</p> <p><u>OR</u></p> <p>E AND S HAVE FIRST AND SECOND LETTERS IN COMMON</p> <p>e.g. E shot run there always theirs S shouted running through also them</p> <p><u>OR</u></p> <p>E AND S HAVE FIRST AND LAST LETTERS IN COMMON</p> <p>e.g. E dodged said 'treatenly' called S dashed shouted threateningly could</p>

SCALE	CRITERIA
1	E AND S HAVE FIRST LETTER IN COMMON
	e.g. E was but back hoping
	S were because branches hesitation
0	E FITS S ON NONE OF THE ABOVE CRITERIA
	e.g. E reach afraid looked wandered
	S stretched frightened stared owned
	All omissions) Since none of these types of error All additions) bear any identifiable orthographic relationship to the text, they all score 0

Note: Sequential position is, in all cases, taken from the beginning of the words concerned: i.e. forward sequence.

The sum of all orthographic scores was expressed as a percentage of the total possible orthographic score for all errors (i.e., total errors x 5). This represented the average orthographic acceptability of that subjects' errors: the relative tendency of that subject to use available orthographic information in his word identification strategy.

v) Self-Corrections

The number of self-corrections was expressed as a percentage of the total number of errors (corrected and uncorrected). This represented the rate at which errors were being spontaneously and overtly detected and corrected by the subject.

Comprehension

The comprehension question for all six stories, together with examples of answers at the various levels of scoring for each question in each story are fully set out in appendix 3. In the present context, the criteria used for scoring each type of comprehension question will be set out. Examples will be selected from The Donkey at reading age seven level. Since the scoring criteria apply to the same type of question across all stories, this should be sufficiently representative for the purpose of illustrating the scoring rationale.

vi) Literal Comprehension

This was a composite score comprising the sum of scores on a) Detail, b) Cause-effect, c) Main idea, and d) Related ideas. The score was expressed as a percentage of the total possible score (10) and reflected the reader's literal comprehension of those ideas that were explicitly stated in the text.

a) Detail

Scoring of answers to questions requiring the accurate recall of detail from the text was on the following scale.

TABLE 17SCALE AND CRITERIA FOR SCORINGLITERAL COMPREHENSION : DETAIL

SCALE	CRITERIA
2	<p>THE READER'S ANSWER (A) IS SYNONYMOUS WITH THE TEXT-BASED ANSWER (T) : A CONTAINS ALL SUBSTANTIVE ELEMENTS OF T.</p> <p>Question (Q): Where was the donkey standing when the boys first saw him?</p> <p>Text-based answer (T): Under a tree, in a field.</p> <p>e.g. A: 'Next to a tree in the middle of a big field'</p>
1	<p>A CONTAINS AT LEAST ONE SUBSTANTIVE ELEMENT OF T.</p> <p>e.g. A: 'Under a tree'</p>
0	<p>A CONTAINS NONE OF THE SUBSTANTIVE ELEMENTS OF T.</p> <p>e.g. A: 'In a stable'</p>

b) Cause-Effect

Scoring of answers to questions requiring recall of the cause to an effect as specified in the text was as follows:

TABLE 18

SCALE AND CRITERIA FOR SCORING
LITERAL COMPREHENSION: CAUSE-EFFECT

SCALE	CRITERIA
2	<p>A EXPRESSES THE SAME EXPLICIT CAUSE AS IN T.</p> <p>Q: Why did the donkey run away from Roger?</p> <p>A: Because he dashed over towards the donkey. He waved his arms. He cried 'Stop!'</p> <p>e.g. A: 'Because Roger dashed over and shouted at him.'</p>
1	<p>A IS AN INDIRECT OR NON-EXPLICIT (IN TERMS OF T) EXPRESSION OF THE CAUSE.</p> <p>e.g. A: 'Because he didn't want Roger to ride him.'</p>
0	<p>A IS AN ERRONEOUS, IRRELEVANT OR REDUNDANT (CIRCULAR) CAUSAL STATEMENT</p> <p>e.g. A: 'Because he was fast.'</p>

c) Main Idea

Scoring of answers to questions requiring a summary of the main idea or central series of events in the story was as follows:

TABLE 19

SCALE AND CRITERIA FOR SCORING
LITERAL COMPREHENSION: MAIN IDEA

SCALE	CRITERIA
3	<p>A EXPRESSES T THROUGH REFERENCE TO AT LEAST <u>THREE</u> OF THE SAME IDEAS</p> <p>Q: How did Roger manage to get onto the donkey to ride him?</p> <p>T: Johnny held out an apple. Roger climbed up the tree (<u>or</u>, onto the branch). The donkey came nearer (<u>or</u>, took the apple from Johnny's hand). Roger slipped down onto the donkey's back.</p> <p>e.g. A: 'He climbed a tree and then jumped down when the donkey came to get the apple.'</p>
2	<p>A MAKES REFERENCE TO <u>TWO</u> OF THE IDEAS IN T.</p> <p>e.g. A: 'He climbed onto a branch and jumped onto the donkey's back.'</p>
1	<p>A MAKES REFERENCE TO <u>ONE</u> OF THE IDEAS IN T.</p> <p>e.g. A: 'He climbed a tree.'</p>
0	<p>A MAKES REFERENCE TO NONE OF THE IDEAS IN T OR IS NON-EXPLICIT.</p> <p>e.g. A: 'He jumped on.'</p>

d) Related Ideas

Scoring of answers to questions requiring the grouping of related ideas or events, stated but not necessarily or expressly grouped in the text, was as follows:

TABLE 20

SCALE AND CRITERIA FOR SCORING
LITERAL COMPREHENSION: RELATED IDEAS

SCALE	CRITERIA
3	A MAKES REFERENCE TO AT LEAST <u>THREE</u> OF THE IDEAS IN T.
	Q: What happened in the first part of the story?
	T: Roger and Johnny had climbed a hill. Johnny was flying his kite. The wind blew the kite higher. Johnny had to run after it. He asked Roger to help pull it in. They pulled it in together.
	e.g. A: 'They were flying their kite and the wind blew it so they had to run after it.'
2	A MAKES REFERENCE TO <u>TWO</u> RELEVANT IDEAS
	e.g. A: 'They were flying a kite and Roger helped Johnny but they nearly fell off the hill.'

SCALE	CRITERIA
1	A MAKES REFERENCE TO <u>ONE</u> RELEVANT IDEA. e.g. A: 'They had a kite and they were playing around with it.'
0	A IS IRRELEVANT. e.g. A: 'They were going for a picnic and Johnny brought an apple along.'

vii) Inferential Comprehension

Scoring for answers to questions requiring the drawing and justification of a character inference, beyond the stated information in the text, was as follows:

TABLE 21

SCALE AND CRITERIA FOR SCORING

INFERENCEAL COMPREHENSION

SCALE	CRITERIA
3	A IS A FULLY EXPANDED AND JUSTIFIED EXPRESSION OF THE INFERENCE AS DRAWN IN T. Q: Who do you think is the cleverest of the two boys? Why?

SCALE	CRITERIA
3 (cont.)	<p>T: Johnny. Because it was he who thought of the idea of holding out the apple so that the donkey would come nearer and Roger could jump from the tree onto his back.</p> <p>e.g. A: 'Johnny. Because he thought of how to get the donkey near with an apple so Roger could jump on him.'</p>
2	<p>A IS A PARTIAL EXPRESSION OF THE INFERENCE AS DRAWN IN T: A CONTAINS A RELEVANT JUSTIFICATION BUT THIS IS NOT FULLY EXPANDED.</p> <p>e.g. A: 'Johnny. Because he thought of the idea.'</p> <p><u>OR</u></p> <p>A IS A FULLY EXPANDED AND JUSTIFIED EXPRESSION OF A LEGITIMATE INFERENCE THAT IS <u>NOT</u> DRAWN IN T.</p> <p>e.g. A: 'Johnny. Because he waited to see what would happen to Roger before <u>he</u> tried to ride the donkey'</p>
1	<p>A IS A PARTIAL EXPRESSION OF THE INFERENCE AS DRAWN IN T BUT IS NEITHER EXPANDED NOR JUSTIFIED</p> <p>e.g. A: 'Johnny'</p>
0	<p>A IS AN ERRONEOUS INFERENCE OR IS IRRELEVANT.</p> <p>e.g. A: 'Roger'</p>

SCALE	CRITERIA
0 (cont.)	A: 'Johnny. Because he knew how to fly a kite.'

The score for inferential comprehension was expressed as a percentage of the total possible score (3) and reflected the degree to which the reader could infer and justify a relevant conclusion that had not been specifically stated in the text.

6. Reliability

Three areas of scoring required to be checked for reliability:

a) The first involved those variables that were essentially dependent on the accuracy of transcription; word identification accuracy (number of uncorrected errors) and self-correction (number of spontaneous self-corrections).

b) The second involved the three error analysis scales and their application; Semantic Acceptability; Syntactic Acceptability and Orthographic Acceptability.

c) The third involved the comprehension variables where both the accuracy of transcription and the application of the relevant scoring scales was at stake.

One sixth of the total sample was randomly selected for the reliability check. In terms of the design, the 20 subjects thus selected were distributed as follows:

TABLE 22

DISTRIBUTION WITHIN DESIGN CELLS OF
RANDOMLY SELECTED RELIABILITY SAMPLE

		P1	P0	
7	H	2	3	
	L	2	2	
9	H	2	2	
	L	2	5	TOTAL 20

Given the nature of the variables, particularly the scoring scales involved, the most relevant form of reliability check was judged to be the relationship between two independent (inter-judge) ratings on each variable. The independent judge, a lecturer in psychology with some knowledge of the concepts involved, was trained by the experimenter on the application of the relevant scales.

a) Oral reading transcription

An independent transcription for each of the 20 subjects in the reliability sample was made from the original tape-recorded data. Raw (untransformed) scores for number of errors (word identification accuracy) and number of self-corrections were then compared across the two ratings. The reliability co-efficient for number of errors was $r = .986$ and for number of self-corrections, $r = .960$. Given that errors in transcription, in most cases, occur only

through lapses of concentration, these co-efficients are, as expected, high. The fact of their being high, nevertheless, endorses the confidence with which results for these two variables can be interpreted.

b) Error analysis scales

A choice was available for this group of variables; to base the independent error analysis on either the original transcription or on the independent transcription. The former would give the reliability of applying the scoring scales by two judges on the same set of transcribed data: the latter would give the reliability of applying the scoring scales by two judges on independently transcribed data. Since it was both more rigorous and more representative of a totally independent assessment of a subject's errors, the latter procedure was adopted. Based on percentage error acceptabilities, the following reliability co-efficients were obtained for the two independent assessments over the three respective error analyses:

Semantic acceptability of errors : $r = ,955$

Syntactic acceptability of errors : $r = ,893$

Orthographic acceptability of errors: $r = ,957$

Given the complexity of the scoring procedure, these reliability co-efficients are highly satisfactory. It can be claimed with confidence that the error analysis scales as developed in this study are satisfactorily reliable instruments for assessing the semantic, the syntactic and the orthographic acceptability of errors.

It is of interest to note that all the co-efficients quoted above are, in fact, similar to those obtained by Hood (1976) after extensive training of five judges. For number of errors and number of self-corrections the co-efficients quoted were between ,97 and ,99. For orthographic acceptability based on a simpler, discrete criterion, the co-efficients were also between ,97 and ,99. For contextual acceptability (combined semantic and syntactic acceptability with pre-error context, sentence context and passage context as criteria) the co-efficients were between ,84 and ,94. Although Hood was evaluating reliability over five judges, her co-efficients - on much simpler criteria than those used in this study - do indicate that the present scales, despite their complexity, are sufficiently definitive to allow a similar level of reliability.

c) Comprehension

As with the error analysis scales, the more rigorous option of testing the comprehension scales on the basis of the independent transcription was adopted. For scores on answers to the five different comprehension questions, the following reliability co-efficients were obtained;

- | | | |
|----|----------------|----------|
| 1. | Detail: | r = ,861 |
| 2. | Cause-Effect: | r = ,866 |
| 3. | Main Idea: | r = ,903 |
| 4. | Related Ideas: | r = ,899 |
| 5. | Inference: | r = ,971 |

The co-efficient for the sum of scores on answers to

questions 1 to 4, Literal comprehension, was $r = ,941$. Again, these co-efficients are all satisfactorily high. The relatively high reliability of scoring inferential comprehension was somewhat surprising given its relative complexity in comparison with other comprehension criteria. However, many subjects, particularly at the reading age seven level, scored either 0 or only 1 on this variable and since the criteria for these two scores are clearly definitive (see comprehension scales above) the relatively high reliability is probably, in part, accounted for in this. Despite the high level of variance on this variable (see Results, chapter X), and the reservations (expressed in chapter XII) on the reliability of a measure of comprehension based on one question, it is at least clear that scoring of the variable was highly reliable. This does add some measure of confidence to the results.

CHAPTER XRESULTS AND INTERPRETATION

The results will be presented in two sections. The first and major section will cover the presentation and interpretation of comparative data within the experimental design. Cell and treatment means together with the relevant analysis of variance data will be presented and interpreted for each of the dependent variables in the following order:

- i) Word Identification Accuracy
- ii) Semantic Acceptability of Errors
- iii) Syntactic Acceptability of Errors
- iv) Orthographic Acceptability of Errors
- v) Self-Corrections
- vi) Literal Comprehension
- vii) Inferential Comprehension

The main purpose will be to identify significant illustration (P) effects across the four experimental groups, RA7, High Progress (7H); RA9, High Progress (9H); RA7, Low Progress (7L); and RA9, Low Progress (9L). Although not specifically hypothesised, it is expected that the relative strength of the P effect will differ across the four experimental groups (Chapter VII). A second and important purpose, therefore, will be to identify significant progress level (Pr) and reading age level (RA) effects particularly where these interact with illustration (P) effects. Where

Pr and RA effects are independent of any P effect they will be interpreted in so far as they are of general interest although not of direct relevance to the experimental issue. Furthermore, since the relative strength of P effects is not always apparent in the absolute terms of conventional interpretation, trends, where necessary, will also be interpreted particularly where these are consistent with theoretical expectations or are consistent in pattern over theoretically related dependent variables.

The second section will involve the presentation and interpretation of a correlation matrix over all dependent variables (all subjects, all conditions). This is introduced in so far as it might elucidate inter-relationships among the dependent variables and the significance of these for the theoretical and practical interpretation of the overall results.

For all results, levels of significance will be indicated as follows (critical F values for analysis of variance results with df 1/112 are also indicated):

+++ = $p < .001$ (F = 11,38)

++ = $p < .01$ (F = 6,85)

+ = $p < .05$ (F = 3,92)

Throughout, the illustration condition is represented as P1 and the no-illustration condition as P0; the seven year reading age level is represented as 7 and the nine year reading age level as 9; high progress readers are represented as H and low progress readers as L.

1. Analyses of Variancei. Word Identification Accuracy

Scores on this variable are represented in the following formula:

$$\text{Word identification accuracy} = 100 \left(\frac{a-b}{a} \right)$$

where: a = Total number of words in the text

b = Number of uncorrected errors

TABLE 23a

CELL MEANS/STANDARD DEVIATIONS,
CELL MEAN DIFFERENCES (P1 - P0) AND
TREATMENT MEANS: WORD IDENTIFICATION ACCURACY

		P1	P0	DIFF
7	H	95,91	89,30	6,61
		(2,55)	(8,05)	
	L	94,47	91,38	3,09
		(2,72)	(5,30)	
9	H	96,08	95,60	0,48
		(1,89)	(2,66)	
	L	94,24	91,35	2,89
		(3,20)	(4,26)	

	P1	P0
H	96,00	92,45
L	94,36	91,37

	P1	P0
7	95,19	90,34
9	95,16	93,48

	H	L
7	92,61	92,93
9	95,84	92,80

P1	95,18	H	94,22	7	92,77
P0	91,91	L	92,86	9	94,32

FIGURE 5

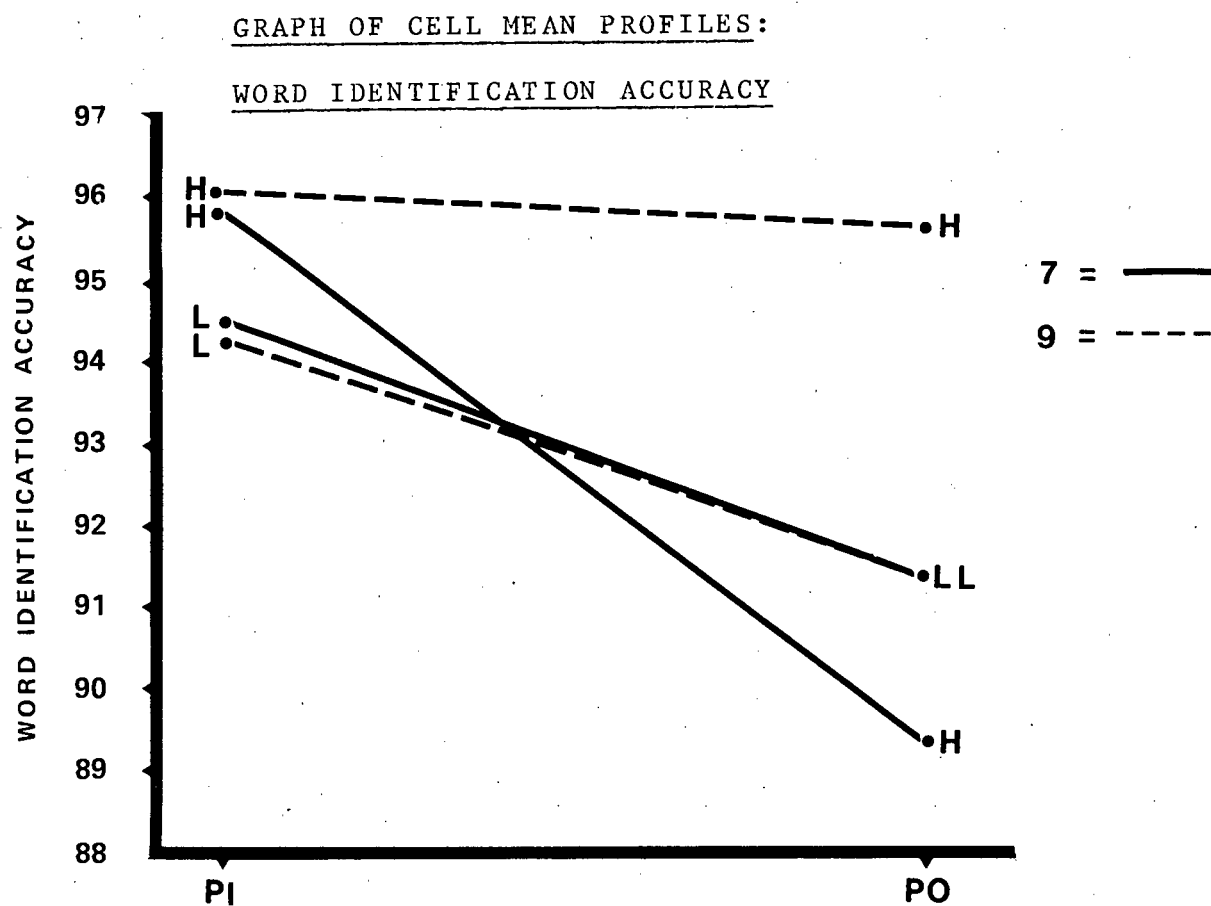


TABLE 23b

ANALYSIS OF VARIANCE SUMMARY TABLE:WORD IDENTIFICATION ACCURACY

SOURCE	SS	df	MS	F RATIO
P	319,61	1	319,61	17,54+++
Pr	56,00	1	56,00	3,07
RA	72,11	1	72,11	3,96+
PxPr	2,30	1	2,30	0,13
PxRA	75,23	1	75,23	4,13+
PrxRA	84,87	1	84,87	4,66+
PxPrxRA	66,09	1	66,09	3,63 ¹
WITHIN	2040,67	112	18,22	

¹ P < ,06

Given the pattern of interaction on other dependent variables; given that group differences at P1 on this particular variable are, a priori, limited by the pre-testing of illustrated stories to be readable at approximately 95% accuracy for all experimental groups (chapter IX); and given that the PxPrxRA interaction is nevertheless very close to the stipulated 5% level significance ($F = 3,63$; $p < ,06$: Table 23b), it seems most logical to analyse the results for this variable on the basis of this interaction.

TABLE 23c

SIMPLE INTERACTION EFFECTS:
WORD IDENTIFICATION ACCURACY

SOURCE	SS	df	MS	F RATIO
PxPr at 7	46,61	1	46,61	2,56
at 9	21,80	1	21,80	1,20
PxRA at H	141,21	1	141,21	7,75++
L	0,13	1	0,13	0,01
PrxRA at P1	0,58	1	0,58	0,03
P0	150,41	1	150,41	8,26++
WITHIN	2040,67	112	18,22	

On the basis of the significant simple interactions P x RA at H and Pr x RA at PO, the following simple-simple main effects are relevant:

TABLE 23d

SIMPLE-SIMPLE MAIN EFFECTS:
WORD IDENTIFICATION ACCURACY

SOURCE	SS	df	MS	F RATIO
P at 7,H	327,76	1	327,76	17,99+++
9,H	1,68	1	1,68	0,09
RA at H,P1	0,20	1	0,20	0,01
H,PO	297,74	1	297,74	16,34+++
Pr at 7,PO	32,43	1	32,43	1,78
9,PO	135,73	1	135,73	7,45
(RA at H,PO	297,74	1	197,74	16,34+++) ¹
L,PO	0,01	1	0,01	0,00
Within	2040,67	112	18,22	

On the basis of the non-significant simple interactions P x Pr at 7; P x Pr at 9; P x RA at L and Pr x RA at P1

¹ Effects that are redundant in the table as a whole will be bracketted but will be included to preserve the deductive pattern from simple interaction to relevant simple-simple main effects.

(Table 23c), the following simple main effects are relevant:

TABLE 23e

SIMPLE MAIN EFFECTS: WORD IDENTIFICATION ACCURACY

SOURCE	SS	df	MS	F RATIO
P at 7	352,51	1	352,51	19,35+++
Pr at 7	1,51	1	1,51	0,08
P at 9	42,35	1	42,35	2,32
Pr at 9	139,37	1	139,37	7,65++
P at L	133,70	1	133,70	7,34++
RA at L	0,27	1	0,27	0,02
Pr at P1	40,56	1	40,56	2,23
RA at P1	0,02	1	0,02	0,00
WITHIN	2040,67	112	18,22	

Interpretation

Illustration (P) Effects:

Given the P x Pr x RA interaction ($F = 3,63$; $p < ,06$: Table 23b), analysis of the simple interaction effects (Table 23c) reveals a significant P x RA interaction at H ($F = 7,75++$). The simple-simple main effect for P at 7H ($F = 17,99+++$) is highly significant while P at 9H ($F = 0,09$) is non-significant (Table 23d).

Going from the non-significant simple interactions (P x Pr at 7, $F = 2,56$; P x Pr at 9, $F = 1,20$; P x RA at L, $F = 0,01$: Table 23c), the simple main effects for P both at 7 and at L are significant ($F = 19,35+++$; $F = 7,34++$ respectively) while P at 9 is non-significant ($F = 2,32$; Table 23c).

It appears, therefore, that there is a strong and significant P effect for the 7H group and a clearly non-significant P effect for the 9H group. On the other hand the P effect is significant for the combined groups at L (7L and 9L) and at 7 (7L and 7H) but non-significant at 9 (9H and 9L). Given the strong P effect for 7H and the clearly non-significant P effect for 9H it would appear that the P effects for 7L and 9L are only moderate (i.e., the strong effect at 7 is heavily influenced by the 7H effect and the non-significant effect at 9 is heavily influenced by the 9H effect). This pattern is consistent with the cell mean differences (Table 23a) and the relative P effects as shown in Figure 5.

On these grounds the results for the P effect may be taken as evidence that reading with or without illustrations has effectively no influence on the word identification accuracy of high progress readers at reading age nine. On the other hand, the influence is strong for high progress readers at reading age seven and moderate for low progress readers at the reading ages of both seven and nine.

Progress level (Pr) effects:

The simple interaction effects (Table 23c) reveal a significant Pr x RA interaction at P0 ($F = 8,26++$). The simple-simple main effect for Pr at 9P0 is significant ($F = 7,45++$) while Pr at 7P0 is non-significant ($F = 1,78$: Table 23d).

Going from the non-significant simple interactions (P x Pr at 7, $F = 2,56$; P x Pr at p, $F = 1,20$ and Pr x RA at P1, $F = 0,03$; Table 23c), the simple main effect for Pr at 9 is significant ($F = 7,65++$) while Pr both at 7 and at P1 is non-significant ($F = 0,08$; $F = 2,23$: Table 23e).

It appears, therefore, that the 7L and 7H groups do not differ significantly overall in their word identification accuracy. Although there appears to be a significant difference overall between the 9H and 9L groups the combined groups at P1 (7H and 9H vs 7L and 9L) do not show a significant difference while at P0 there is a clearly significant difference between 9H and 9L. This apparent anomaly becomes clarified in considering the RA effects as well.

Reading Age (RA) effects:

The simple interaction effects (Table 23e) reveal a significant P x RA interaction at H ($F = 7,75++$) and a significant Pr x RA interaction at P0 ($F = 8,26++$). The simple-simple main effect for RA at HPO ($F = 16,34++$) is significant while RA at HPI and RA at LPO are both non-significant ($F = 0,01$; $F = 0,00$ respectively: Table 23d).

Going from the non-significant simple interactions ($P \times RA$ at L, $F = 0,01$ and $Pr \times RA$ at P1, $F = 0,03$: Table 23c), the simple main effects for RA both at L and at P1 are non-significant ($F = 0,02$; $F = 0,00$ respectively: Table 23e).

Thus, the 7L and 9L groups do not differ significantly overall in their word identification accuracy. Neither do the combined groups at P1 (7H and 7L vs 9H and 9L). At P0, the 7H group differs significantly from the 9H group but the 7L and 9L groups do not differ.

Considered together with the Pr effects, the evidence indicates that there are no differences between groups at P1. At P0, on the other hand, there are no differences between 7L and 9L or between 7L and 7H (or between 7H and 9L by implication) but a significant difference does exist between 9H and 9L and between 9H and 7H (and, by implication, between 9H and 7L). In other words, simply stated, where illustrations are available, word identification accuracy is similar for all groups: where illustrations are not available, word identification accuracy is significantly lower for all groups compared to the high progress readers at reading age nine. Two observations need to be made here.

First, although the difference between 7H and 7L at P0 is non-significant, the trend for high progress readers to have a lower level of word identification accuracy than low progress readers under this condition is counter to what might be expected. Since the design necessitated the

selection of high progress readers from a lower grade than low progress readers (chapter VIII), it is likely that a grade level effect may account for the trend. In other words, 7H (grade I) readers tend to be more susceptible to the removal of illustrations in word identification than 7L (grade III) readers. The tendency, on this variable alone, is not distinguishable from a chance effect ($F = 1,78$; $p < ,20$: Table 23d). It is mentioned, however, because it appears on subsequent variables and may therefore constitute a genuine effect.

Second, given that stories had been constructed and pre-tested to be readable at an 'instructional' level of accuracy, with illustrations, for all experimental groups (chapter IX) the lack of differences under the illustration condition is entirely expected. Thus, the general pattern of results suggests that under the no-illustration condition the high progress readers at reading age nine are as accurate in their word identification as they are under the illustration condition. On the other hand, the low progress readers at reading ages seven and nine and, particularly, the high progress readers at reading age seven achieve a significantly lower level of word identification accuracy where the contextual information in illustrations is not available: for these groups it appears that word identification without the support of the illustrations is a less accurate process than it is for the high progress readers at reading age nine, and than it is for these same groups where the illustration information is available.

Based on a similar pattern that appears on other variables (see below), it seems likely that the role of illustrations is to increase the accessibility of semantic information, in particular. Thus, on word identification accuracy, when illustrations are not available, the high progress readers at reading age nine are relatively unaffected since semantic information is accessible to them through the text itself. By contrast, high progress readers at reading age seven (grade I) appear to be strongly dependent on illustrations and to find word identification difficult where the relevant semantic information is not made accessible through the extra-textual context of illustrations. Low progress readers at both reading age levels appear to experience an intermediate effect between these two extremes.

ii. Semantic Acceptability of Errors

Scores on this variable are represented in the following formula:

$$\text{Semantic acceptability} = 100 \left(\frac{\sum \text{SEM}}{4d} \right)$$

where: SEM = Score (0-4) for each error¹ on the semantic acceptability scale

d = Total number of errors¹

¹ 'Errors' here include those that were subsequently self-corrected.

TABLE 24a

CELL MEANS/STANDARD DEVIATIONS, CELL MEAN
DIFFERENCES (P1-PO), AND TREATMENT MEANS:
SEMANTIC ACCEPTABILITY

	P1	PO	DIFF
7	H 47,72	28,01	19,71
	(11,40)	(13,00)	
	L 48,88	43,00	5,88
	(11,70)	(8,62)	
9	H 47,02	40,47	6,55
	(10,17)	(7,69)	
	L 56,39	44,07	12,32
	(10,85)	(11,59)	

	P1	PO
H	47,37	34,24
L	52,64	43,54
	P1	PO
7	42,27	32,51
9	51,71	48,30

	H	L		
7	37,86	45,94		
9	43,75	50,23		
PI	50,00	H 40,81	7	41,90
PO	38,89	L 48,09	9	46,99

FIGURE 6

GRAPH OF CELL MEAN PROFILES:

SEMANTIC ACCEPTABILITY

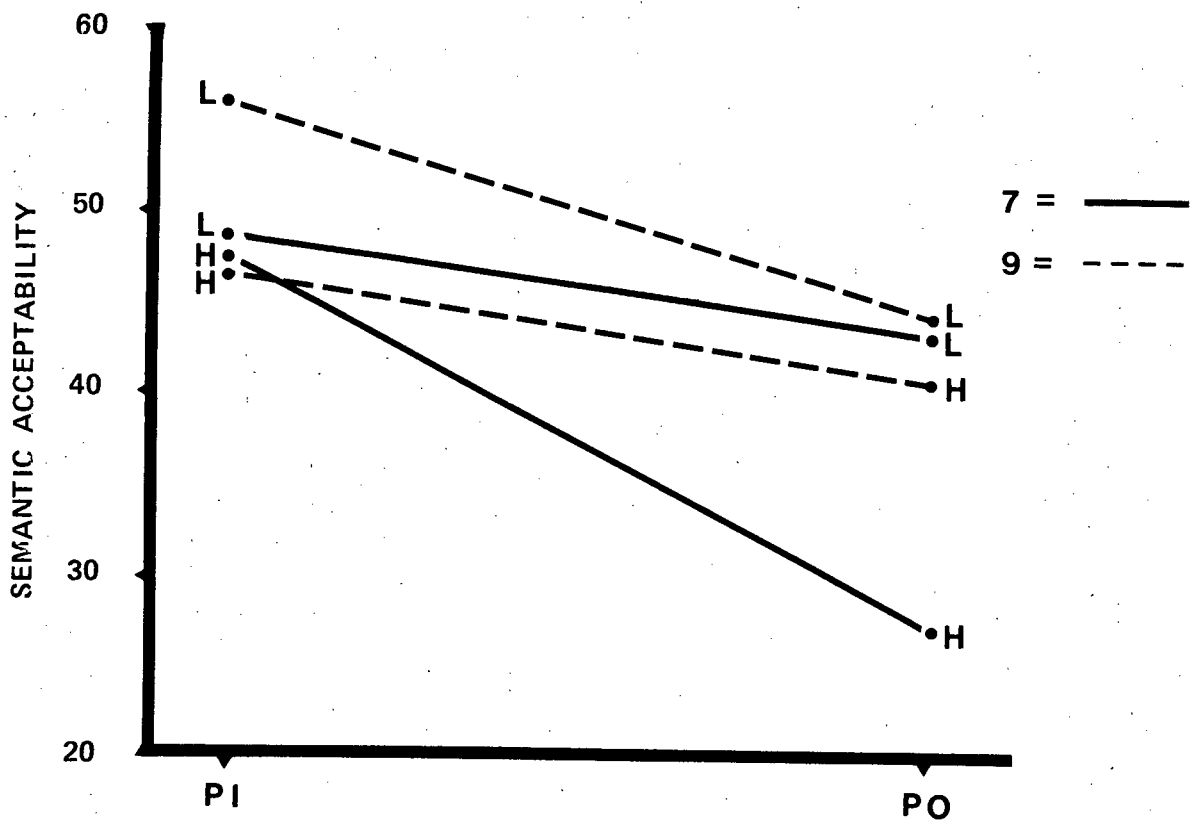


TABLE 24b

ANALYSIS OF VARIANCE SUMMARY TABLE:SEMANTIC ACCEPTABILITY

SOURCE	SS	df	MS	F RATIO
P	3705,52	1	3705,52	32,06+++
Pr	1589,88	1	1589,88	13,76+++
RA	776,48	1	776,48	6,72+
PxPr	121,74	1	121,74	1,05
PxRA	84,55	1	84,55	0,73
PrxRA	19,05	1	19,05	0,17
PxPrxRA	720,54	1	720,54	6,24+
WITHIN	12943,28	112	115,57	

Given the significant P x Pr x RA interaction, analysis of the simple interaction effects is relevant.

TABLE 24c

SIMPLE INTERACTION EFFECTS:
SEMANTIC ACCEPTABILITY

SOURCE	SS	df	MS	F RATIO
PxPr at 7	717,33	1	717,33	6,21+
at 9	124,96	1	124,96	1,08
PxRA at H	649,38	1	649,38	5,62+
at L	155,72	1	155,72	1,35
PrxRA at P1	252,64	1	252,64	2,19
at P0	486,95	1	486,95	4,21+
WITHIN	12943,28	112	115,57	

On the basis of the significant simple interactions P x Pr at 7, P x RA at H and Pr x RA at P0, the following simple-simple main effects are relevant:

TABLE 24d

SIMPLE-SIMPLE MAIN EFFECTS:SEMANTIC ACCEPTABILITY

SOURCE	SS	df	MS	F RATIO
P at 7,H	2913,04	1	2913,04	25,21+++
7,L	259,07	1	259,07	2,24
Pr at 7,P1	10,11	1	10,11	0,09
7,PO	1685,70	1	1685,70	14,59+++
(P at 7,H	2913,04	1	2913,04	25,21+++)
9,H	321,64	1	321,64	2,78
RA at H,P1	3,63	1	3,63	0,03
H,PO	1165,14	1	1165,14	10,08++
(Pr at 7,PO	1685,70	1	1685,70	14,59+++)
9,PO	97,02	1	97,02	0,84
(RA at H,PO	1165,14	1	1165,14	10,08++)
L,PO	8,57	1	8,57	0,07
WITHIN	12943,28	112	115,57	

On the basis of the non-significant simple interactions P x Pr at 9, P x RA at L and Pr x RA at P1 (Table 24c), the following simple main effects are relevant:

TABLE 24e

SIMPLE MAIN EFFECTS: SEMANTIC ACCEPTABILITY

SOURCE	SS	df	MS	F RATIO
P at 9	1335,29	1	1335,29	11,55+++
Pr at 9	630,44	1	630,44	5,46+
P at L	1241,97	1	1241,97	10,75++
RA at L	276,15	1	276,15	2,39
Pr at P1	415,86	1	415,86	3,60
RA at P1	174,29	1	174,29	1,51
WITHIN	12943,28	112	115,57	

InterpretationIllustration (P) effects:

Given the significant P x Pr x RA interaction (F = 6,24+: Table 24b), analysis of the simple interaction effects (Table 24c) reveals a significant P x Pr interaction at 7 (F = 6,21+) and a significant P x RA interaction at H (F = 5,62+). The simple-simple main effect for P at 7H is highly significant (F = 25,21+++), while for P at 7L and for P at 9H the effects are non-significant (F = 2,24; F = 2,78 respectively: Table 24d).

Going from the non-significant simple interactions (P x Pr at 9, $F = 1,08$; P x RA at L, $F = 1,35$: Table 24c), the simple main effects for P both at 9 and at L are significant ($F = 11,55+++$; $F = 10,75++$ respectively: Table 24e).

What emerges from this is a strong and significant P effect for the 7 H group and lesser but also highly significant P effects for the combined groups at 9 (9H and 9L) as well as at L (7L and 9L). Given the relative strength of the P effect at 9L that is apparent in the cell mean differences (Table 24a) and in Figure 6 it would appear that the strength of the P effect both at 9 and at L is heavily influenced by the effect at 9L. In other words, it is likely that the P effect is strong at 9L while only moderate at both 9H and 7L.

On this interpretation, the results for the P effect may be taken as evidence that illustrations strongly increase the semantic acceptability of errors for high progress readers at reading age seven as well as for low progress readers at reading age nine. For high progress readers at reading age nine and for low progress readers at reading age seven illustrations also increase the semantic acceptability of errors but the increase is only moderate.

Progress level (Pr) Effects

The simple interaction effects (Table 24c) reveal a significant P x Pr interaction at 7 ($F = 6,21+$) and a significant Pr x RA interaction at P0 ($F = 4,21+$). Only the

group where there are no illustrations.

Interpreting the Pr and RA results together, it can be said that the semantic acceptability of errors for the 7H group is significantly lower than for the 7L, the 9H and (by implication) the 9L groups where there are no illustrations.

No significant differences between groups are apparent where illustrations are available. Overall, the semantic acceptability of errors is significantly higher for the 9L group than for the 9H group.

Apart from supporting the interpretation of the relative strength of the P effect for the 7H group, and adding some reservation on the relative strength of the P effect for the 9L group, when these are compared with the 7L and 9H groups, the Pr and RA results are particularly interesting in terms of the grade level effect suggested under the results for word identification accuracy. For instance, if the P effect is temporarily suspended, the mean semantic acceptability for the 7H (grade I) group is 37,86; for the 7L and 9H (grade III) groups it is 45,94 and 43,75 respectively; and for the 9L (grade V) group it is 50,23 (Table 24a). As a tendency, this is not unexpected in terms of, particularly, Burke's (1976) finding that semantic acceptability of errors increases with grade level.

The influence of the P effect on this basic pattern is to decrease the semantic acceptability of errors of the 7H (grade I) and 9L (grade V) groups relative to the 7L and 9H (grade III) groups under the no-illustration condition;

and to increase the semantic acceptability of errors for all groups under the illustration condition but particularly for the 7H and 9L groups relative to the 7L and 9H groups.

Thus the general pattern of the results suggests that a grade level effect operates for the semantic acceptability of errors but that this is modified by the relative accessibility of semantic information as determined by the informational value of illustrations to the various groups. In other words, with no illustrations, semantic information appears to be far less accessible to high progress readers at reading age seven (grade I) than it is to the older readers who appear to make relatively more use of text-based semantic information under this condition. For all readers the addition of illustrations to the text improves the accessibility of semantic information, but this effect is particularly strong for the high progress readers at reading age seven and the low progress readers at reading age nine. For the former group, the addition of illustrations appears to make a difference from near meaningless reading (word calling) to relatively meaningful reading (mean semantic acceptability from 28,01 to 47,72): for the latter group it appears to make a difference from relatively meaningless reading to substantially meaningful reading (mean semantic acceptability from 44,07 to 56,39).

iii. Syntactic Acceptability of Errors

Scores on this variable are represented in the following formula:

$$\text{Syntactic Acceptability} = 100 \left(\frac{\sum \text{SYN}}{4d} \right)$$

where: SYN = Score (0-4) for each error on the syntactic acceptability scale

d = total number of errors.

TABLE 25a

CELL MEANS/STANDARD DEVIATIONS, CELL MEAN DIFFERENCES (P1-P0), AND TREATMENT MEANS:

SYNTACTIC ACCEPTABILITY

		P1	P0	DIFF
7	H	66,92	54,46	12,46
		(14,05)	(20,04)	
	L	70,38	62,40	7,98
		(7,14)	(10,88)	
9	H	77,71	72,31	5,40
		(13,90)	(11,38)	
	L	77,15	71,21	5,94
		(6,63)	(12,42)	

	P1	P0
H	72,32	63,38
L	73,76	66,80

	P1	P0
7	68,65	58,43
9	77,43	71,76

	H	L
7	60,69	66,39
9	75,01	74,18

P1	73,04	H	67,85	7	63,54
P0	65,09	L	70,28	9	74,60

FIGURE 7

GRAPH OF CELL MEAN PROFILES:

SYNTACTIC ACCEPTABILITY

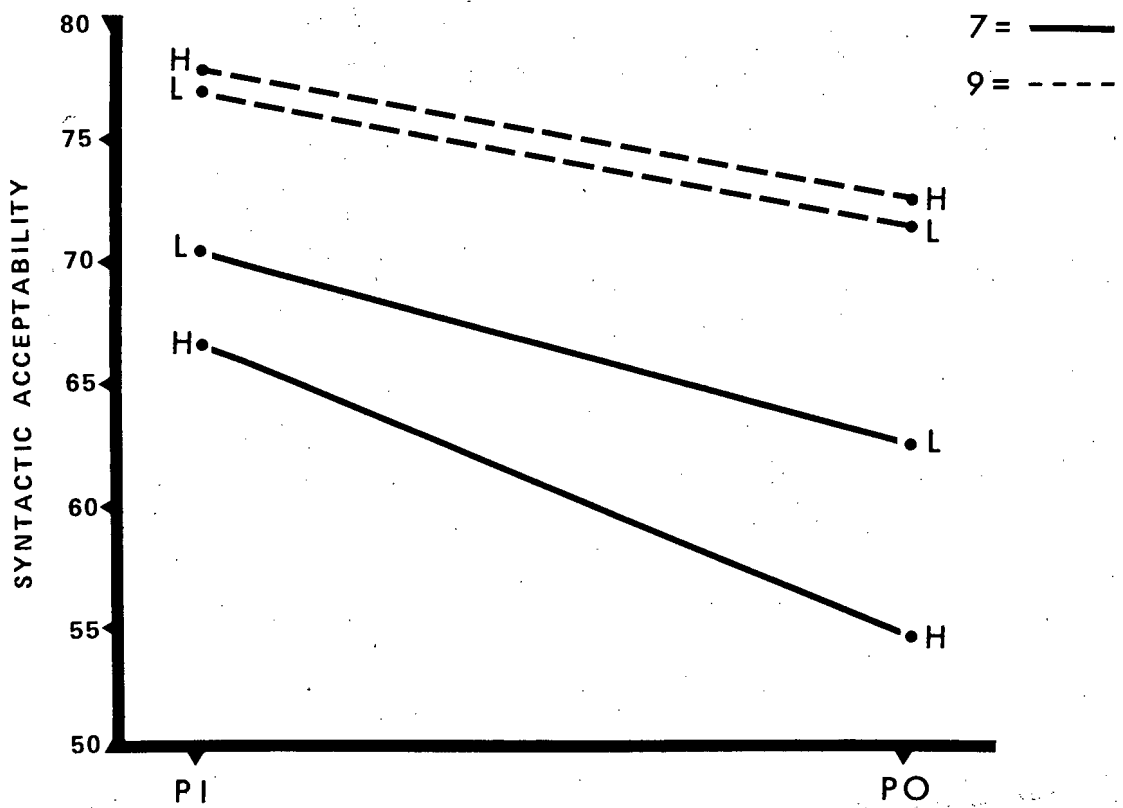


TABLE 25b

ANALYSIS OF VARIANCE SUMMARY TABLE:SYNTACTIC ACCEPTABILITY

SOURCE	SS	df	MS	F RATIO
P	1894,80	1	1894,80	11,76++
Pr	177,68	1	177,68	1,10
RA	3666,61	1	3666,61	22,75+++
PxPr	29,28	1	29,28	0,18
PxRA	155,41	1	155,41	0,96
PrxRA	319,42	1	319,42	1,98
PxPrxRA	47,18	1	47,18	0,29
WITHIN	18049,31	112	161,16	

InterpretationIllustration (P) effect

With no significant interactions, the significant P main effect may be taken as generally applicable across experimental groups. It appears that illustrations significantly increase the syntactic acceptability of errors irrespective of progress or reading age levels.

It is interesting to note, however, that the error term is somewhat higher than that for semantic acceptability and while the cell mean differences for the 7L and 9H groups are similar, those for the 7H and 9L groups are

considerably less substantial than for semantic acceptability.

TABLE 26

COMPARISON OF CELL MEAN DIFFERENCES (P1-P0)
AND ERROR TERMS ACROSS SEMANTIC AND
SYNTACTIC ACCEPTABILITIES OF ERRORS

		SEMANTIC	SYNTACTIC
7	H	19,71	12,46
	L	5,88	7,98
9	H	6,55	5,40
	L	12,32	5,94
MS, WITHIN		115,57	161,16

Although data from the two scales are, strictly speaking, not directly comparable, there is some logical basis for comparison in so far as both scales score from a maximum of 4 to a minimum of 0 on similar, if not identical, levels of criterion (chapter IX). With the higher error variance for syntactic acceptability it may be concluded that P effect mean differences on this measure are generally of relatively less significance than those on semantic acceptability (semantic P main effect, $F = 32,06+++$; syntactic P main effect, $F = 11,76++$). In particular, where the P effect was found to be strong at 7H and 9L for semantic acceptability and only moderate for 7L and 9H, the P effect

over all groups may be regarded as only moderate for syntactic acceptability.

In terms of the influence of illustrations, therefore, it would appear that there is an undifferentiated and moderate increase in the appearance of syntactic constraints in the errors of all readers where illustrations are available. Since it is unlikely that illustrations have a direct influence on the use of syntactic information it would appear that this effect is linked to the influence of illustrations on the accessibility of semantic information. In other words, where illustrations have a direct effect on the accessibility and use of semantic information, it is likely that errors that are highly semantically acceptable will also tend to be linguistically acceptable in a general sense (i.e., syntactically acceptable as well). Where errors are not semantically acceptable, however, there may nevertheless be a degree of syntactic constraint (Chapter IX). Thus for those groups that appeared to be particularly influenced by the relative non-accessibility of semantic information where no illustrations were available (the 9L and, particularly, the 7H group) it would not necessarily follow that use of syntactic information should be equally reduced under this condition. The undifferentiated illustration effect for syntactic acceptability would appear to be at least partly explained in this. Equally the overall moderate effect may be explained in the partial link that must necessarily exist between semantic acceptability and syntactic acceptability

of errors.

Reading age (RA) effect

The highly significant main effect for reading age on syntactic acceptability of errors ($F = 22,75+++$; Table 25b) is not unexpected. This is consistent with other findings (chapter VII) that children's use of syntactic information increases with development. The trend towards a grade level effect, as in semantic acceptability, is again apparent. In this case, however, the high progress readers at reading age nine (grade III) make use of syntactic constraints to the same high level as low progress readers at reading age nine (grade V). This would appear to be a combination of both grade level and progress level effects and is interesting in the light of Mackworth's (1972b) finding that good readers make relatively more progress in their use of syntactic information than poor readers beyond the fourth grade level.

iv. Orthographic acceptability of Errors

Scores on this variable are represented in the following formula:

$$\text{Orthographic acceptability} = 100 \left(\frac{\sum \text{ORT}}{5d} \right)$$

where: ORT = Score (0-5) for each error on the orthographic acceptability scale

d = number of errors

TABLE 27a

CELL MEANS/STANDARD DEVIATIONS, CELL
MEAN DIFFERENCES (P1-P0), AND TREATMENT
MEANS: ORTHOGRAPHIC ACCEPTABILITY

		P1	P0	DIFF
7	H	44,56	58,89	-14,33
		(5,74)	(11,19)	
	L	45,85	48,87	- 3,02
		(3,39)	(11,03)	
9	H	45,16	48,83	- 3,67
		(11,23)	(9,06)	
	L	43,18	46,10	- 2,92
		(8,04)	(10,55)	

	P1	P0
H	44,86	53,86
L	44,52	47,48

	P1	P0
7	45,21	53,88
9	44,17	47,47

	H	L
7	51,73	47,36
9	47,00	44,64

P1	46,49	H	49,36	7	49,54
P0	50,67	L	46,00	9	45,82

FIGURE 8

GRAPH OF CELL MEAN PROFILES:
ORTHOGRAPHIC ACCEPTABILITY

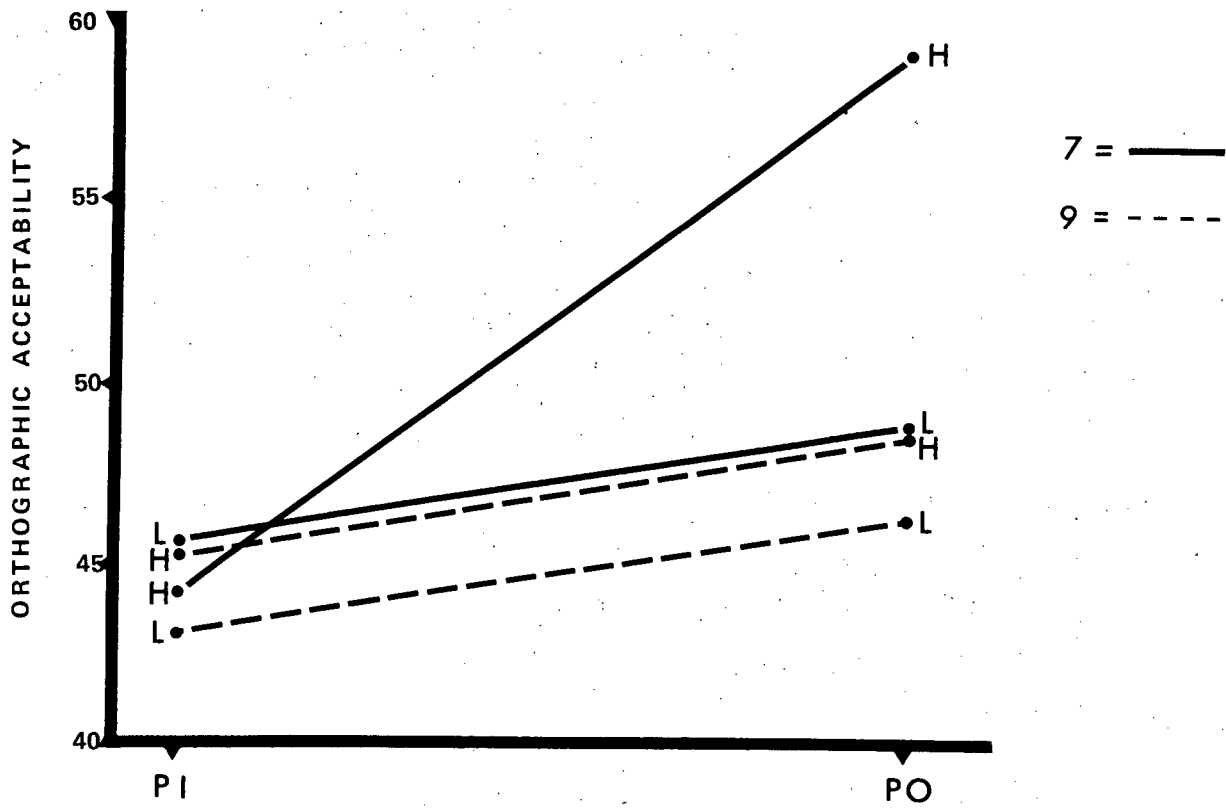


TABLE 27b

ANALYSIS OF VARIANCE SUMMARY TABLE:ORTHOGRAPHIC ACCEPTABILITY

SOURCE	SS	df	MS	F RATIO
P	1072,87	1	1072,87	11,69++
Pr	338,72	1	338,72	3,69
RA	415,86	1	415,86	4,53+
PxPr	272,62	1	272,62	2,97
PxRA	216,79	1	216,79	2,36
PrxRa	30,41	1	30,41	0,33
PxPrxRA	209,16	1	209,16	2,28
WITHIN	10278,40	112	91,77	

InterpretationIllustration (P) Effect

With no significant interactions, the significant P main effect may be taken as generally applicable across the experimental groups. It appears that illustrations significantly reduce the orthographic acceptability of errors irrespective of progress or reading age levels.

Nevertheless, given the interaction pattern on other results and the predicted inverse effect of illustrations on orthographic and semantic acceptabilities of errors, the trend to P x Pr x RA interaction ($F = 2,28$; $p < ,20$: Table 27b)

may be taken as possibly having some significance. The chance probability is too high to warrant formal analysis of the results in terms of this trend. However, examination of the cell mean differences (Table 27a) and the cell mean profiles (Figure 8) suggests that the P effect is similar and barely moderate for the 9H, 7L and 9L groups but that it is quite substantial for the 7H group. Equally, it would appear that the orthographic acceptability of errors at P0 is considerably higher for the 7H group than for the other groups while at P1 there is little difference between all groups. It would appear, therefore, that the P main effect must be substantially influenced by the strong effect at 7H. Although the interaction does not reach significance, the trend would seem to indicate that the P effect is moderate, and probably only marginal, for the 9H, 7L and 9L groups compared to the relatively strong effect at 7H.

On this interpretation, the inverse effect of illustrations on orthographic and semantic acceptabilities of errors appears likely to apply strongly to the high progress readers at reading age seven: illustrations strongly increase the accessibility and use of semantic information and this is matched by what appears to be a relatively strong reduction in the use of, or reliance on, orthographic information. For the other three groups the inverse effect is present but the reduction in use of orthographic information appears to be only marginal.

Reading age (RA) effect

The significant RA main effect ($F = 4,53+$; Table 27b) is also of interest. Within the trend to interaction it is again apparent that the effect must be substantially influenced by the considerably higher orthographic acceptability of the 7H group's errors at P0 when compared to the other three groups (Table 27a; Figure 8). At P0, a grade level trend, in inverse order, is also apparent (mean orthographic acceptability at P0: 7H (grade I), 38,39; 7L, 9H (grade III), 48,87, 48,83; 9L (grade V), 46,10). At P1, however, all differences are marginal and neither RA nor grade level trends are apparent. On the same argument as used in the essentially parallel but inverse pattern for semantic acceptability, this would appear to relate to the relative accessibility of semantic information. In other words, the grade level/RA effect is apparent where there are no illustrations such that semantic information is relatively inaccessible to the younger readers. Under this condition, the younger the reader the more orthographic information needs to be relied on; whereas, the older the reader the more text-based semantic information is accessible and can be used to reduce reliance on orthographic information. Where illustrations are available, however, the effect disappears. Although all readers reduce their use of orthographic information with the available semantic information in illustrations, it is the high progress, grade I, readers who, in particular, benefit from the greater accessibility

of semantic information such that their reliance on orthographic information becomes equivalent to that of the other groups where illustrations are available.

v. Self-Corrections

Scores on this variable are represented in the following

formula: Self-corrections = 100

$$\text{Self-corrections} = 100 \left(\frac{c}{b+c} \right)$$

where: c = number of self-corrected errors

b = number of uncorrected errors

TABLE 28a

CELL MEANS/STANDARD DEVIATIONS, CELL
MEAN DIFFERENCES (P1-P0), AND
TREATMENT MEANS: SELF-CORRECTIONS

	P1	P0	DIFF
7	H 38,37 (9,86)	16,82 (11,53)	21,55
	L 30,88 (11,97)	23,98 (15,07)	6,90
9	H 27,25 (16,34)	25,40 (16,07)	1,85
	L 26,72 (13,82)	12,90 (6,67)	13,82

	P1	P0
H	32,81	21,11
L	28,80	18,44

	P1	P0
7	34,62	20,40
9	26,99	19,15

	H	L
7	27,60	27,43
9	26,33	19,81

P1	30,81	H	26,96	7	27,51
P0	19,78	L	23,62	9	23,07

FIGURE 9

GRAPH OF CELL MEAN PROFILES : SELF-CORRECTIONS

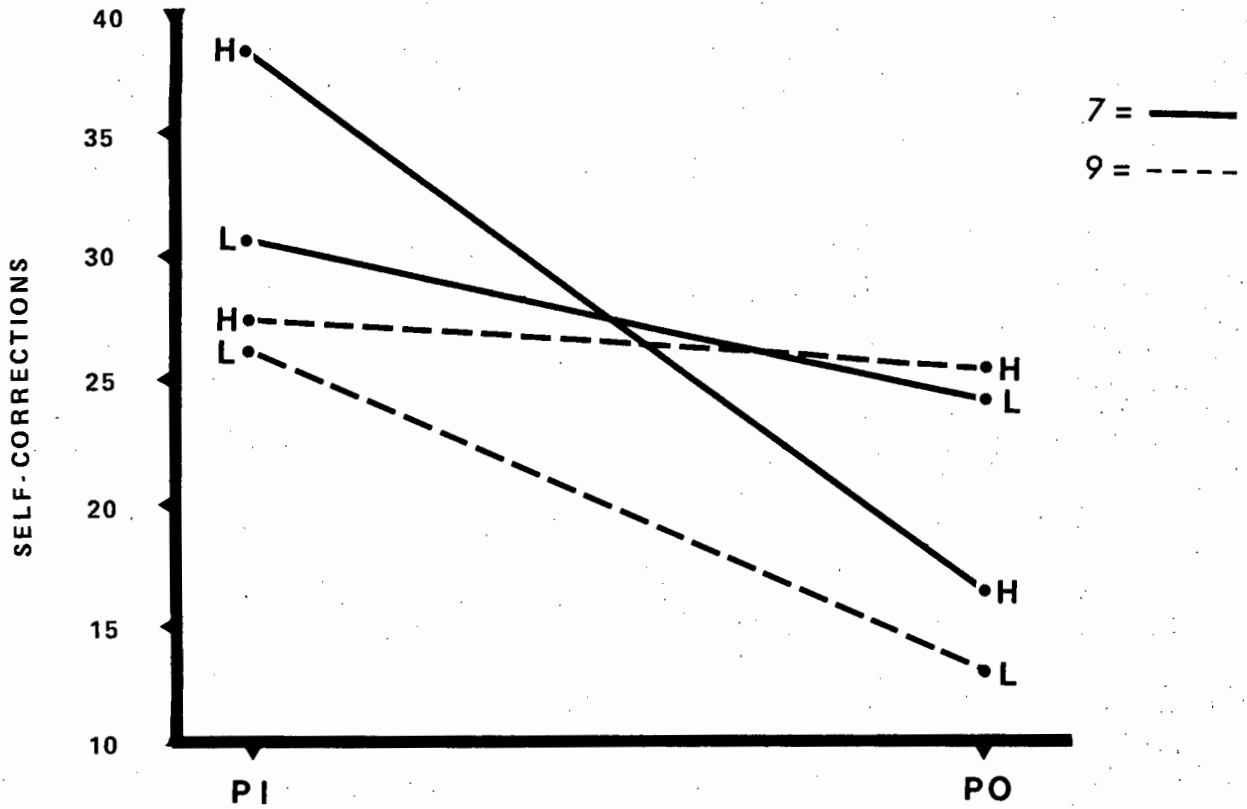


TABLE 28b

ANALYSIS OF VARIANCE SUMMARY TABLE:
SELF-CORRECTIONS

P	3649,83	1	3649,83	21,457++
Pr	335,07	1	335,07	1,97
RA	591,94	1	591,94	3,48
PxPr	13,35	1	13,35	0,08
PxRA	305,86	1	305,86	1,80
PrxRA	301,85	1	301,85	1,78
PxPrxRA	1328,57	1	1328,27	7,81++
WITHIN	19051,42	112	170,10	

TABLE 28c

SIMPLE INTERACTION EFFECTS : SELF-CORRECTIONS

SOURCE	SS	df	MS	F RATIO
PxPr at 7	803,85	1	803,85	4,73+
at 9	537,78	1	537,78	3,16
PxRA at H	1454,45	1	1454,45	8,55++
L	179,71	1	179,71	1,06
PrxRA at P1	181,87	1	181,87	1,07
P0	1448,40	1	1448,40	8,52++
WITHIN	19051,42	112	170,10	

On the basis of the significant simple interactions P x Pr at 7, P x RA at H and Pr x RA at P0, the following simple-simple main effects are relevant:

TABLE 28d

SIMPLE-SIMPLE MAIN EFFECTS : SELF CORRECTIONS

SOURCE	SS	df	MS	F RATIO
P at 7,H	3480,76	1	3480,76	20,46+++
7,L	357,21	1	357,21	2,10
Pr at 7,P1	420,83	1	420,83	2,47
7,P0	383,53	1	282,52	2,26
(P at 7,H	3480,76	1	3480,76	20,46+++)
9,H	25,63	1	25,63	0,15
RA at H,P1	926,85	1	926,85	5,45+
H,P0	551,87	1	551,87	3,24
(Pr at 7,P0	383,53	1	383,53	2,26)
9,P0	1172,25	1	1172,25	6,89++
(RA at H,P0	551,87	1	551,87	3,24)
L,P0	919,97	1	919,97	5,41+
WITHIN	19051,42	112	170,10	

On the basis of the non-significant simple interactions P x Pr at 9; P x RA at L and Pr x RA at P1 (Table 27c), the following simple main effects are relevant:

TABLE 28c

SIMPLE MAIN EFFECTS : SELF-CORRECTIONS

SOURCE	SS	df	MS	F RATIO
P at 9	921,12	1	921,12	5,42+
Pr at 9	636,55	1	636,55	3,74
P at L	1610,77	1	1610,77	9,47++
RA at L	869,59	1	869,59	5,11+
Pr at P1	241,04	1	241,04	1,42
RA at P1	874,32	1	874,32	5,14+
WITHIN	19051,42	112	170,10	

Illustration (P) effects:

Given the significant $P \times Pr \times RA$ interaction ($F = 7,81++$: Table 27b), analysis of the simple interaction effects (Table 27c) reveals a significant $P \times Pr$ interaction at 7 ($F = 4,73+$) and a significant $P \times RA$ interaction at H ($F = 3,55++$). The simple-simple main effect for P at 7H is highly significant ($F = 20,46+++$) while for P at 7L and for P at 9H the effects are non-significant ($F = 2,10$; $F = 0,15$ respectively: Table 27d).

Going from the non-significant simple interactions ($P \times Pr$ at 9, $F = 3,16$; $P \times RA$ at L, $F = 1,06$: Table 27c), the simple main effects for P both at 9 and at L are significant

($F = 5,42+$; $F = 9,47++$ respectively: Table 27e).

What emerges, therefore, is a strong and significant P effect for the 7H group and a more moderate but significant effect for the combined groups at 9 (9H and 9L) and at L (7L and 9L). Given the clear trend to interaction (P x Pr at 9, $F = 3,16$; $p < ,10$: Table 27c), however, as well as the clear non-significance of the simple-simple main effect for P at 9H ($F = 0,15$: Table 27d), it would appear that the P effect at 9 is strongly influenced by a substantial effect for the 9L group and that the P effect at 9H is effectively non-significant. This is consistent with the pattern of cell mean differences (Table 27a) and the relative P effects as shown in Figure 9.

On these grounds, the results for the P effect may be taken as evidence that illustrations strongly increase the rate of self-correction for high progress readers at reading age seven; and that the increase is moderate for all low progress readers and for all readers at reading age nine. Within this there are clear trend indications, however, that at reading age nine, the increase is relatively strong for low progress readers and only marginal for high progress readers.

Progress level (Pr) effects

The simple interaction effects (Table 27c) reveal a significant P x Pr interaction at 7 ($F = 4,73+$) and a significant Pr x RA interaction at P0 ($F = 3,52++$). Only the simple-simple main effect for Pr at 9P0 ($F = 6,89++$) is significant

while Pr at 7P1 ($F = 2,47$) and Pr at 7P0 ($F = 2,26$) are both non-significant (Table 27d).

Going from the non-significant simple interactions (P x Pr at 9, $F = 3,16$ and Pr x RA at P1, $F = 1,07$: Table 27c) the simple main effects for Pr both at 9 and at P1 are non-significant ($F = 3,74$; $F = 1,42$ respectively: Table 27e).

It appears that the rate of self-correction is significantly lower in the 9L than in the 9H group where there are no illustrations. Apart from this there are no significant progress level differences either generally or at P1 and P0. The trend to interaction of P x Pr at 9 ($F = 3,16$, $p < ,10$) mentioned earlier merely confirms, in this context, the difference between 9H and 9L at P0 but not at P1.

Reading age (RA) effects

The simple interaction effects (Table 27c) reveal a significant P x RA interaction at H ($F = 3,55++$) and a significant Pr x RA interaction at P0 ($F = 3,52++$). The simple, simple main effects for RA both at HP1 ($F = 5,45+$) and at LP0 ($F = 5,41+$) are significant while RA at HP0 is non-significant ($F = 3,24$: Table 27d).

Going from the non-significant simple interactions (P x RA at L, $F = 1,06$ and Pr x RA at P1, $F = 1,07$: Table 27c), the simple main effects for RA both at L and at P1 are significant ($F = 5,11+$; $F = 5,14+$ respectively: Table 27e).

This would result in a higher overt self-correction rate for readers at reading age seven than at reading age nine as, in fact, appears at P1, and also overall for the low progress readers at reading age seven and nine respectively. That the same pattern is not as clear at P0 and for the high progress readers is attributable to the 9H group retaining virtually the same self-correction rate at P0 as at P1 while the 7H group appears to be particularly susceptible to the removal of illustrations at P0 so that their overt self-correction rate, while expectedly higher than for the 9H's at P1, is considerably lower than for the 9H's at P0.

Although the difference between 7H and 7L is not significant at P0, there is a clear trend for the self-correction rate to be lower for 7H than for 7L. Ostensibly this is counter intuitive as it might be expected that high progress readers would self-correct at a higher rate than low progress readers with or without illustrations. The grade level effect, however, appears to be influencing the pattern again as it appears that 7H (grade I) readers are relatively less able to generate self-corrections without the contextual support of illustrations than are the 7L (grade III) readers.

Thus, the general pattern of results on self-corrections would appear to suggest that for high progress readers at reading age nine, illustrations have little or no effect on the rate of overt self-correction. Those overt self-corrections that are made appear to be constrained

mainly be text-based information and are relatively independent of the complimentary contextual information in illustrations. For both low and high progress readers at reading age seven, however, the overt self-correction rate is higher with illustrations than without. Particularly for the high progress, grade I, readers it would appear that the relative accessibility of semantic information (see under results for semantic acceptability of errors) across illustration conditions has a very marked effect on the rate at which errors are detected and corrected.

For the low progress readers at reading age nine there appears to be a dual effect. Not only are self-corrections more likely to be covert in terms of reading age and grade level but, without illustrations, overt self-corrections, where they do occur, are also significantly lower than for the high progress readers at reading age nine. This would seem to indicate that the low progress readers are relatively less able than the high progress readers at this developmental level to utilize text-based information in the overt detection and correction of errors. Within this, however, the addition of illustrations makes a substantial difference to the self-correction rate of these readers and it is again likely that the greater accessibility of semantic information creates a more meaningful contextual framework within which relatively more errors are detected and overtly corrected.

vi Literal Comprehension

Scores on this variable are represented in the following formula:

$$\text{Literal comprehension} = 100 \left(\frac{a+b+c+d}{10} \right)$$

where: a = Score (0-2) for 'Detail'

b = Score (0-2) for 'Cause-Effect'

c = Score (0-3) for 'Main Idea'

d = Score (0-3) for 'Related Ideas'

TABLE 29a

CELL MEANS/STANDARD DEVIATIONS, CELL
MEAN DIFFERENCES (P1-PO) AND TREATMENT
MEANS : LITERAL COMPREHENSION

		P1	P0	DIFF
7	H	56,67	29,36	27,31
		(15,89)	(15,80)	
	L	62,67	48,00	14,67
		(12,80)	(19,35)	
9	H	58,67	51,33	7,34
		(19,22)	(17,27)	
	L	78,00	50,67	27,33
		(13,20)	(21,20)	

	P1	P0
H	57,67	40,34
L	70,33	49,33

	P1	P0
7	59,63	38,68
9	68,33	51,00

	H	L
7	43,01	55,33
9	55,00	64,33

P1	64,00	H	49,01	7	49,17
P0	44,84	L	59,83	9	59,67

FIGURE 10

GRAPH OF CELL MEAN PROFILES:

LITERAL COMPREHENSION

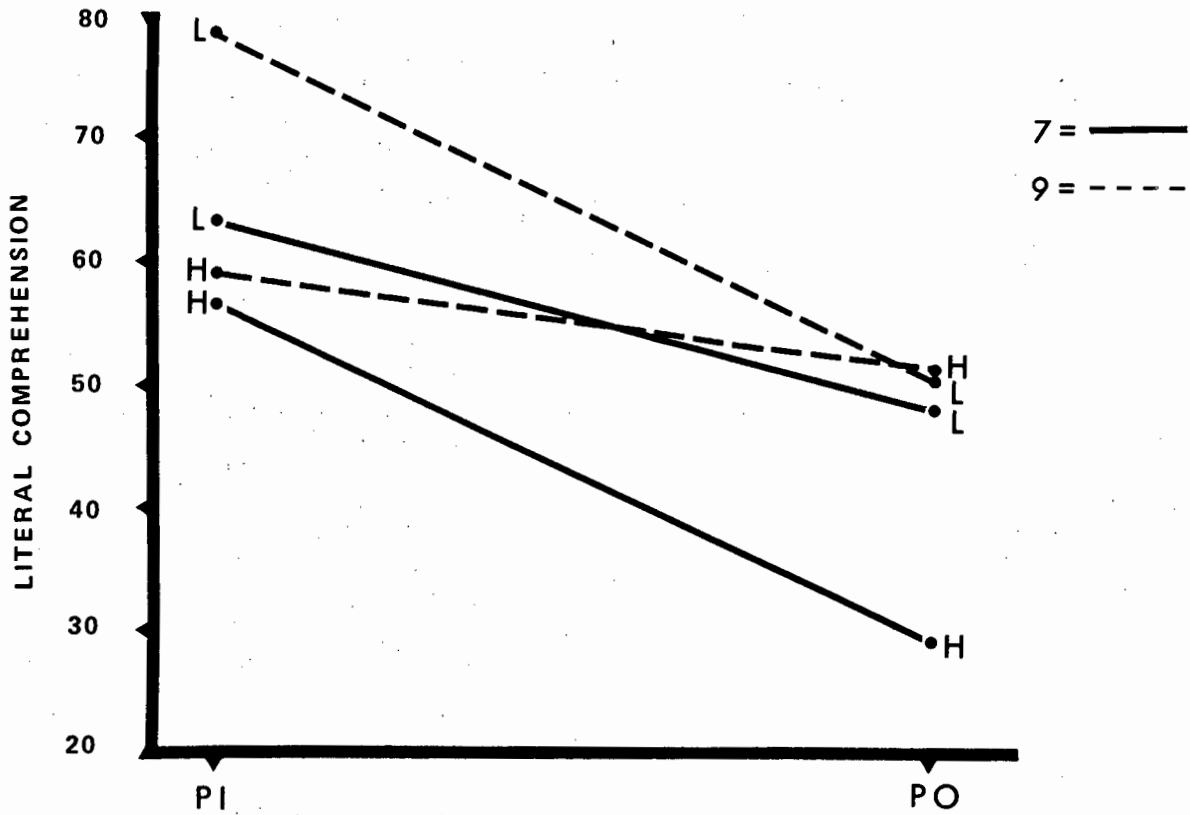


TABLE 29b

ANALYSIS OF VARIANCE SUMMARY TABLE:

LITERAL COMPREHENSION

SOURCE	SS	df	MS	F RATIO
P	11014,51	1	11014,51	37,79+++
Pr	3517,26		3517,26	12,07+++
RA	3304,04	1	3304,04	11,34++
PxPr	101,44	1	101,44	0,35
PxRA	100,23	1	100,23	0,34
PrxRA	67,01	1	67,01	0,23
PxPrxRA	1998,14	1	1998,14	6,86++
WITHIN	32640,54	112	291,43	

Given the significant P x Pr x RA interaction, analysis of the simple interaction effects is relevant:

TABLE 29c

SIMPLE INTERACTION EFFECTS:

LITERAL COMPREHENSION

SOURCE	SS	df	MS	F RATIO
PxPr at 7	599,58	1	599,58	2,06
at 9	1500,00	1	1500,00	5,15+
PxRA at H	1496,70	1	1496,70	5,14+
at L	601,67	1	601,67	2,07
PrxRA at P1	666,67	1	666,67	2,29
P0	1398,48	1	1398,48	4,80+
WITHIN	32640,54	112	291,43	

On the basis of the significant simple interactions P x Pr at 9, P x RA at H and Pr x RA at P0, the following simple-simple effects are relevant:

TABLE 29d

SIMPLE-SIMPLE MAIN EFFECTS:LITERAL COMPREHENSION

SOURCE	SS	df	MS	F RATIO
P at 9,H	403,33	1	403,33	1,38
9,L	5603,34	1	5603,34	19,23++
Pr at 9,P1	2803,33	1	2803,33	9,62++
9,P0	3,33	1	3,33	0,01
P at 7,H	5594,32	1	5594,32	19,20+++
(9,H	403,33	1	403,33	1,38)
RA at H,P1	30,00	1	30,00	0,10
H,P0	3622,74	1	3622,74	12,43+++
Pr at 7,P0	2607,18	1	2607,18	8,95++
(9,P0	3,33	1	3,33	0,01
(RA at H,P0	3622,74	1	3662,74	12,43+++
L,P0	53,33	1	53,33	0,18
WITHIN	32640,54	112	291,43	

On the basis of the non -significant simple interactions,
 P x Pr at 7, P x RA at L and Pr x RA at P1 (Table 29c)
 the following simple main effects are relevant:

TABLE 29e

SIMPLE MAIN EFFECTS : LITERAL COMPREHENSION

SOURCE	SS	df	MS	F RATIO
P at 7	6608,08	1	6608,08	22,67+++
Pr at 7	2277,60	1	2277,60	7,82++
P at L	6615,00	1	6615,00	22,70+++
RA at L	1215,00	1	1215,00	4,17+
Pr at P1	2406,67	1	2406,67	8,26++
RA at P1	1126,67	1	1126,67	3,87
WITHIN	32640,54	112	291,43	

InterpretationIllustration (P) effects:

Given the significant P x Pr x RA interaction (F = 6,86++: Table 29b), analysis of the simple interaction effects (Table 29c) reveals a significant P x Pr interaction at 9 (F = 5,15+) and a significant P x RA interaction at H (F = 5,15+). The simple-simple main effects for P at 9 H (F = 19,23+++), and for P at 7H (F = 19,20+++), are highly significant while P at 9 H (F = 1,38) is non-significant (Table 29d).

Going from the non-significant simple interactions (P x Pr at 7, F = 2,06; P x RA at L, F = 2,07; and Pr x RA

at P1, $F = 2,29$: Table 29c), the simple main effects for P both at 7 and at L are highly significant ($F = 22,67+++$; $F = 22,70+++$ respectively: Table 29e).

What emerges is a strong and significant P effect for both the 9L and the 7H groups, a non-significant P effect for the 9H group; and a strong and significant P effect for the combined groups at 7 (7H and 7L) as well as at L (7L and 9L). There is some indication that the relative strength of the P effect for both 7H and 9L is largely responsible for the strength of the effect at 7 and L respectively since a more moderate effect for 7L is apparent in the pattern of all mean differences (table 29a) and cell mean profiles (Figure 10). Further evidence to confirm this interpretation emerges from the Pr and RA effects, below.

On this interpretation, it appears that illustrations strongly increase the level of literal comprehension for high progress readers at reading age seven as well as for low progress readers at reading age nine. For high progress readers at reading age nine, however, illustrations appear to have no effect on literal comprehension while, for low progress readers at reading age seven, the effect appears to be only moderate.

Progress level (Pr) effects:

The simple interaction effects (Table 29c) reveal a significant P x Pr interaction at 9 ($F = 5,15+$) and a significant Pr x RA interaction at P0 ($F = 4,80+$). The

simple-simple main effects for Pr at 9 P1 ($F = 9,62++$) and for Pr at 7 P0 ($F = 8,95++$) are both significant while Pr at 9 P0 ($F = 0,01$) is non-significant (Table 29d).

Going from the non-significant simple interactions (P x Pr at 7, $F = 2,06$ and Pr x RA at P1, $F = 2,29$: Table 29c), the simple main effects for Pr both at 7 and at P1 are significant ($F = 7,82++$; $F = 8,26++$: Table 29e).

It appears that where illustrations are available, literal comprehension is significantly higher for the 9L group than the 9H group as it is for the combined groups at L compared to the combined groups at H. Where there are no illustrations, literal comprehension is significantly lower for the 7H group than the 7L group as it is overall for these two groups.

Reading age (RA) effects:

The simple interaction effects (Table 29c) reveal a significant P x RA interaction at H ($F = 5,14+$) and a significant Pr x RA interaction at P0 ($F = 4,80+$). The simple-simple main effect for RA at H P0 is significant ($F = 12,43+++$) while RA at H P1 and RA at L P0 ($F = 0,10$; $F = 0,18$ respectively: Table 29d) are both non-significant.

Going from the non-significant simple interactions (P x RA at L, $F = 2,07$ and Pr x RA at P1, $F = 2,29$: Table 29c), the simple main effects for RA at L is significant ($F = 4,17+$) while RA at P1 is non-significant ($F = 3,87$: Table 29e).

It appears that where no illustrations are available literal comprehension is significantly lower for the 7H group than the 9H group. It is also significantly higher for 9L than 7L overall.

Interpreting the Pr and RA effects together, the following pattern emerges:

At P0: Literal comprehension for the 7H group is significantly lower than for the 7L, the 9H and (by implication) the 9L groups.

At P1: Literal comprehension for the 9L group is significantly higher than for the 9H group. It is also significantly higher for the combined groups at L (7L and 9L) than for the combined groups at H (7H and 9H). On the pattern of cell mean profiles (figure 10) it is apparent that the latter effect is largely due to the difference between 9L and 9H and that 7L and 7H are only marginally different. In other words, the interpretation of a moderate P effect for 7L in comparison with the strong effects for 7H and 9L (see above) is confirmed in the significantly lower literal comprehension of 7H in relation to 7L at P0 and the substantially higher literal comprehension of 9L in relation to 7L at P1.

Overall: Literal comprehension is significantly higher for 9L than for 7L which, in turn, is significantly higher than for 7H.

Once again, the grade level effect is apparent in these results. For instance, if the P effect is temporarily

suspended, the mean literal comprehension for the 7H (grade I) group is 43,01; for the 7L and 9H (grade III) groups it is 55,33 and 55,00 respectively; and for the 9L (grade V) group it is 64,33 (Table 29a). As a tendency this is not surprising. With the story content and the nature of the comprehension questions (apart from minor wording differences at the two reading age levels) being the same for all groups it might be expected that literal comprehension should improve with grade level.

The influence of the P effect on this basic pattern is to decrease the literal comprehension of the 7H (grade I), the 9L (grade V) and, to a lesser extent, the 7L (grade III) groups relative to the 9H (grade III) group under the no-illustration condition; and to increase the literal comprehension of the 7H, the 9L and, to a lesser extent, the 7L groups relative to the 9H group under the illustration condition.

Thus, the general pattern of the results suggests that a grade level effect operates for literal comprehension but that this is modified by the relative informational value of illustrations in the process of literal comprehension for the various groups. What is of particular interest is the remarkable similarity of this pattern to that found for the semantic acceptability of errors. It seems highly likely, therefore, that the relative influence of illustrations on literal comprehension is also related to the accessibility of semantic information. In other words, with no illustrations, semantic information would appear to be far less accessible to high progress

readers at reading age seven (grade I), thus affecting their literal comprehension, than it is to the older readers who appear able to make more use of text-based semantic information and to comprehend significantly better. Thus the expected grade level inferiority operates for this group where illustrations are not available but where they are available, and access to semantic information is increased, high progress readers at this level comprehend at a similar level to grade III readers. For high progress readers at reading age nine (grade III) the addition of illustrations makes no effective difference to their literal comprehension. Although this group showed a moderate increase in use of semantic information across illustration conditions, this does not appear to have materially affected their literal comprehension. On the other hand, low progress readers at reading age seven (grade III) showed a similar, moderate increase in use of semantic information across illustration conditions and this is echoed in a moderate increase in literal comprehension. Since the structure and vocabulary of the stories differed between RA7 and RA9 (chapter IX) it is possible that the semantic information in illustrations was substantially more directly relevant to literal comprehension for the 7L than for the 9H group. (The difference in level of story difficulty also explains why high progress readers (grade III) do not comprehend at a higher level than low progress readers (Grade III)). Finally, for low progress readers at reading age nine (grade V) a substantial increase in the accessibility

of semantic information (from a moderate to a high level across the illustration conditions) would appear to be related to the similar and substantial increase in literal comprehension. In other words, the expected grade level effect does not operate where no illustrations are available but, where semantic information is made more accessible through illustrations, low progress readers are able to demonstrate their grade level superiority on literal comprehension.

vii. Inferential Comprehension

Scores on this variable are represented in the following formula:

$$\text{Inferential Comprehension} = 100 \left(\frac{a}{3} \right)$$

where: a = Score (0-3) for 'Inference'

TABLE 30a

CELL MEANS/STANDARD DEVIATIONS, CELL MEAN DIFFERENCES (P1-P0), and TREATMENT MEANS: INFERENCE COMPREHENSION

	H	31,11	24,44	6,67
7		(32,04)	(38,76)	
	L	57,78	46,67	11,11
		(40,76)	(43,28)	
	H	68,89	46,67	22,22
9		(19,79)	(37,37)	
	L	55,56	53,33	2,23
		(29,99)	(37,38)	

	P1	P0
H	50,00	35,56
L	56,67	50,00

	P1	P0
7	44,44	35,56
9	62,22	50,00

	H	L
7	27,78	52,22
9	57,78	54,45

P1	53,33	H	42,78	7	40,00
P0	42,78	L	53,33	9	56,11

FIGURE 11

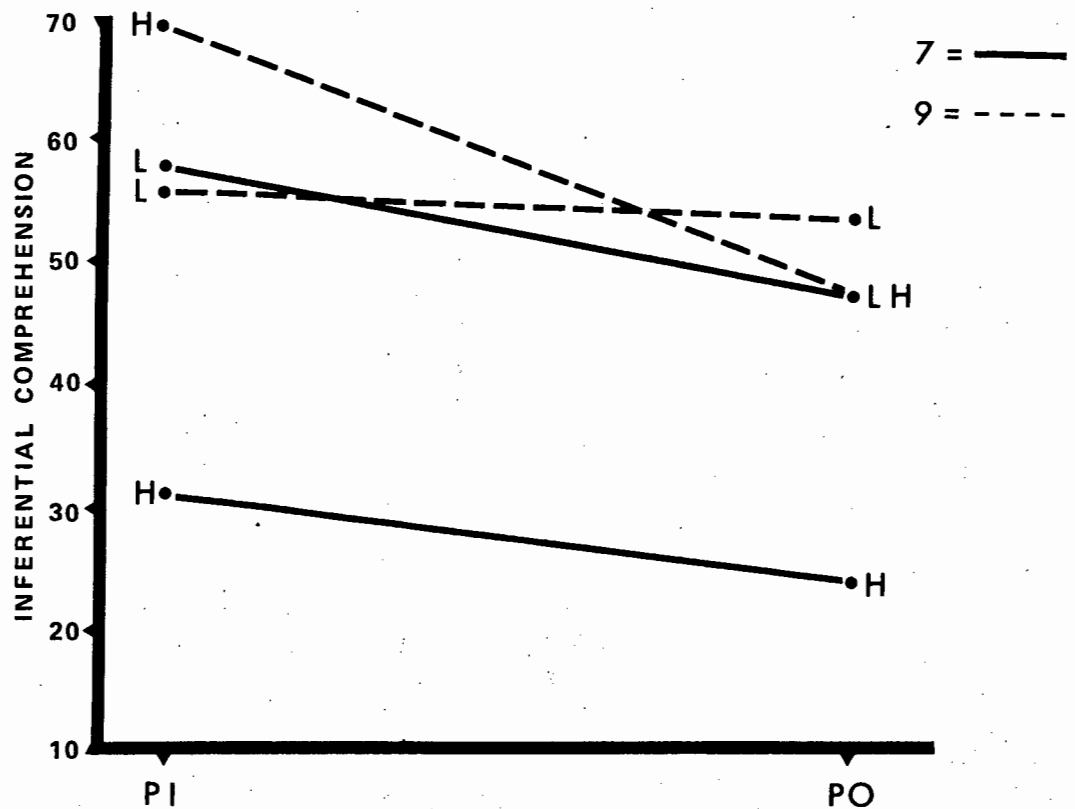
GRAPH OF CELL MEAN PROFILES:INFERENCEAL COMPREHENSION

TABLE 30b

ANALYSIS OF VARIANCE SUMMARY TABLE:INFERENCEAL COMPREHENSION

SOURCE	SS	df	MS	F RATIO
P	3342,99	1	3342,99	2,64
Pr	3343,20	1	3343,20	2,64
RA	7788,07	1	7788,07	6,14+
PxPr	453,68	1	453,68	0,36
PxRA	83,41	1	83,41	0,07
PrxRA	5786,98	1	5786,98	4,56+
PxPrxRA	1120,53	1	1120,53	0,88
WITHIN	142079,09	1	1268,56	

On the basis of the significant Pr x RA interaction, the following simple main effects are relevant.

TABLE 30c

SIMPLE MAIN EFFECTS: INFERENCEAL COMPREHENSION

Pr at 7	8963,61	1	8963,61	7,07++
at 9	166,57	1	166,57	0,13
Ra at H	13500,90	1	13500,90	10,64++
at L	74,15	1	74,15	0,06
WITHIN	142079,09	112	1268,56	

Interpretation

Illustration (P) effects:

Given the non-significance of all P interactions as well as the P main effect, it can be concluded that illustrations have no significant effect on inferential comprehension.

Progress level (Pr) effects:

Given the significant Pr x RA interaction ($F = 4,56+$: Table 30b), the simple main effect for Pr at 7 is significant ($F = 7,07++$). While Pr at 9 is non-significant ($F = 0,3$: Table 30c). Inferential comprehension for the 7H group is significantly lower than for the 7L group.

Reading age (RA) effect:

On the same basis, the simple main effect for RA at H is significant ($F = 10,64++$) while RA at L is non-significant ($F = 0,06$: Table 30c). Inferential comprehension for the 7H group is significantly lower than for the 9H group.

Interpreting the Pr and RA results together, this simply conveys that inferential comprehension is significantly lower for high progress readers at reading age seven than for the other three groups (9L as well, by implication). This is probably best explained in terms of a partial grade level effect where the youngest group (grade I) are relatively less sophisticated or experienced in answering this sort of comprehension question than children at the higher grade levels.

Although not to be taken as significant, what is nevertheless interesting in this result is the trend for the high progress readers at reading age nine to show the only relatively clear difference in inferential comprehension across the illustration conditions. This is in contradiction to the results on all other variables. It may be a chance effect as the analysis of variance would seem to indicate. On the other hand it may be a genuine tendency that might be explained in the high progress reader at this level, in contrast to other groups, being able to integrate information from illustrations into the higher level of comprehension required in the process of inference. Although this cannot be deduced on the present evidence, it is

suggestive and may be worth further investigation on the basis of a wider range of inferential comprehension questions.

viii Summary of the Illustration Effect over Analysis of Variance Results.

Since the principal aim is to identify the direction and significance of the P effect over the four experimental groups and within this to compare the relative strength of the effect, the relevant results as interpreted above will be summarised for the seven dependent variables in tabular form.

TABLE 31

SUMMARY TABLE OF DIRECTION, SIGNIFICANCE, AND RELATIVE STRENGTH OF ILLUSTRATION (P) MAIN EFFECTS OVER ALL DEPENDENT VARIABLES

D.V.	GROUP	DIRECTION	SIGNIFICANCE		STRENGTH ^{1.}	
			S-S-MAIN	S-MAIN		
i. Word identi- fication accuracy	7H	P1 > P0	+++	}+++	S	
	7L	P1 > P0			M	
	9L	P1 > P0			M	
	9H	P1 > P0	NS	}NS	O	
ii. Seman- tic Accep- tability of Errors	7H	P1 > P0	+++	}+++	S	
	7L	P1 > P0			NS	M
	9L	P1 > P0			}++	S
	9H	P1 > P0	NS		}+++	M

D.V.	GROUP	DIRECTION	SIGNIFICANCE			STRENGTH
			S-S- MAIN	S- MAIN	MAIN	
cont..						
iii. Syn- tactic Acceptab- ility of Errors	7H	P1 > P0			} ++	M
	7L	P1 > P0				M
	9L	P1 > P0				M
	9H	P1 > P0				M
iv. Ortho- graphic Acceptab- ility of Errors						
iv. Ortho- graphic Acceptab- ility of Errors	7H	P1 < P0			} ++	M(S)
	7L	P1 < P0				M(O)
	9L	P1 < P0				M(O)
	9H	P1 < P0				M(O)
v. Self- corrections						
v. Self- corrections	7H	P1 > P0	+++		} ++	S
	7L	P1 > P0	NS			M
	9L	P1 > P0				S
	9H	P1 > P0	NS			0
vi. Liter- al Compre- hension						
vi. Liter- al Compre- hension	7H	P1 > P0	+++		} +++	S
	7L	P1 > P0				M
	9L	P1 > P0	+++			S
	9H	P1 > P0	NS			0
vii. Inf- erential Compre- hension						
vii. Inf- erential Compre- hension	7H	P1 > P0			} NS	0
	7L	P1 > P0				0
	9L	P1 > P0				0
	9H	P1 > P0				0

¹ STRENGTH: S = strong
M = moderate
0 = non-significant
() = trend interpretation

LIT : Literal Comprehension

INF : Inferential Comprehension.

The correlations between word identification accuracy and most other variables are moderate and significant. Two exceptions are the correlations with orthographic acceptability of errors and inferential comprehension. It would appear, in particular, that the degree of accuracy achieved in word identification is related to the degree of semantic (ACC x SEM, $r = ,449++$) and syntactic (ACC x SYN, $r = ,453++$) acceptability of errors (i.e., use of semantic and syntactic information) but not to the degree of orthographic (ACC x ORT, $r = -,139$) acceptability of errors (i.e., use of orthographic information). Since accuracy of word identification must be taken as a basic criterion of reading development, this set of relationships is particularly interesting in the light of the inverse effect of illustrations on the use of contextual information (syntactic acceptability and, particularly, semantic acceptability) versus orthographic information (orthographic acceptability). In other words, it would appear that the negative effect of illustrations on the use of orthographic information, despite the claims of the focal attention hypothesis, is of little consequence in determining word identification accuracy. However, the positive effect of illustration on the use of semantic and syntactic information would appear to have significant consequences in determining word identification accuracy.

The relatively high correlation between word identification accuracy and self-correction (ACC x COR, $r = ,587++$) is of considerable interest. Not only does this support the idea that the strategy of self correction is related to optimal reading development (chapter IV) but, since illustrations have a positive effect on rate of self-correction, the importance of this effect in the general context of reading development is emphasised. The significant, although somewhat lower, correlation between literal comprehension and self-correction (LIT x COR, $r = ,259++$) lends support to this interpretation.

The relationships between word identification accuracy and literal comprehension (ACC x LIT, $r = ,424++$) as well as inferential comprehension (ACC x INF, $r = ,185+$) are not unexpected. Correlations of a moderate order are common for most reading tests that measure accuracy as well as comprehension (e.g. Neale, 1958) and it is clear that at least literal comprehension must, under most circumstances, be influenced by the degree to which the words in the text have been correctly identified. What is of interest is that the relationship is much lower and barely of significance for inferential comprehension. This lends some support to the idea that the processing involved in inferential comprehension is, at least relatively, of a different order to that involved in literal comprehension (chapter IV). Again, the relatively low correlation between literal and inferential comprehension (LIT x INF, $r = ,317++$), although significant lends further support to this interpretation: on a priori assumptions of relatedness, the correlation between the

The significant and relatively strong negative correlation between semantic acceptability and orthographic acceptability (SEM x ORT, $r = -.559++$) has two important implications.

First, it endorses the basic contextual hypothesis that the more contextual information the reader has available, the less need he rely on orthographic information.

Second, when compared with the low and non-significant correlation between syntactic and orthographic acceptability (SYN x ORT, $r = -.103$) it is apparent that it is, in this context, mainly the availability of semantic information that creates the contextual effect and reduces the reader's reliance on orthographic information. Interpretation of the illustration effect as principally affecting the relative accessibility of semantic information would appear to be given some support in this particularly as the data on which the correlation is based are derived from a situation in which it is illustration information that is systematically varied¹.

The significant and moderate correlation of semantic acceptability with self-correction (SEM x COR, $r = .417++$) is interesting, particularly when compared with the non-significant correlation of syntactic acceptability with self-correction

¹ Separate correlations at P0 and at P1 (N = 60) support this. At P0, SEM x ORT = $-.540++$. In other words, with no illustration information, use of orthographic information is high and use of semantic information is low. At P1, SEM x ORT = $-.457++$. With illustrations increasing semantic access, use of semantic information is high and use of orthographic information is reduced.

(SYN x COR, $r = ,089$) and the significant and low negative correlation of orthographic acceptability with self-correction (ORT x COR, $r = -,247++$). In other words, the rate of error detection and correction is positively related to the degree to which semantic information is used; is negatively related to the degree to which orthographic information is used; and bears no relation to the degree to which syntactic information is used. This indicates that self-correction, in this context at least, is more directly based on the reader's access to, and use of, semantic information than on his use of syntactic information and that a high use of orthographic information is, in fact, negatively related to self-correction. Since there was a strong reading age effect on syntactic acceptability of errors with the reading age nine group using syntactic constraints more than the reading age seven group; and since there was a similar, although less clear and inverse effect on self-corrections with the reading age seven group tending to make more (overt) self-corrections than the reading age nine group, it is possible that the low correlation between use of syntactic information and self-correction is partly a product of this inverse effect. It is likely that syntactic information is used in the overt self-corrections of the reading age seven group (Clay, 1969; 1972) but for those at reading age nine it is possible that their high sensitivity to syntactic constraints (74,60%) is the source of more covert self-corrections. In support of this, the relevant correlation ($N = 60$) for the RA7 group (SYN x COR, $r = ,413++$) is moderate and significant while that for the RA9 group (SYN

is positively related to the use of both semantic and syntactic information - although possibly more to the former - while it is negatively related to the use of orthographic information. Since literal comprehension necessarily relies on identifying the meaning of the text and since this must also rely, to some extent, on following the linguistic structure of the text, the correlations with semantic and syntactic acceptabilities are not unexpected. What is of interest is that a high reliance on orthographic information appears to reduce the effectiveness of literal comprehension. On the evidence that there is a relatively strong negative relationship between use of semantic and orthographic information, this becomes explicable and further reduces the claim under the focal attention hypothesis that reduced attention to orthographic information is necessarily damaging to reading development. With inferential comprehension, on the other hand, which relies less directly on the explicit content of the text, all the respective relationships are weaker and, in the case of orthographic acceptability, the relationship is non-significant. If anything, it is somewhat surprising to find the semantic and syntactic relationships with inferential comprehension as significant at all. However, some degree of text-based information is obviously necessary to generate appropriate inferences and the relationships probably reflect no more than this.

The remaining correlation to be discussed is between self-corrections and inferential comprehension. Since this is non-significant (COR x INF, $r = .121$) little more can be

said other than comparing it with the correlation between self-corrections and literal comprehension (COR x LIT, $r = .259++$) where it appears that self-corrections are related to literal comprehension while they have little or no relation to inferential comprehension. Again, given that literal comprehension depends partly on explicit identification of the content of the text while inference is less dependent on this, the finding is not surprising.

Attempting to summarise the patterns in the set of correlations, it can be said that there is a moderate and significant positive relationship between all of word identification accuracy (ACC), semantic acceptability (SEM), self-corrections (COR) and literal comprehension (LIT). Conversely there is a consistent, and occasionally significant, negative relationship between orthographic acceptability (ORT) and all other variables. The relationship of syntactic acceptability (SYN) to word identification accuracy and to semantic acceptability is moderate and significant but to other variables it is generally weak and only occasionally significant. The relationship of inferential comprehension (INF) to other variables is also significant on occasion but generally weak.

If word identification accuracy and literal comprehension are taken as central achievement criteria in reading development, then it is clear that the strategies of self-correction and the use of semantic information are strong predictors of these criteria and must, themselves, be regarded as important in evaluating the relevance of the

illustration effect on reading development. By the same argument, the strategy of use of syntactic information appears less relevant and the strategy of use of orthographic information appears to be a consistent although weak negative predictor. The third achievement criterion of inferential comprehension appears to be relatively independent of both word identification accuracy and literal comprehension as well as the measures of strategy.

The clear inter-relationship between word identification accuracy, semantic acceptability, self-correction and literal comprehension would also seem to endorse the interpretation, in the analysis of variance results, that the accessibility of semantic information is common and crucial to relative performance on all these variables.

Finally, although this set of correlations is legitimate in its own right, it is interesting to compare the generality of the patterns across the illustration (P1) and no-illustration (P0) conditions. Without developing this in detail, comparisons of the relevant correlations revealed, in fact, a similar positive and negative pattern with correlations under the P1 condition being generally somewhat lower than those under the P0 condition. Comparisons of all correlations under the two conditions revealed only three that were significantly different. These were related and are, in themselves, of some interest to interpret. The first was between word identification accuracy and syntactic acceptability (P1 ACC x SYN, $r = ,089$; P0 ACC x SYN, $r = ,514++$; $z = 2,57+$). This difference indicates

that use of syntactic information is more important in determining word identification accuracy under the no-illustration condition than under the illustration condition. The second difference was between correlations for literal comprehension and syntactic acceptability (PI LIT x SYN, $r = -.190$; PO LIT x SYN, $r = .455++$: $z = 3.65++$). This also indicates that literal comprehension is less constrained by the use of syntactic information under the illustration condition. Both differences emphasise that text-based linguistic information is more important where the additional contextual information in illustrations is not available. They also indicate that reliance on syntactic information is possibly greater where semantic information is less accessible and vice-versa. Intuitively this makes sense and is an interesting possibility. The third difference was between correlations for word identification accuracy and literal comprehension (PI ACC x LIT, $r = -.068$; PO ACC x LIT, $r = .491++$: $Z = 2.51+$). On a similar pattern this also indicates that literal comprehension is less constrained by accurate word identification where the extra-textual information in illustrations is available.

CHAPTER XI

CONCLUSIONS

Since the conclusions will be discussed in relation to the contextual hypothesis it is convenient to state it again:

'Where illustrations are sufficiently and relevantly related to continuous text, they constitute a source of contextual information that, in particular, compliments and is redundant within the semantic dimension of information. In illustrated reading, therefore, the efficiency and accuracy of message identification and the effectiveness of comprehension will be facilitated through reducing the reader's reliance on orthographic information while increasing his use of the available contextual information.'

Initially, conclusions will be discussed in terms of the major headings of 1. Accuracy, 2. Strategy, and 3. Comprehension, and within each of these, conclusions applicable to each experimental group, or combined groups where relevant, will be discussed. Following this, wider or more general conclusions will be discussed.

1. Accuracy

The criterion under this heading is the accuracy of message identification. As argued in chapter IV, the most valid

variables to consider in evaluating the illustration effect on this criterion are word identification accuracy moderated by the degree of semantic acceptability of errors.

Interpretation of the results for word identification accuracy indicated that there was a positive significant and strong illustration effect for the high progress readers at reading age seven. The same applied to the semantic acceptability of errors. For this group of readers, therefore, it can be concluded that the illustration effect is substantial and that not only is the accuracy of word identification negatively affected by the lack of illustrations but errors that are made are more meaningful where illustrations are present than where they are absent. The accuracy with which the textual message is identified is therefore clearly facilitated by illustrations.

Essentially the same conclusion applies to the low progress readers at the reading age developmental levels of seven and nine. Rather than strong however, the effects appear to be only moderate while yet significant.

For the high progress readers at the nine year reading age developmental level it appears that illustrations have a positive but non-significant effect on word identification accuracy and a positive, significant but only moderate effect on the meaningfulness of errors. It can be concluded that the accuracy with which the textual message is identified is effectively not influenced by illustrations at this level.

Given the significant and moderately high correlation between semantic acceptability of errors and word identification accuracy it would be reasonable to conclude that, where illustrations facilitate word identification, this is largely due to greater accessibility of semantic information such that words are more semantically predictable where illustrations are available than when they are not. Equally, where illustrations have no effect on word identification, as is the case for high progress readers at reading age nine, it may be concluded that sufficient semantic information is accessible to these readers from the text itself and that the information in illustrations is not substantively relevant to word identification.

2. Strategy

The criterion under this heading is the efficiency of message identification. The evidence for evaluating the influence of illustrations on this criterion will be drawn from the illustration effect as it appears on each of the strategies of information use as well as in a comparison of the effect across the strategies. The influence of illustrations on the strategy of self-correction will also be included in this evaluation,

i. Use of Semantic information

Interpretation of the analysis of variance results for semantic acceptability of errors indicated that the illustration effect was positive, significant and strong for

high progress readers at reading age seven as well as for low progress readers at reading age nine; and that it was significant but only moderate for low progress readers at reading age seven as well as for high progress readers at reading age nine. It can be concluded from this that illustrations do constitute a source of contextual information that compliments the available semantic information. Moreover, it can be concluded that the availability of this information encourages a word identification strategy that makes a higher use of semantic information than where the illustration information is not available. This applies to all groups but is strongly evident for the high progress readers at reading age seven and the low progress readers at reading age nine. For these groups, in particular, it would appear that the accessibility of semantic information, and therefore the development of a strategy that makes use of this information, is crucially dependent on the contextual support of illustrations. Further, the importance of this strategy within the general efficiency of message identification is suggested in the positive, significant and moderate correlations between semantic acceptability and both word identification accuracy and literal comprehension.

ii Use of syntactic information

Analysis of variance results for syntactic acceptability of errors indicated that the illustration effect was positive, significant but only moderate across combined groups.

Although illustrations might have a direct influence on the use of syntactic information where an explicit 'relational' illustration is paired with a single sentence expressing that relationship (e.g. Denberg's, 1976-77, sort of task where sentences - and illustrations - were of the order of 'The donkey pulls the wagon' : p.180), it is unlikely that illustrations accompanying continuous text, as in the present context, should have a direct influence on the use of syntactic information. However, the influence may well be indirect. The evidence suggests that this is, in fact, so. The influence of illustrations on the accessibility and use of semantic information is clear and logically explicable and, where the use of semantic information is high, it is likely that appropriate syntactic constraints will also operate. Where the use of semantic information is low, however, it does not necessarily follow that syntactic constraints will not operate. From this it might be expected that illustrations would have a moderate influence on syntactic acceptability of errors but that this would be indirect and largely attributable to the semantic influence. Apart from the generally more moderate illustration effect on syntactic acceptability of errors when contrasted with the effect on semantic acceptability of errors, support for this interpretation comes from the significant but only moderate correlation between semantic and syntactic acceptability of errors. Indirectly, it also comes from contrasting the substantial negative correlation between semantic and orthographic acceptability of errors with the minimal and non-significant correlation

between syntactic and orthographic acceptability of errors. This contrast would seem to indicate that the contextual effect (reduced reliance on orthographic information with increased contextual information) is largely attributable to the semantic factor and that the syntactic factor is only incidental within this.

The significant and moderate correlations between syntactic acceptability of errors and both word identification accuracy and literal comprehension suggest that this strategy is also important in the general efficiency of message identification. The clear non-significance of these relationships for correlations under the illustration condition, as opposed to the no-illustration condition, however, suggests that the strategy tends to be important to efficiency only where sufficient semantic information is not accessible.

iii) Use of Orthographic Information

Analysis of variance results for orthographic acceptability of errors indicated that the illustration effect was negative, significant but only moderate across combined groups. Within this there was a trend for the effect to be strong for high progress readers at reading age seven and only marginal for the other groups. From this it can be concluded that where the accessibility and use of contextual - particularly semantic - information is increased through the addition of illustrations to the text, there is an inverse reduction in the reader's reliance on orthographic infor-

mation. Although true across combined groups, the inverse effect tended to be particularly strong for high progress readers at reading age seven. Since the correlations of orthographic acceptability with word identification accuracy and literal comprehension were non-significant in the former case and negative, low but significant in the latter case the suggestion is that, in itself, a strategy that makes a high use of orthographic information is at best unimportant in the efficiency of message identification and, at worst, a negative influence.

iv) Comparison of strategies

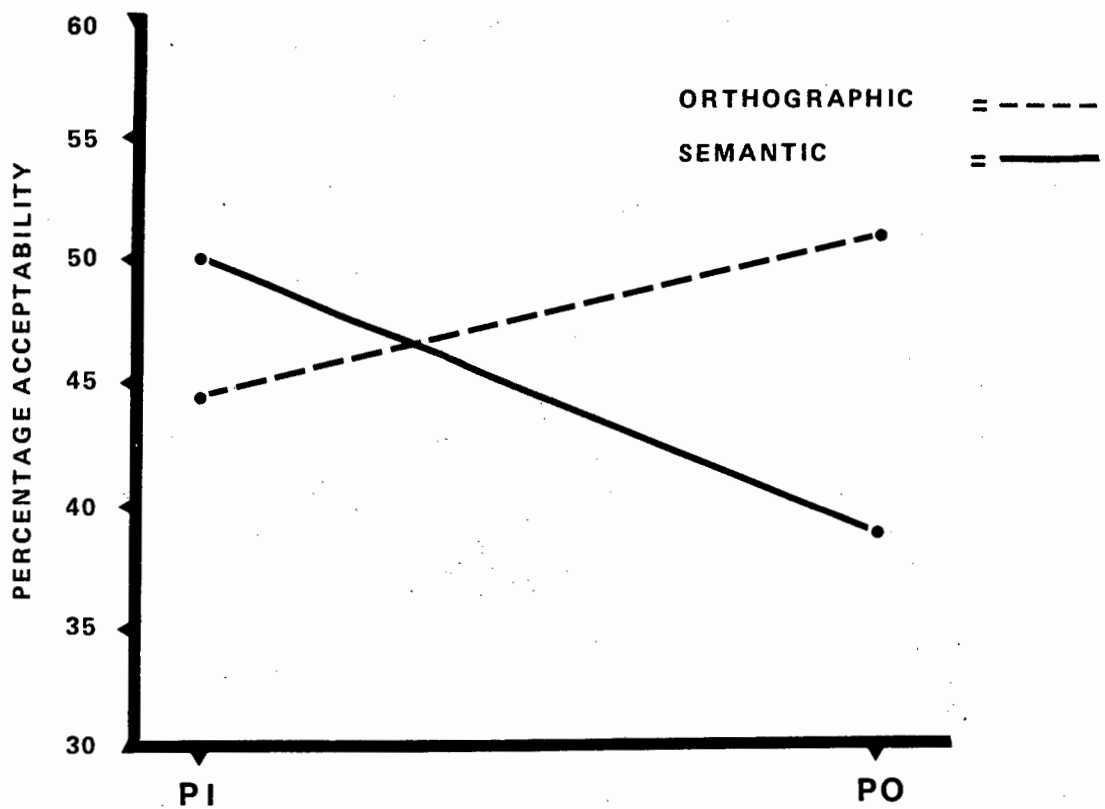
From what has been concluded so far it is clear that the influence of illustrations is most directly manifested in an increase in the accessibility and use of semantic information and that, as predicted in the contextual hypothesis, this results in a reduced reliance on orthographic information. Within this there appears to be an incidental or indirect increase in the use of syntactic information. Although the importance of each strategy in the efficiency of message identification has been suggested through correlational evidence, the relative efficiency of the balance of strategies under the influence of illustrations is crucial not only in terms of the contextual hypothesis but also for its theoretical and practical implications.

Since, under the influence of illustrations, there is an inverse relationship between use of semantic and orthographic information, it is most crucial to consider this particular

relationship in evaluating relative efficiency. Moreover, although the relationship is consistent, it is not uniform across experimental groups so that it needs to be considered as a general phenomenon as well as for each group.

FIGURE 12

GRAPH OF P1 AND PO TREATMENT MEANS FOR
ORTHOGRAPHIC AND SEMANTIC ACCEPTABILITIES



As a general pattern, the cross-over effect is clearly apparent in figure 12: where semantic acceptability is high under the illustration condition, orthographic acceptability is low; and where orthographic acceptability is high under the no-illustration condition, semantic acceptability is low. The question is, which is the more efficient

balance of strategy? Apart from evidence that use of semantic information is positively related to word identification accuracy and literal comprehension while use of orthographic information is not (or negatively so), it is also apparent that use of orthographic information reduces less from P0 to P1 than use of semantic information does from P1 to P0 (P1-P0 mean difference: orthographic, -5,98; semantic, 11,11). This is not only consistent with Donald's finding (1979a: Chapter VI), but also with Denberg's (1976-77) finding that only a partial 'trade-off' of orthographic information occurs where contextual information is made more available through illustrations. In other words, what appears to happen is that where more contextual information is available, reliance on orthographic information is reduced to a level where only that which is necessary, within the redundancy of information from other sources, is used. Finally if the absolute levels of error acceptability on the respective scales (chapter IX) are considered it is apparent that, under the illustration condition, 50% semantic acceptability means that errors are, on average, acceptable within local semantic context while 45% orthographic acceptability means that errors have, on average, at least first and second or first and last letters in common. This balance is clearly more efficient than under the no-illustration condition where 39% semantic acceptability means that errors are, on average, verging on being only lexically acceptable while 51% orthographic acceptability means that (on the five point scale) errors are, on average, only marginally more orthographically constrained

than under the illustration condition. On the general pattern, therefore, it can be concluded that the balance of strategy is more efficient under the illustration condition than under the no-illustration condition.

Since use of syntactic information is also increased, albeit indirectly, under the illustration condition this necessarily reinforces the conclusion.

For the different experimental groups, the pattern is consistent but differs in degree:

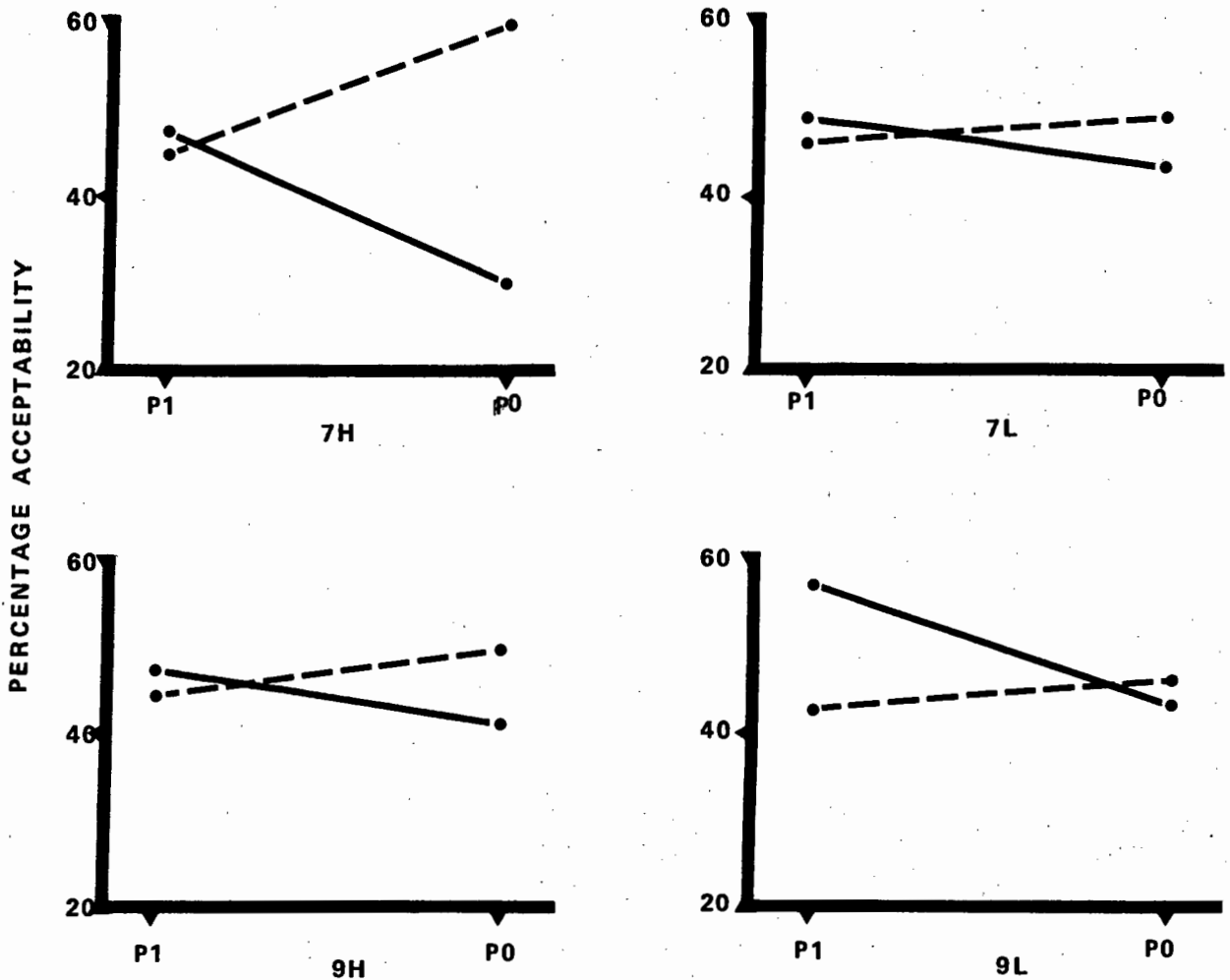
FIGURE 13

GRAPHS OF P1 AND P0 CELL MEANS FOR ORTHO-

GRAPHIC AND SEMANTIC ACCEPTABILITIES

AT 7H; 7L; 9H; and 9L

ORTH = - - - -
SEM = ————



From figure 13, it is apparent that the patterns for low progress readers at reading age seven and for high progress readers at reading age nine are essentially similar. As in the general pattern the reduction in use of orthographic information is less from P0 to P1 than the reduction in use of semantic information from P1 to P0 (P1 - P0 mean difference: orthographic; 7L, -3,02; 9H, -3,67; Semantic; 7L, 5,88; 9H, 6,55). Since the increase in use of syntactic information from P0 to P1 is similar to that for semantic information, the conclusion that applied to the general pattern therefore is equally applicable here but the difference in efficiency between the illustration and no-illustration conditions is relatively less evident.

For high progress readers at reading age seven, the difference in efficiency is clearly evident. Not only is the reduction in use of orthographic information less marked than the inverse reduction in use of semantic information (P1 - P0 mean difference: 7H orthographic, -14,33; semantic 19,71), but it is apparent that the high level of use of orthographic information under the no-illustration condition (59%) must be relatively inefficient when combined with the very low level of use of semantic information (28%): bearing the relevant scales in mind (chapter IX) this implies that there is, on average, a heavy reliance on orthographic information with very little use of textually relevant semantic information. When the level of use of syntactic information (54%) is also taken into account it appears that, on average, only local syntactic constraints to the point of the error are used. In effect, the balance of strategy implies a near word for word level of reading that, as

a process of message identification, is clearly inefficient. By contrast, under the illustration condition, the relevant levels are orthographic 45%; semantic 48%; syntactic 67%. On average, this implies a use of semantic information that is at least locally relevant in the text; a use of syntactic information that is constrained by total clause or sentence structure; and a use of orthographic information that includes at least first and second or first and last letters. As a process of message identification this is clearly a more efficient balance of strategy.

For the low progress readers at reading age nine the difference in efficiency is also apparent but the pattern is somewhat different. The reduction in use of orthographic information is, in this case, considerably less than the inverse reduction in use of semantic information (P1 - P0 mean difference: 9L: orthographic, -2,92; semantic, 12,32). Effectively there is a reasonably efficient balance of strategy under the no-illustration condition, the relevant levels being, orthographic 46%; semantic 44%; syntactic 71%. Under the illustration condition, the orthographic level stays much the same (43%); the semantic level increases substantially (56%); and the syntactic level shows a moderate increase (77%). This balance implies a high level of use of semantic and syntactic information combined with a still adequate level of orthographic constraint and may be regarded as near optimally efficient.

That this occurs for low progress readers is somewhat unexpected. However, several factors need to be borne in

mind. First, the group are not very low progress readers (see sample selection; chapter IX); second, they are chronologically the oldest and therefore the most experienced of all the groups (grade V); and third, what does emerge and is of real significance to helping readers at this level, is that the optimal balance of strategy is only achieved under illustration conditions.

From all this it may be concluded that the addition of illustrations to continuous text has a generally positive influence on the balance of strategies, increasing the use of both semantic and syntactic information and partially reducing the use of orthographic information such that the efficiency of message identification is greater under the illustration than under the no-illustration condition. This effect is particularly evident for high progress readers at reading age seven and for low progress readers at reading age nine.

v) Self-Correction

Analysis of variance results for self-corrections indicated that the illustration effect was positive, significant and strong for high progress readers at reading age seven as well as for low progress readers at reading age nine; that it was positive, significant but only moderate for low progress readers at reading age seven; and that it was positive but non-significant for high progress readers at reading age nine. From this it may be concluded that the addition of illustrations to continuous text has a generally

positive influence on the detection and (at least, overt) correction of errors but that this varies in degree across the experimental groups.

For high progress readers at reading age seven, it would appear that, without illustrations, the rate of self-correction is low but that, with illustrations, it increases to the highest for all groups. Since this is matched by a similar marked increase in the use of semantic information for this group, as well as being reflected in a general positive, significant and moderate correlation between semantic acceptability and self-correction, it seems likely that illustrations provide an enriched semantic frame of reference against which errors may be detected, evaluated and corrected. For the low progress readers at reading age seven, the same conclusion would seem to apply but this group appears to be relatively less dependent on, as well as less able to benefit from, illustrations than the high progress readers. For the high progress readers at reading age nine it would seem that self-correction, where it occurs overtly at this level, is relatively independent of illustration information. It would seem that text-based information is sufficient to provide a frame of reference for error detection and correction and illustrations make only a marginal difference in this. For low progress readers at reading age nine the rate of self-correction is the lowest for all groups under both conditions. It would seem likely that not only are self-corrections more covert at this level, but, since there is a substantial increase

in the rate of those self-corrections that are overt from the no-illustration condition to the illustration condition, that this group is also dependent on the semantic enrichment provided through illustrations and less able to rely on text-based information than the corresponding high progress readers.¹

Thus, in general, the strategy of self-correction is positively influenced by illustrations. The correlations of self-correction with word identification accuracy and literal comprehension are both positive and significant, and in the case of the former, relatively high. On Thompson's (1981) argument, the correlation with word identification accuracy could simply indicate that fewer errors make self-correction more possible. On the other hand, on Clay's (1969) evidence that high progress readers show a higher self-correction rate than low progress readers, it could be argued that self-correction is instrumental in accuracy and that high progress readers make relatively less errors at least partly because they self-correct at a higher rate. This would also produce a relatively high positive correlation. The answer is that there is probably truth in both arguments. The positive, significant although somewhat lower correlation between self-correction and literal comprehension, however, would seem to lend support to the instrumental

1. As explained in the interpretation of results, apparent anomalies in differences between high and low progress readers are explicable in terms of the grade level effect.

view. It is unlikely, although possible, that literal comprehension influences self-correction. It is more likely that the rate of self-correction has a positive influence on literal comprehension.

On this evidence, and intuitively, it can be assumed that a strategy of self-correction is important within the efficiency of message identification. Since illustrations generally facilitate the strategy this may be taken as further evidence that illustrations increase the efficiency of message identification.

3. Comprehension

The criterion under this heading is the effectiveness of comprehension. The evidence for evaluating the influence of illustrations on this criterion will be drawn from the results for literal and inferential comprehension.

i) Literal comprehension

Analysis of variance results for literal comprehension indicated that the illustration effect was positive, significant and strong for high progress readers at reading age seven as well as for low progress readers at reading age nine; that it was positive, significant but only moderate for low progress readers at reading age seven; and that it was positive but non-significant for high progress readers at reading age nine. From this it may be concluded that the addition of illustrations to continuous text has a generally positive influence on literal comprehension but

that this varies in degree across the experimental groups.

For high progress readers at reading age seven, it would appear that, without illustrations, the level of literal comprehension is very low and that, with illustrations, this increases substantially to a moderate level. Since this increase is matched by a similar marked increase in the use of semantic information for this group, as well as being reflected in a general positive, significant and moderate correlation between semantic acceptability and literal comprehension, it appears possible that illustrations increase the accessibility of semantic information and that this directly facilitates the increase in literal comprehension. More indirectly, but equally plausible, word identification accuracy is increased through the addition of illustrations and it appears likely that this is due to increased semantic predictability of words in the text. This is matched with increased semantic acceptability of errors (meaningfulness of errors) which, together with increased accuracy, indicates greater accuracy in the identification of the textual message. Where the textual message is identified with greater accuracy it follows that literal comprehension must also improve. The thread that is common to this process, however, is again, the greater accessibility of semantic information under the illustration condition. Whichever process is involved in the increase in literal comprehension, and it is likely to be a combination, it is clear that the role of illustrations in facilitating the accessibility of semantic information is central to the process.

For low progress readers at reading age seven the same conclusion would seem to apply but, again, this group appears to be relatively less able to benefit from illustrations than the corresponding high progress readers¹. For the high progress readers at reading age nine it appears that literal comprehension is relatively independent of illustration information. Effectively, it seems that it is text-based information that is used by this group in their literal comprehension and that illustrations have only a marginal positive effect. For low progress readers at reading age nine literal comprehension increases from a moderate level without illustrations to a very high level with illustrations. Given the expected grade level effect, this indicates that this group can only achieve their expected superiority in literal comprehension with the addition of illustration information and that, as with the high progress readers at reading age seven, it is the greater accessibility of semantic information that facilitates this. Since the influence of illustrations on word identification accuracy was only moderate for this group, however, it is possible that the influence of illustrations on literal comprehension is more likely to be direct. In other words, the semantic information in illustrations is directly used in literal comprehension and the effect is less via the indirect route of improved accuracy of message identification. Being low

1. The fact that the literal comprehension of this group is generally somewhat higher than for the 7H group is, again, explicable in terms of the grade level effect.

progress readers this is quite plausible since the indirect route would, in fact, be more adaptive in general reading development and would lead ultimately, as it does for high progress readers, to the relative independence of illustrations by reading age nine.

In general, therefore, it appears that a series of illustrations in relevant relation to continuous text has a positive influence on literal comprehension but that the degree, and possibly the reasons for the effect, vary across the four experimental groups included in this study. In essence this is consistent with the relative nature of other findings on the effects of illustration on comprehension (chapter V).

ii) Inferential Comprehension

Analysis of variance results for inferential comprehension indicated that there was a positive but non-significant illustration effect across combined groups. It may be concluded from this that illustrations do not effectively influence comprehension at the level of inference. This is consistent with Donald's (1979b; chapter VI) earlier finding for seven-year old average readers. It may be assumed that the reasons for the non-significant effect at this level of comprehension are in the nature of inferential comprehension which necessarily requires establishing logical relationships that go beyond the information that is explicitly given in the text or in illustrations.

The reliability of this result is, however, subject to some doubt (see chapter XII below) and it is possible that with a wider range of data on inferential comprehension, the positive tendency that was apparent across all groups could represent a significant if only moderate effect. Since inferential comprehension must involve at least a minimal base of literal comprehension this would appear quite likely. The tendency for high progress readers at reading age nine, in contradiction to all other results, to show the clearest trend towards a positive illustration effect is suggestive and could, with more reliable data, indicate the value of illustrations to this group at higher levels of processing. The positive results of Bransford and Johnson (1972) on high school students who were required to disambiguate a complex passage with or without illustration would tend to suggest that this may be so.

4. General Conclusions

Apart from the particular conclusions applicable to the individual or sub-grouped dependent variables that have been discussed, several general conclusions are of relevance in the study as a whole.

First, the study was a specific attempt to replicate the test of the contextual hypothesis under more representative and rigorous conditions than in Donald (1979a: Chapter VI). Although there were relative differences in subjects, materials, task and scoring criteria, the general pattern of results is essentially similar to that found in the

earlier study. It can therefore be concluded, with some confidence, that as a general phenomenon the illustration effect manifests as predicted in the contextual hypothesis. The only reservations to this are that, although consistently positive, the effect is generally marginal or non-significant for high progress readers at reading age nine and that the effect is non-significant for comprehension at the level of inference.

Second, in terms of the negative illustration effect predicted, although not consistently proved (chapter V) under the focal attention hypothesis, the study was in no way an attempt to prove or disprove this particular hypothesis. As pointed out in the introduction (chapter I), the task requirements for testing the two hypotheses are mutually exclusive. Nevertheless, the study was an attempt to show that, if the focal attention hypothesis is true, it is only a particular case within the more general case of reading development. Moreover, if, under the contextual hypothesis, illustrations can be shown to have a positive influence on word identification accuracy - in context; on the efficiency of word identification strategies, and on comprehension, then the predicted negative effect of illustrations on isolated word recognition could be taken as of no consequence in the wider and more representative context of learning to read and comprehend continuous text. The results of this experiment, particularly when taken together with Donald's (1979b: chapter VI) earlier results and with those of Denberg (1976-77), indicate, quite clearly, that this is the case. Furthermore, the focal

attention effect appears to be demonstrated in contextual reading in the reduced use of orthographic information under illustration conditions. However, both the import and the explanation of the effect is at variance with the focal attention hypothesis. In terms of import, reduction in use of orthographic information is seen as positively adaptive in contextual reading not only on correlational evidence, but also on the grounds that the balance of word identification strategies is improved. In terms of explanation, the evidence from this study as well as Donald (1979b: chapter VI) and Denberg (1976-77) is that there is not total attentional competition between orthographic and contextual information but that there is a partial 'trade off' of information such that where contextual information is made more accessible, reliance on orthographic information is partially reduced to a level where only that information which is necessary within the redundancy of information is attended. As pointed out earlier, that this does not occur under the test conditions of the focal attention paradigm is hardly surprising since, with isolated words, there is no redundancy of information and reliance on orthographic information must be total. It may be concluded, therefore, that even if the focal attention hypothesis is true for the learning of isolated words paired with pictures, it is not relevant in the more normal context of learning to read continuous text; and where the effect does occur in contextual reading it is neither maladaptive nor is it explicable in terms of attentional competition.

Third, the most general and logically tenable explanation for the positive illustration effect over all the dependent variables used in this study is that illustrations basically, affect the accessibility of semantic information.

Essentially, this may take two possible forms. The first is a specific and substantive cueing function where semantic information represented in an illustration has a direct and textually unprocessed cueing function on word identification or comprehension responses. The second is a more generalized and non-specific cueing function where illustrations provide a semantic 'set' or framework (Neisser, 1967) within which the meaning of the text is 'foregrounded' (Gibson and Levin, 1975) and which is fluid enough to be cumulatively integrated and modified within the textual information as it is read. The evidence in this study is insufficient to specify which form the illustration effect takes for different dependent variables or for different groups of readers. Nevertheless the pattern of results obtained is suggestive and would appear to be centrally related to this issue. It would appear most likely that the two forms are not mutually exclusive and that in most situations there is an interaction between them and that they may even be additive. Nevertheless, for high progress readers and developmentally more advanced readers the non-specific and more integrated form is more likely to manifest while for low progress readers and those at lower developmental levels the specific and more direct form is more likely to manifest (Denberg, 1976-77; Rusted and Coltheart, 1979a). In combination

information such as illustrations, then it would follow that the influence of illustrations on semantic access would be relatively reduced for both high and low progress readers at reading age nine. In combination with the pattern already suggested, this would result in a pattern that is remarkably similar to the actual obtained pattern (Table 33). It may tentatively be concluded therefore that the dominant pattern of relative influence of illustrations obtained in this study is probably explicable in these terms.

This introduces the fourth general conclusion. In terms of the tenuous and circumstantial evidence from previous research, a general interactive pattern was loosely predicted for the illustration effect over the four experimental groups (chapter VIII). Within this the respective weightings assigned were:

		READING AGE	
		7	9
PROGRESS LEVEL	H	4	1
	L	2	3

(Table 2: page 187)

In the actual results, over each of the four variables where access to semantic information may be regarded as crucial, significant $P \times Pr \times RA$ interactions were obtained. Given this, an estimate of the overall relative strength of the illustration effect for each group can be obtained from the average illustration (P) effect for each group over all

four variables. This is given in the formula:¹

$$P = \frac{\sum \left[\frac{n}{2} \frac{(P_i - P_0)^2}{MS \text{ within}} \right]_{a b c d}}{4}$$

where: a, b, c, and d represent the variables of word identification accuracy; semantic acceptability; self-corrections; and literal comprehension.

The respective P values are:

$$7H : P = 20,72$$

$$7L : P = 3,45$$

$$9H : P = 1,10$$

$$9L : P = 10,23$$

Although the results for the high progress readers at reading age nine were, on several variables, non-significant, the tendency in all cases was positive and a general although marginal effect can probably be assumed on the basis of consistency. If the relative illustration effect for the 9H group is weighted as 1 (marginal), therefore, then the equivalent weights for the other three groups, on the bases of the P values calculated above, emerge as:

¹ i.e. The formula represents, effectively, the average F value for the relevant simple-simple main P effects.

$$\begin{array}{rcl}
 7H & : & \frac{20,72}{1,10} = 19 \\
 7L & : & \frac{3,45}{1,10} = 3 \\
 (9H & : & \frac{1,10}{1,10} = 1) \\
 9L & : & \frac{10,23}{1,10} = 9
 \end{array}$$

These weights are not meant to be taken as absolute. As opposed to a totally arbitrary weighting, however, they do reflect, with some validity, the relative overall illustration effect as obtained for each group in the results of the four variables concerned.

Although the obtained weightings for the 9L group and, particularly, the 7H group are considerably higher than those estimated in the predicted pattern, the relative patterns are essentially similar with, in both cases, the illustration effect being strongest for the 7H group, weakest for the 9H group and intermediate for the 7L and 9L groups (with a relatively stronger tendency for the 9L group). It can be concluded, therefore, that the obtained pattern of the relative illustration effect over the four variables of word identification accuracy; semantic acceptability; self-corrections; and literal comprehension is consistent with the pattern predicted, albeit loosely, from previous research. Since most previous research had focussed on either accuracy or literal comprehension, the fact that the variables of syntactic

acceptability, orthographic acceptability and inferential comprehension do not follow the pattern is not inconsistent.

A further relevant factor in the obtained pattern was the influence of grade level. The truest representation of the relative obtained effects should reflect this and should therefore appear as follows: Table 33.

TABLE 33

WEIGHTINGS FOR AVERAGE OBTAINED ILLUSTRATION
EFFECTS OVER GRADE, PROGRESS AND READING
AGE DEVELOPMENTAL LEVELS FOR THE VARIABLES
OF WORD IDENTIFICATION ACCURACY; SEMANTIC
ACCEPTABILITY; SELF-CORRECTIONS; AND LITERAL
COMPREHENSION

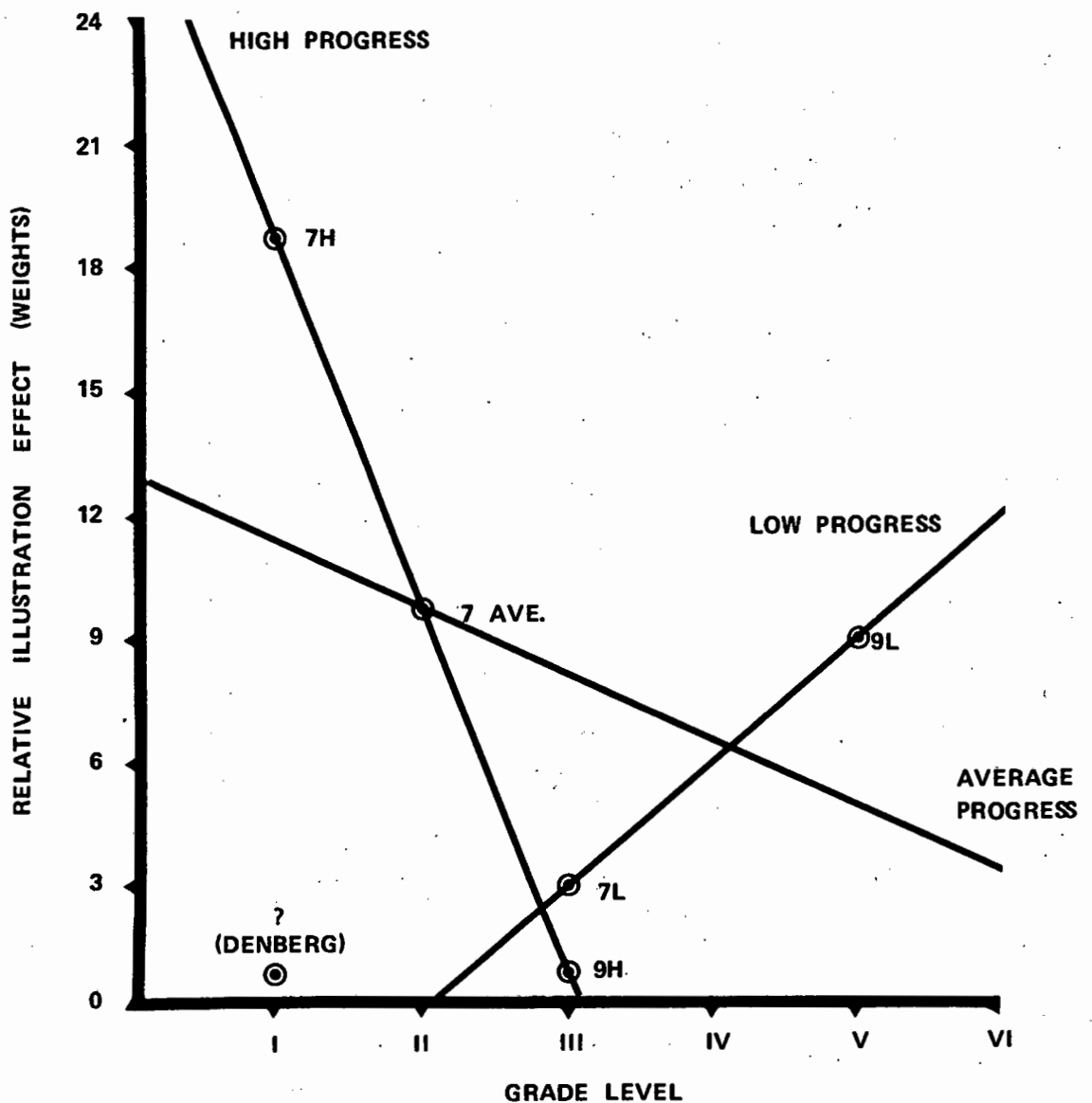
GRADE	PROGRESS	R E A D I N G A G E						
		5	6	7	8	9	10	11
I	H	-	-	19				
	AVE	-	?	-				
	L	?	-	-				
III	H			-	-	1		
	AVE			-	?	-		
	L			3	-	-		
V	H					-	-	?
	AVE					-	?	-
	L					9	-	-

The influence of grade level in the pattern of results will be discussed further under the critique (chapter XII). For the present it is important to point out that the relative obtained illustration effects apply to the specific groups indicated and that low progress readers in grade I, for instance (who would not be reading at reading age seven), or high progress readers in grade V (who would not be reading at reading age nine), are not represented in the obtained pattern. Equally all average progress groups are not represented. For groups such as these the obtained relative illustration effects would have to be extrapolated to complete the developmental picture. This is mentioned as the results are otherwise liable to be misinterpreted in terms of their applied implications.

Fifth, finally, and related to the above, when results from Donald (1979b: chapter VI) on average second year readers, and from Denberg (1976-77) on high and low 'graphemic skill' (equated with progress level for this purpose) readers in grade I are combined with results from this study, a very tentative extrapolation can be made that suggests a general developmental progression in the relative influence of illustrations on reading development for high, average and low progress readers.

FIGURE 14

EXTRAPOLATED DEVELOPMENTAL PROGRESSION
FOR THE RELATIVE INFLUENCE OF ILLUSTRATIONS
ON THE READING DEVELOPMENT OF HIGH, AVERAGE
AND LOW PROGRESS READERS



The pattern suggested in Figure 14 is that the influence of illustrations is high for high progress readers in grade I and that this declines steeply through high progress readers in grade III to effectively disappear by grade IV.

Conversely, for low progress readers in grade I the suggested evidence from Denberg (1976-77) is that the effect is clearly demarcated between children with high and low graphemic skill and that although the influence of illustrations is manifested for all readers, the effect is less marked for those with low graphemic skill. Since this is combined with the findings in this study on low progress readers in grade III and grade V a trend towards progressive increase in the influence of illustrations is apparent from a low (?) level in grade I through a moderate level in grade III to a moderately high level in grade V. For average progress readers, the evidence is slim and uncertain. The results from Donald (1979b: chapter VI) however, suggest that in grade II the relative influence of illustrations on average readers is somewhat less marked than for the high progress grade I readers and relatively more marked than for low progress grade III readers in this study¹. There is no clear evidence to extrapolate beyond the moderately high influence indicated in this but it can reasonably be expected that average progress readers would follow a trend that is intermediate between that for the high and low progress readers respectively.

Several reservations need to be placed on this extrapolation. First, the weightings on which the pattern is

¹ On the same basis as outlined above, the relative weighting for results over the four variables concerned in experiment I emerges as 10 (7 AVE : P = 11,27).

based are relative and not absolute. Alternative relative interpretations are therefore possible. This could result in somewhat different gradients for the three extrapolations but would be unlikely to alter the basic pattern. Second, as mentioned, the evidence for average progress readers is particularly tenuous and it is possible that a horizontal gradient would be more appropriate. However, the general developmental expectation is that more advanced readers become relatively less dependent on extra-textual information (chapter VII) and a gradual declining gradient is therefore more likely. Third, on the same argument, it is unlikely that the gradient for low progress readers would continue in a linear fashion to higher levels of illustration influence. What is more likely is that there would be a gradual decline at higher grade levels. Where the effect peaks is uncertain and since the development of reading for low progress readers is of special concern to most educators, it would appear important to attempt to establish this through further more focussed research. There is also some evidence (Rowher and Matz, 1975; and, particularly, Levin, 1963) that low progress 'difference' readers (essentially a social class factor) are more influenced by illustrations than low progress 'deficit' readers at the grade IV-V level. Since the low progress groups in this study had a generally lower socio-economic status than the high progress groups (chapter IX: Table 9), and since 'deficit' readers are more likely to be at the lower extreme of the normal distribution, the low progress samples selected in this study are most likely

to fall into the 'difference' category. The extrapolated tendency, therefore, may be taken as applicable to this group but it is quite possible that, for the 'deficit' group, a different tendency would apply. Finally, it is uncertain on the data available whether the gradients are in fact linear. This is the simplest and only possible form of extrapolation under the circumstances. However, it is quite likely that, apart from the expected decline in the low progress gradient beyond grade V that there is some illustration effect for low progress readers in grade I (in fact indicated in Denberg's results). This would imply a non-linear gradient for low progress readers overall.

CHAPTER XII

CRITIQUE AND RESEARCH IMPLICATIONS

1. Critique

The principal point of criticism of the experiment must necessarily relate to the grade level effect. For reasons given in chapter VIII, grade level was not incorporated as a separate independent variable in the design. As explained, only two alternatives were possible. The first and rejected alternative was to vary grade level only and to select high and low progress readers at two grade levels. This would have invalidated comparisons between high and low progress readers since the difficulty level of the material being read would have differed too much between low progress and high progress readers in any one grade. Thompson (1981b), as mentioned earlier, has criticised most research on the differences between good and poor readers specifically on these grounds. It was specifically to avoid this situation that the second alternative was selected whereby high and low progress samples were selected at a common reading age level such that the material to be read would be at the instructional level of difficulty for both groups. This, however, necessitated selecting low progress readers from a grade level two years ahead of the high progress readers. It was argued that since grade

level differences were consistent their effect could be logically accommodated in the interpretation of results. This has, in fact, been achieved. Nevertheless, it is also true that although the illustration effect could be interpreted with clarity as it applied to the various groups in themselves, the relative differences between reading age and progress level often appeared anomalous precisely because of the confounding grade level effect. Fortunately, absolute reading age and progress level effects were not, in themselves, central to the issues in this study. Where their interaction with the illustration effect was concerned they were, indeed, central but with appropriate logical adjustment these effects were also interpretable within the relative effects of illustrations. (Figure 14). Whether the alternative chosen was the most appropriate, therefore, must remain in some doubt. At least, however, it was consciously chosen and not the result of an oversight!

The second criticism applies to the measure of inferential comprehension. As opposed to literal comprehension only one question was asked at the inferential level and this was scored on a scale of 0-3. The resulting variance was therefore extremely high and the reliability of the results must remain in some doubt when compared with literal comprehension (four questions with a total possible score of 10). The reason for using only one question was again conscious and based on the practical difficulty of generating more than one reasonable inferential question from the given

material to be comparable in difficulty across three different stories. The difficulty or frustration that grade I readers were expected to experience on this level of question was also a factor and was, in fact, demonstrated in the results and in the experience of testing. Nevertheless with hindsight, it might have been more appropriate to have come to terms with the practical problems and to have included more inferential questions.

Finally, it could be argued that a repeated measures design would have been more appropriate and more powerful for testing the illustration effect. This was in fact the reason that such a design was used in experiment I (Donald, 1979a; chapter VI). However, as pointed out in chapter VIII, a repeated measures design, given the confounding factor of necessarily different stories would not have been interpretable in this study as it was in the simpler 2 x 2 factor design of the earlier study. Delayed presentation of the same story to the same subject under randomly varied first or second presentation with illustration would have been possible but the carry-over effect of meaningful stories is strong and would undoubtedly have nullified the results. The matched samples design that was finally selected, therefore, was within the circumstances, probably the most appropriate.

2. Research Implications

The research implications to be discussed comprise two reservations about research in general in this area and

three clear indications for further research.

The first reservation applies to the nature of the experimental task in which readers are involved. Where a complex cognitive act such as reading is concerned it is clear that even minor task variations may alter the nature of the experimental effects obtained. This theme has been repeatedly emphasised throughout this thesis and the clearest example of apparent contradiction in obtained effects, that are, in fact, explicable in terms of task variation, is the focal attention versus contextual controversy. Apart from this obvious example, there are many more minor variations that are equally relevant. Story length; instructions to readers and their relevant task expectations; the nature of reading material (including its difficulty, its content and its format); and behaviours expected of readers (particularly oral versus silent reading) are only some of the task variables that might affect the relative influence of illustrations or any other factor on the process of reading. Within this reservation, and where the purpose is to investigate the process of learning to read, the most relevant task to set for readers must be one which most closely approximates normal learning-to-read conditions. The attempt in this study was to do this, and results from studies where this criterion is not met must necessarily be interpreted with circumspection if the declared purpose is to draw conclusions about the normal process of learning to read.

The second reservation refers to the doubtful validity of drawing general conclusions about the influence of illustrations on reading without reference to, at least, the three central factors of developmental level, progress level and the relationship of illustration to text. In the present study, the relationship of illustration to text is defined as a series of illustrations relevantly related to continuous, narrative text. The same conclusions reached in this study may not apply, however, if this basic relationship is altered. Within the study, however, replication was built into the design on the factor of story content and it can safely be concluded that the reading of narrative stories in general, even with different contents, is subject to the influence of illustrations as found. On the factors of developmental level and progress level and their interaction with the illustration effect it is clear from this study that significant variation in the relative influence of illustrations does occur on the basis of these factors. It has also been suggested that the effect may be different within at least two separable groups of low progress readers. In general, it is clear that research and, particularly, reviews in this area (e.g., Samuels, 1970; Goldstein and Underwood, 1981) must be wary of overgeneralizing their conclusions. Not only is overgeneralization clearly invalid but there is a distinct danger that over-general conclusions, because of their simplicity appeal, will be quoted in educational texts and adopted in practice. The importance of this should become apparent in the next chapter.

The first of the three areas that clearly requires further research is the relationship of illustration to text.

The area has, in fact, received some attention (chapter V) but this has necessarily been within the specific influence of illustration on comprehension (oral or read) and has not, with a few exceptions, focussed on the early stages of reading development. Educationally, perhaps the most important variations to investigate are the influence of illustrations on the oral reading of non-fiction prose, the influence of varying ratios of illustrations to text; and the influence of heavily detailed or 'obscure perspective' illustrations (currently somewhat in vogue with publishers' artists) on, especially, early oral reading. These are factors that have considerable relevance to publishers of basal, supplementary and remedial reading materials. It is unfortunate, in fact, that so much attention has traditionally gone into the careful construction of early reading material on the basis of word frequency, phonic structure and syntactic structure but little or no attention seems to have been given to the construction of illustrations and their relationship to the text on which children learn to read.

The second area that requires further investigation is the way illustrations are used to gain access to semantic information. As pointed out in the general conclusions there appear to be two alternatives; a specific and direct cueing function, and a non-specific and integrated, 'foregrounding' function. In this study, the evidence of

relative differences in the influence of illustrations on certain variables as well on the four experimental groups appeared to be partly explicable in terms of these two semantic access functions. The interpretation was based only on logical grounds, however. Since the non-specific, integrated function must clearly be more adaptive in the process of reading development, it would be important to establish on more empirical grounds what function is operative at what level and for what reading variables. This would have not only important practical implications but would constitute an intriguing and relevant theoretical issue as regards the processing of verbal and pictorial information.

Finally, the full trend on the relative influence of illustrations for low progress readers at, particularly, higher developmental levels than those investigated here would seem to be indicated as a research priority. As pointed out earlier, the form of the trend is more important for low progress readers than for any other group and yet, ironically, it is this trend and its extrapolation beyond the grade V level, and to some extent down to the grade I level, that is most uncertain from the evidence presented in this study (chapter XI). Within this, the possible difference in the influence of illustrations on low progress 'difference' readers as opposed to low progress 'deficit' readers at various developmental levels would appear to be a relevant and unresolved issue. More specifically, the issue might be posed as the possible difference or not

between readers with low linguistic performance on the sort of language that is required in early reading material (Stubbs, 1980) and readers with particular sets of decoding disabilities such as those with poor visuo-perceptual discrimination skills (re:decoding illustrations as well as print) or those with poor auditory retention skills but adequate visual retention skills (re:comprehension of illustrated text). The permutations are many but where sources of help to low progress readers are continually being sought the dimension of illustrations would appear to be one that has been insufficiently investigated or developed.

In the South African context, in particular, the findings in this study on the relative effects of illustrations on low progress readers taken together with those of Levin (1973) and of Rowher and Matz (1975) could suggest that illustrations have a potentially important function in teaching reading, particularly beyond the very early stages, to the very large population of 'culturally different' children in this country. The real and practical problem in this context is immense. Since the implications are broad and profound, however, it is clear that further specific research is required to establish whether the suggested effect is, indeed, real for such children. Certainly a number of additional variables enter the situation (chapter IX) and, although the suggestion would appear to be logically likely it cannot be assumed without more population specific evidence.

The suggestion was also made (chapter VII) that in so far

as deaf or hard-of-hearing children experience difficulty in reading beyond, approximately, the nine year reading age level that this is largely attributable to their relative inability to make use of linguistic contextual information. On the present evidence, it may be that these children would also benefit from the extra-textual (extra-linguistic) contextual information available in illustrations. Again, the effect cannot be assumed but, given its likelihood, it would appear to be worth investigating as it may have significant implications for the teaching of reading to this particular group of handicapped readers.

CHAPTER XIII

THEORETICAL AND EDUCATIONAL IMPLICATIONS

1. Theoretical Implications

The principal theoretical issue that has been raised in this thesis refers to the process of reading and whether reading may legitimately be regarded as a constructive, synthetic process. This view of the process of reading is based on the assumption that reading involves an on-going process of prediction and confirmation of the textual message. This depends, in turn, on the availability of redundant sources of information and the reader's ability to select efficiently, within these sources of information, those cues that are most relevant to uncertainty reduction and message prediction (chapter II).

In so far as the contextual hypothesis is generated directly from this theoretical view point, the successful validation of the hypothesis in the research study offers considerable support to the theoretical view point itself. In particular, the evidence that illustrations increased the accessibility of semantic information and that this resulted in not only more accurate message identification and more effective literal comprehension but that the efficiency of message identification, as reflected in the balance of strategies, was also clearly improved constitute significant and specific

support for the theoretical view of the process involved. Put quite simply, there is no way that illustrations could positively influence the factors mentioned unless the illustration information, directly or indirectly, was effectively increasing not only the availability of redundant information but also its accessibility to the readers concerned. There must, in other words, have been improved information selection and message prediction on the basis of increased redundancy: the results are otherwise inexplicable.

Furthermore, the evidence that as use of semantic information increased, the use of orthographic information was partially and inversely decreased constituted very clear proof of the operation of redundancy and the readers' ability to predict the textual message on the basis of increased access to semantic information with less reliance on orthographic information.

In terms of the transition from initial reading to more fluent reading the developmental trends were interesting. As suggested in chapter III, the progressive acceleration model of the transitional process assumes that some use of redundancy must be operative from even the initial stages of reading if it is to involve true message identification and comprehension and not be crippled in the cognitive processing demands of a cumulative, sequential decoding process. However, as was also pointed out, the beginning reader has limited access to the linguistic constraints in written text that are necessary in the operation of redundancy. The results from this study indicate that

illustrations strongly increase the accessibility of semantic information for high progress readers in grade I and that, as a result, these beginning readers are enabled to make use of redundancy. The measure of this is that where the semantic information is relatively inaccessible (under the no illustration condition), and use of redundancy limited, the levels of word identification accuracy and literal comprehension were both substantially reduced to levels where neither could be regarded as effective. This evidence, therefore, supports the view that some use of redundancy (especially from extra-textual sources) is necessary for effective reading in the initial stages.

The progressive acceleration model also assumes that use of redundancy increases with development. This was demonstrated in the results in that a grade level effect was apparent on both semantic and syntactic acceptability of errors. This indicated that use of these sources of information, within which orthographic information is redundant, increases progressively with development and reading experience. Within this the relative independence, of high progress readers in grade III, of illustrations indicated that the need for extra-textual support in the use of redundancy declines sharply for high progress readers. For low progress readers, on the other hand, it appears that use of redundancy is initially limited and although it increases with grade level, illustrations continue and in fact, increase, in their extra-textual support function. The results could not reveal the parallel developing role

of immediate access in the acceleration model. However, the grade level effect in literal comprehension could, indirectly, be taken as evidence of a progressive development in immediate access to meaning. The declining role of mediated access could also not be directly assessed. Nevertheless, although use of orthographic information cannot be equated with mediated access it is a necessary process within it, and the inverse grade level trend for use of orthographic information would seem to be an indication that mediated access declines progressively with development and reading experience. In general, the results of the study, although not specifically designed to test the progressive acceleration model of transitional processing, would seem to lend considerable support to the model. Certainly, the alternative models would find it difficult to accommodate the dramatic redundancy effect for high progress, grade I readers under the illustration condition.

2. Educational Implications

The primary and most general educational implication is that, contrary to the general and influential conclusions reached by Samuels (1970) on the basis of the focal attention hypothesis and its paradigm, illustrations do not interfere with learning to read. The extent to which they positively and significantly facilitate reading development is indeed relative but, in line with conventional wisdom, there appears no need to deprive children of illustrations since they are motivationally relevant (chapter I), and the

result of this study and many of its predecessors outside the focal attention paradigm, do not indicate that there is any cognitive interference provided it is contextual reading, and not isolated word recognition, that is taken as the criterion of learning to read.

Beyond this most general statement, the implications are necessarily far more explicit. First, as indicated earlier, the implications of this study at least must be limited to situations in which children are reading continuous narrative text, whatever the content, and where a sequence of contextually relevant illustrations is related to that text. Since these requirements are met in a very wide range of early reading materials the implications do, in fact, have quite wide relevance. One reservation is pertinent, however, and that is that most reading material at the nine year reading age level does not have as high a ratio of illustration to text as was used in this study (chapter IX). Since several implications apply to low progress (grade V) readers at this level, it should be borne in mind that optimum illustration effects for these readers should not be expected from the, normally, rather sparsely illustrated text available at this level.

First, for high progress readers at reading age seven (last quarter of grade I) the illustration effects were, throughout, (bar inferential comprehension), significant and, indeed, quite dramatic. There can be no doubt that on all aspects of reading development evaluated in this study, illustrations have a crucially important function for

these beginning readers. Perhaps the most important implication is that it appears to be more than mere achievement that is benefitted: the balance of strategy that goes into message identification is itself substantially improved under illustrated reading. It is on the basis of this balance that efficiency of processing will be achieved and the transition to fluent reading progressively effected. On these grounds alone the importance of illustrations for these readers cannot be over-emphasised. Furthermore, since the 'foregrounding' function of illustrations would appear to be the more adaptive, it would seem worth while discussing illustrations with children before embarking on the reading and generally encouraging them to integrate this information and to predict in terms of general meaning expectations.

By reading age nine, high progress readers (grade III) appear to be relatively independent of illustration information. This need not imply that illustrations should be removed but it does imply that through grade II, high progress readers should be encouraged to minimise their reliance on illustrations and make progressively more use of text-based prediction. This does not contradict what was said for grade I high progress readers but implies that, when illustrations are obviously necessary to the reader, their use should be encouraged; while the more developed process of using text-based information should be phased in as the reader shows the ability to dispense with the extra-textual support.

For low progress readers the pattern is somewhat different. First, by extrapolation, it would appear that low progress readers in grade I are unlikely to be able to make much constructive use of illustration information. Nevertheless, in so far as this is possible, appropriate illustrations should be made available. It is also possible that readers at this level are only using illustrations on the basis of their direct, textually unprocessed cueing function (chapter XI). This may help low progress readers to get through the text and, to some extent, to extract meaning and therefore has a limited function. However, it would be more productive and adaptive if low progress readers could be helped to use illustrations on the basis of the more integrated and non-specific function. If this were actively engaged as a remedial goal there is no doubt that appropriate activities could be designed to develop this function.

What has been said for low progress readers in grade I would apply equally to low progress readers in grades III and V and in intermediate grades. The difference would be that, according to the results in this study, there would be a progressive development in the ability to benefit from illustrations together with a progressive and parallel development in the ability to use text-based linguistic information. If the specific, textually unprocessed cueing function of illustrations is allowed to dominate - as, under normal circumstances, it probably does for low progress readers throughout - then the value of illustrations

as information would be limited. However, if the non-specific, 'foregrounding' function is encouraged as a relatively more adaptive process, and is integrated with the developing ability to use text-based linguistic information, then there can be no doubt, again, that illustrations might have a very real remedial function for low progress readers.

A direct corollary to the remedial potential of illustrations for low progress readers, particularly at higher developmental levels, is the need for the development of suitably illustrated (especially in terms of ratio and relevance) remedial reading material. Recently a number of publishers have brought out remedial texts that have a relatively older content/interest level while presenting text that is structurally relatively easy to read. This is proving immensely valuable for those whose concern is to interest low progress older readers in the pleasures and possibilities of reading. Nevertheless, although such texts are usually quite reasonably illustrated one suspects that illustrations have been designed with only motivation in mind. If the more conscious goal were adopted of including a high ratio of informationally relevant illustrations considerable benefit might accrue to these low progress readers.

An important reservation on all that has been said in the educational implications so far is the fact of individual, as opposed to group, differences. The analysis of variance is, specifically, a test of whether an effect is significantly general despite individual variations. The group results can therefore be taken with the levels of confidence

actually obtained. Nevertheless, the fact that there is variation at all - and on some variables the variance was considerable - means, in educational terms, that there are individuals in the groups discussed who do not follow the general pattern demonstrated by the group as a whole. Taking one example, in practical terms, a high progress reader in grade I may well be at the more advanced stage of being relatively independent of illustrations and it would be an error to assume that because he is still in grade I he should necessarily still be reliant on illustrations. This may seem obvious but it is the logical extension of not drawing generalized or absolute conclusions about the role of illustrations in reading development and it is, therefore, important to say it.

On a different theme, one of the aims of the thesis was to justify and operationalize measures of the relative efficiency of the process of reading development as opposed to measures of absolute achievement. Measuring the strategies; use of semantic information; use of syntactic information; use of orthographic information; and the strategy of self-correction were attempts to do this. For this reason considerable attention was given to the theoretical justification for these measures as well as to the development of reliable scales for scoring them. Beyond the measures in themselves, emphasis was given to the importance of evaluating the balance of strategies as an indication of the relative efficiency of the process of message identification at any one level of development.

- and a conveniently controllable source - of extra-textual information that may facilitate access to the semantic information in the text. In a very real sense, the applied significance of the results in this study relate only incidentally to illustrations but very directly to the concept of semantic accessibility. It may be assumed, with reasonable certainty, that any other extra-textual means of increasing the reader's access to relevant semantic information would demonstrate the same relative effects as were found for illustrations. This has wide implications in the methodology of teaching reading and it is to be hoped that ultimately it is those readers, struggling with the incomprehensibilities of written text, who will benefit from this contribution.

PART C

APPENDICES

1. REFERENCES

- Arlin, M., Scott, M. and Webster, J. The effects of pictures on rate of learning sight words: a critique of the focal attention hypothesis. Reading Research Quarterly, 1978-79, 14, 645-657.
- Atkinson, R.C., Fletcher, J.D., Chetin, H.C. and Stauffer, C.M. Instruction in initial reading under computer control. Technical Report No 158, Stanford University, 1970.
- Betts, E.A. Foundations of reading instruction. New York: American Book Company, 1950.
- Bianco, R.M. The purposes of illustrations in fourth grade basal readers and their effect on children's comprehension, word recognition and interest in non-fiction passages. Unpublished Ph.D. dissertation, University of Georgia, 1980.
- Biemiller, A. The development of the use of graphic and contextual information as children learn to read. Reading Research Quarterly, 1970, 6, 75-96.
- Bransford, J.D., and Johnson, M.K. Contextual prerequisites for understanding: some investigations of comprehension and recall. Journal of Verbal Learning and Verbal Behaviour, 1972, 11, 717-726.
- Braun, C. Interest loading and modality effects on textual response acquisition. Reading Research Quarterly, 1969, 4, 428-444.

- Brown, R. Psychology and reading. In H. Levin and J. Williams (Eds.), Basic studies on reading. New York: Basic Books, 1970
- Burke, E. A developmental study of children's reading strategies. Educational Review, 1976, 29, 30-46.
- Burke, E. The probing of children's reading strategies. Educational Studies, 1977, 3, 137-143.
- Cape Education Department. The median age of pupils to Std VI on 1st March, 1977. Report of the Director of Education. Cape Education Department, 1977.
- Chall, J.S. Learning to read : the great debate. New York: McGraw Hill, 1967.
- Christenson, A. Oral reading errors of intermediate grade children at their independent, instructional and frustration reading levels. In J.A. Figurel (Ed.), Reading and realism. Proceedings of the International Reading Association, 1969, 13, 674-677.
- Clay, Marie. A syntactic analysis of reading errors. Journal of Verbal Learning and Verbal Behaviour, 1968, 7, 434-438.
- Clay, M. Reading errors and self-correction behaviour. The British Journal of Educational Psychology, 1969, 39, 47-56.
- Clay, Marie. Reading: The patterning of complex behaviour. Auckland: Heinemann Educational Books, 1972.
- Clay, M.M. and Imlach, R.H. Juncture, pitch and stress as reading behaviour variables. Journal of Verbal Learning and Verbal Behaviour, 1971, 10, 133-139.

- Clymer, T. What is 'reading'? : Some current concepts. In A. Melnik and J. Merritt (Eds.), Reading: Today and tomorrow. London: Open University Press, 1972.
- Cohen, A.S. Oral reading of first grade children taught by a code emphasis approach. Reading Research Quarterly, 1975, 10, 616-651.
- Cooper, C.R., and Petrosky, A.R. A psycholinguistic view of the fluent reading process. Journal of Reading, 1976, 20, 184-207.
- Craik, F.I.M., and Lockhart, R.S. Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 671-684.
- Dale, E. A comparison of 2 word lists. Educational Research Bulletin, 1931, 10, 484-489.
- Davis, F.B. Research in comprehension in reading. Reading Research Quarterly, 1968, 3, 499-545.
- Dechant, E., and Smith, H. Psychology in teaching reading. Englewood Cliffs, N.J.: Prentice-Hall, 1977.
- Denberg, S. The interaction of picture and print in reading instruction. Reading Research Quarterly, 1976-77, 12, 176-189.
- Donald, D.R. An analysis of the effects of contextually relevant illustrations on the learning-to-read strategies, performance and comprehension of seven year old average readers. Unpublished M.Ed. dissertation, Faculty of Education, University of Exeter, 1978.

- Donald, D.R. Effects of illustrations on early oral reading accuracy, strategies and comprehension. British Journal of Educational Psychology, 1979, 49, 282-289.
- Donald, D.R. The remedial evaluation of oral reading errors: a methodological synopsis. Reading, 1979, 13(2), 16-24.
- Donald, D.R. Analysis of children's reading errors: a current perspective. Journal of Research in Reading, 1980, 3, 106-115.
- Donald, D.R. Learning to read: a psycholinguistic analysis of transitional processing. Journal of Research in Reading, 1981, 4, 34-42.
- Downing, J. (Ed.) Comparative Reading. New York: Macmillan, 1973.
- Durrell, D.D. Durrell analysis of reading difficulty. Yonkers-on-Hudson: World Book, 1955.
- Dwyer, F.H. Exploratory studies in the effectiveness of visual illustrations. Audio-Visual Communication Review, 1970, 18, 235-249.
- Ekwall, E.E. Should repetitions be counted as errors? The Reading Teacher, 1974, 27, 365-367.
- Felzen, E., and Anisfeld, M. Semantic and phonetic relations in the false recognition of words by third and sixth grade children. Developmental Psychology, 1970, 3, 163-168.
- Ferguson, G.A. Statistical analysis in psychology and education. New York: McGraw Hill, 1959.

- Fillmore, V. The case for case. In E. Bach and R. Harms (Eds.), Universals in linguistic theory. New York: Holt, Rinehart and Winston, 1968.
- Flavell, J.H. Cognitive development. Englewood Cliffs, New Jersey: Prentice-Hall, 1977.
- Flood, J. and Lapp, D. Language /Reading instruction for the young child. New York: Macmillan, 1981
- Francis, H. Childrens' strategies in learning to read. The British Journal of Educational Psychology, 1977, 47, 117-125.
- Fromkin, V.A. Speech errors as linguistic evidence. The Hague: Mouton, 1973.
- Furth, H.G. Thinking without language. New York: The Free Press, 1966.
- Geyer, T.J. Comprehensive and partial modes related to the reading process. Reading Research Quarterly, 1972, 7 541-587.
- Gibson, E.J. Learning to read. Science, 1965, 148, 1066-1072.
- Gibson, E.J. Perceptual learning and the theory of word perception. Cognitive Psychology, 1971, 2, 351-368.
- Gibson, E.J., and Guinet, L. The perception of inflections in brief visual presentations of words. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 182-189.
- Gibson, E.J. Reading for some purpose. In J.F. Kavanagh and I.G. Mattingly (Eds.), Language by ear and by eye. Cambridge, Mass: M.I.T. Press, 1972.

Gibson, E.J. and Levin, H. The psychology of reading.

Cambridge, Mass: M.I.T. Press, 1975.

Gilliland, J. Readability. London: University of London

Press, 1972.

Ginn Reading Programme. Reading 360. Aylesbury, Bucks:

Ginn, 1978.

Goldstein, R., and Underwood, G. The influence of pictures on the derivation of meaning from children's reading materials. Journal of Research in Reading, 1981, 4, 1, 6-16.

Goodacre, E.J. Hearing children read. Reading: Centre for the Teaching of Reading (n.d.)

Goodman, K.S. A linguistic study of cues and miscues in reading. Elementary English, 1965, 42, 639-643.

Goodman, K.S. Reading: A psycholinguistic guessing game. Journal of the Reading Specialist, 1967, 6, 259-264.

Goodman, K.S. (Ed.) The psycholinguistic nature of the reading process. Detroit: Wayne State University Press, 1968.

Goodman, K.S. Analysis of oral reading miscues: applied psycholinguistics. Reading Research Quarterly, 1969, 5, 9-30.

Goodman, K.S. Psycholinguistic universals in the reading process. Journal of Typographic Research, 1970, 4, 103-110.

Goodman, K.S. Miscues: windows on the reading process.

In K.S. Goodman (Ed.), Miscue analysis: Applications to reading instruction. Urbana, Ill: National Council of Teachers of English, 1973.

Goodman, K.S. Miscue analysis: theory and reality in reading.

In J.E. Merritt (Ed.), New horizons in reading.

Newark, Delaware: International Reading Association, 1974.

Goodman, K.S., and Burke, C.L. Theoretically based studies

of patterns of miscues in oral reading performance:

Final report, U.S. Office of Education, 1973.

Goodman, Yetta. Reading diagnosis: qualitative or quantitative.

The Reading Teacher, 1972, 26, 32-37.

Goodman, Yetta and Burke, C. Reading miscue inventory kit.

New York: Macmillan, 1972.

Gurney, R. Language, learning and remedial teaching.

London: Arnold, 1976.

Guttman, J., Levin, J.R., and Pressley, M. Pictures,

partial pictures and young children's oral prose learning.

Journal of Educational Psychology, 1977, 69, 473-480.

Haring, M.J. The effect of pictures on reading comprehension.

Unpublished Ph.D. dissertation, 1978, Arizona State University.

Hartley, R.N. Effects of list types and cues on the learning

or word lists. Reading Research Quarterly, 1970, 6,

97-121.

- Harzem, P., Lee, I., and Miles, T.R. The effects of pictures on learning to read. British Journal of Educational Psychology, 1976, 46, 318-322.
- Hochberg, J. Components of literacy. In H. Levin and J. Williams (Eds.) Basic studies on reading. New York: Basic Books, 1970.
- Hood, Joyce. Qualitative analysis of oral reading errors: the interjudge reliability of scores. Reading Research Quarterly, 1976, 11, 577-598.
- Hood, J. & Kendall, J. A qualitative analysis of oral reading errors of reflective and impulsive second graders. Journal of Reading Behaviour, 1975, 7, 269-281.
- Huey, E.B. The psychology and pedagogy of reading. New York: Macmillan, 1908. (Reprinted: Cambridge, Mass: M.I.T. Press, 1968).
- Keir, G. The use of pictures as an aid to reading. Reading, 1970, 4, 6-11.
- Kellmer-Pringle, M.L., Butler, N.R., and Davie, R. 11 000 Seven-year-olds. London: Longmans, 1966.
- Ketcham, C.H., and Heath, R.W. Teaching effectiveness of sound with pictures that do not embody the material being taught. Audio-Visual Communication Review, 1962, 10, 89-93.
- Kibby, M.W. Passage readability affects the oral reading strategies of disabled readers. The Reading Teacher, 1979, 32, 4, 390-96.

- King, E., and Muehl, S. Different sensory cues as aids in beginning reading. The Reading Teacher, 1965, 19, 163-168.
- Kiraly, J., and Furlong, A. Teaching words to kindergarten children with picture, configuration, and initial sound cues in a prompting procedure. The Journal of Educational Research, 1974, 67, 295-298.
- Kolers, P.A. Reading and talking bilingually. American Journal of Psychology, 1966, 79, 377-396.
- Kolers, P.A., and Perkins, D.N. Orientation of letters and errors in their recognition. Perception and Psychophysics, 1969, 5, 265-269.
- Kolers, P.A. Three stages of reading. In H. Levin and J. Williams (Eds.) Basic studies on reading. New York: Basic Books, 1970.
- Kosslyn, S.M., Holyoak, K.J. and Huffman, C.S. A processing approach to the dual coding hypothesis. Journal of experimental Psychology: Human Learning and Memory, 1976, 2, 223-233.
- Kosslyn, S.M. and Pomerantz, J.R. Imagery, propositions, and the form of internal representations. Cognitive Psychology, 1977, 9, 52-76.
- Kramer, C.Y. A first course in methods of multivariate analysis. Virginia Polytechnic Institute and State University, 1972.

- Kulhavy, R.W. and Swenson, I. Imagery instructions and the comprehension of text. British Journal of Educational Psychology, 1975, 45, 47-51.
- La Berge, D., and Samuels, S.J. Towards a theory of automatic information processing in reading. Cognitive Psychology, 1974, 6, 293-323.
- Levin, J.R. Induced comprehension in poor readers: A test of a recent model. Journal of Educational Psychology, 1973, 65, 19-24.
- Lunzer, E., Waite, M., and Dolan, T. Comprehension and comprehension tests. In E. Lunzer and K. Gardner (Eds.) The effective use of reading. London: Heinemann Educational Books, 1979.
- Mackworth, J.F. Some models of the reading process: learners and skilled readers. Reading Research Quarterly, 1972, 7, 701-733.
- Mackworth, N.H. Verbal and pictorial comprehension by children with reading or speech disorders. Paper presented at the 20th International Congress of Psychology, Tokyo, 1972. Reported in: Gibson, E.J. and Levin, H. The Psychology of Reading, Cambridge, Mass: M.I.T. Press, 1975.
- Madden, M., and Pratt, M. An oral reading survey as a teaching aid. Elementary English Review, 1941, 18, 122-126.
- Magne, O., and Parknas, L. The learning effects of pictures. The British Journal of Educational Psychology, 1963, 33, 265-275.

- Marchbanks, G., and Levin, H. Cues by which children recognise words. Journal of Educational Psychology, 1965, 56, 57-61.
- McCullagh, Sheila. One, two, three and away. London: Hart-Davis. 1965.
- Miller, W. Reading with and without pictures. Elementary School Journal, 1938, 38, 676-682.
- Montare, A., Elman, E., and Cohen, J. Words and pictures: a test of Samuel's findings. Journal of Reading Behavior, 1977, 9, 269-285.
- Neale, M.D. Neale Analysis of Reading Ability. London: Macmillan, 1958.
- Neely, J.H. Semantic priming and retrieval from lexical memory: evidence for facilitatory and inhibitory processes. Memory and Cognition, 1976, 4, 648-654.
- Neisser, U. Cognitive psychology. Engelwood Cliffs: Prentice-Hall, 1967.
- Oller, J. Language tests at school. London: Longman, 1979.
- Ollila, L.O., and Olson, J.H. The effect on learning rate of pictures and realia in the presentation of words to kindergarteners. The Journal of Educational Research, 1972, 65, 312-314.
- O'Donnell, M., and Munro, R. Janet and John. London: Nisbet, 1950.
- Paivio, A. Mental imagery in associative learning and memory. Psychological Review, 1969, 76, 241-263.

- Rasco, R.W., Tennyson, R.D., and Boutwell, R.C. Imagery instructions and drawings in learning prose. Journal of Educational Psychology, 1975, 67, 188-192.
- Reid, J.F. Breakthrough in action: An independent evaluation of 'Breakthrough to Literacy'. London: Longman, 1974.
- Roos, W.L. The intellectually superior pupil: a background description at standard six level. South African Human Sciences Research Council, Pretoria, 1970.
- Rowher, W.D. Elaboration and learning in childhood and adolescence. In H.W. Riese (Ed.) Advances in Child Development and Behavior. (Vol 8). New York: Academic Press, 1973.
- Rowher, W.D., and Matz, R. Improving aural comprehension in white and black children: pictures versus print. Journal of Experimental Childhood Psychology, 1975, 19, 23-36.
- Ruch, M.D., and Levin, J.R. Pictorial organization versus verbal repetition of children's prose: evidence for processing differences. Audio-visual Communication Reviews, 1977, 25, 269-280
- Rusted, J., and Coltheart, V. The effect of pictures on the retention of novel words and prose passages. Journal of Experimental Child Psychology, 1979, 28, 516-524.

Rusted, J., and Coltheart, M. Facilitation of children's prose recall by the presence of pictures. Memory and Cognition, 1979, 7, 354-359.

Samuels, S.J. Attentional process in reading: the effect of pictures on the acquisition of reading responses. Journal of Educational Psychology, 1967, 58, 337-342.

Samuels, S.J. Can pictures distract students from the printed word: a rebuttal. Journal of Reading Behavior, 1977, 9, 361-364.

Samuels, S.J. Effects of pictures on learning to read, comprehension and attitudes. Review of Educational Research, 1970, 40, 397-407.

Samuels, S.J., Begy, G., and Chen, C.C. Comparison of word recognition speed and strategies of less skilled and more highly skilled readers. Reading Research Quarterly, 1976, 11, 72-87.

Schonell, F.J. Happy venture. London: Oliver and Boyd, 1939.

Siler, E.R. Syntactic and semantic constraints on oral reading. Reading Research Quarterly, 1973, 9, 583-602.

Shankweiler, D., and Liberman, I.Y. Misreading : a search for causes. In J.F. Kavanagh and I.G. Mattingly (Eds.) Language by ear and by eye. Cambridge, Mass: M.I.T. Press, 1972.

Singer, H., Samuels, S.J., and Spiroff, J. The effects of pictures and contextual conditions on learning responses to printed words. Reading Research Quarterly, 1973-74, 9, 555-567.

- Singer, H. Sight word learning with and without pictures: a critique of Arlin, Scott and Webster's research. Reading Research Quarterly, 1980, 15, 290-298.
- Smith, F. Understanding reading. New York: Holt, Rinehart & Winston, 1971.
- Smith, Frank (Ed.) Psycholinguistics and reading. New York: Holt, Rinehart & Winston, 1973.
- Smith, Frank. Comprehension and learning. New York: Holt, Rinehart & Winston, 1975.
- Smith, F. Reading. Cambridge: Cambridge University Press, 1978.
- Spache, G. A new readability formula for primary grade reading materials, Elementary School Journal, 1953, 53, 410-413.
- Spache, G.D. Diagnostic reading scales. Monterey: McGraw-Hill, 1963.
- Spearritt, D. Identification of subskills in reading comprehension by maximum likelihood factor analysis Reading Research Quarterly, 1972, 8, 92-111.
- Spooner, F.A. The evaluation of reading. London: Hodder and Stoughton, 1976.
- Stauffer, R. The language-experience approach to the teaching of reading. New York: Harper and Row, 1970
- Stubbs, M. Language and literacy: the socio-linguistics of reading and writing. London: Routledge and Kegan Paul, 1980.

- Thomson, M. A psycholinguistic analysis of reading errors made by dyslexics and normal readers. Journal of Research in Reading, 1979, 1, 7-20.
- Thompson, G. Individual differences attributed to self-correction in reading. British Journal of Education Psychology, 1981, 51, 228-229.
- Thompson, G. Towards a theoretical account of individual differences in the acquisition of reading skill. Reading Research Quarterly, 1981, in press.
- Tulving, E., and Gold, C. Stimulus information and contextual information as determinants of tachistoscopic recognition of words. Journal of Experimental Psychology, 1963, 66, 319-327.
- Van der Merwe, I.J., and Zietsman, H.C. Sosiale Atlas van die Kaapstad Metropolitaanse Gebied. Institute of Cartography, Stellenbosch University, 1977, 39.
- Vernon, M.D. The value of pictorial illustration. The British Journal of Educational Psychology, 1953, 23, 180-187.
- Vernon, M.D. The instruction of children by pictorial illustration. The British Journal of Educational Psychology, 1954, 24, 171-179.
- Weber, R. The study of oral reading errors - a survey of the literature. Reading Research Quarterly, 1968, 4, 96-119.

Weber, R. A linguistic analysis of first grade reading errors. Reading Research Quarterly, 1970, 5, 427-451.

Weber, R. First graders' use of grammatical context in reading. In H. Levin and J. Williams (Eds.) Basic studies on reading. New York: Basic Books, 1970

Webster, J. Rescue reading. Aylesbury, Bucks: Ginn, 1970

Williamson, L.E., and Young, Freda. The IRI and RMI diagnostic concepts should be synthesised. Journal of Reading Behaviour, 1974, 6, 183-194.

— Willows, D.M. A picture is not always worth a thousand words: pictures as distractors in reading. Journal of Educational Psychology, 1978, 70, 255-262.

— Willows, D.M. Individual differences in distraction by pictures in a reading situation. Journal of Educational Psychology, 1978, 70, (5), 837-847.

Willows, D.M. Reading between the lines: A study of selective attention in good and poor readers. Child Development, 1974, 45, 408-415.

Young, D. Manual for the group reading test. London: Hodder and Stoughton, 1968.

2a. STORIES AND ILLUSTRATIONS : EXPERIMENT I

The day I had a visit.

A little cat came to my house.

She went right down the garden.

She stood at the door.

Then she climbed on the top step.

I heard her cry.

She was lost.

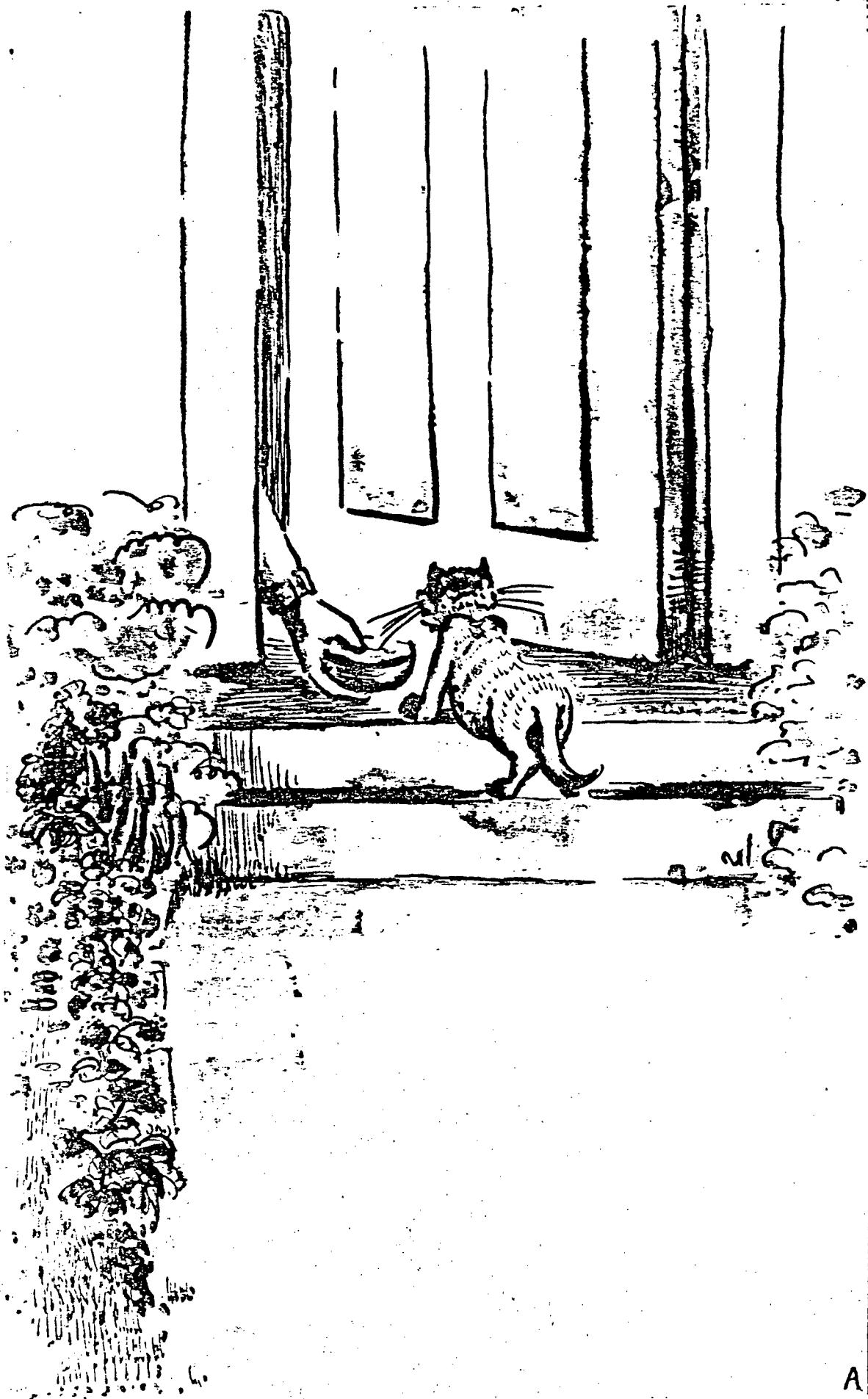
I gave her something to drink.

She felt very thirsty.

Soon she had finished it up.

She was happy again.

Now she stays with us in our home.



What can be in the box?

Father gives Jane a present.

He puts the parcel down.

Then he stands at the table.

There is pretty paper round the box.

The ribbon is in a bow.

Jane is very happy.

It is a surprise for her.

She wonders what can be inside.

She wants a dress.

Father does not tell her.

He waits until she opens it.



2b. STORIES AND ILLUSTRATIONS: EXPERIMENT II

A. STORIES FOR READING AGE SEVEN¹

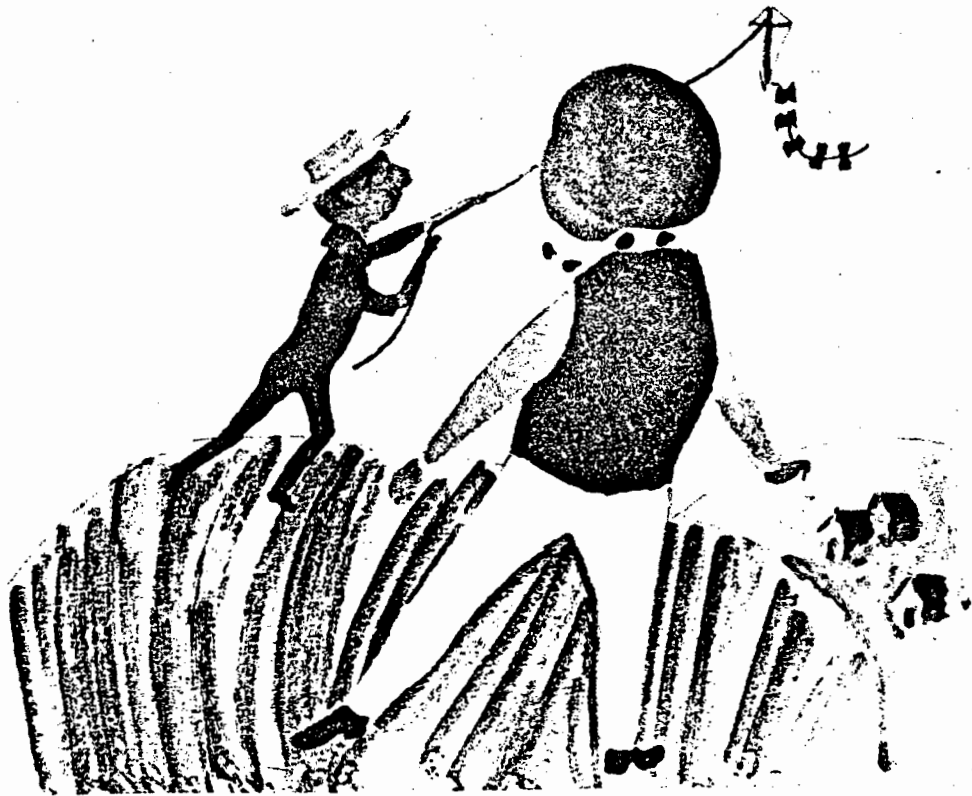
STORY I

'THE DONKEY'

¹ Stories for reading age seven are presented in their illustration format. In the no-illustration format, the sections of text appearing on each page for the illustration format were presented continuously, with each section represented in an independent paragraph. The no-illustration format is presented below for reading age nine stories. Illustrations in the actual experimental material were in colour, one example of which is given on P. 221 : chapter IX.



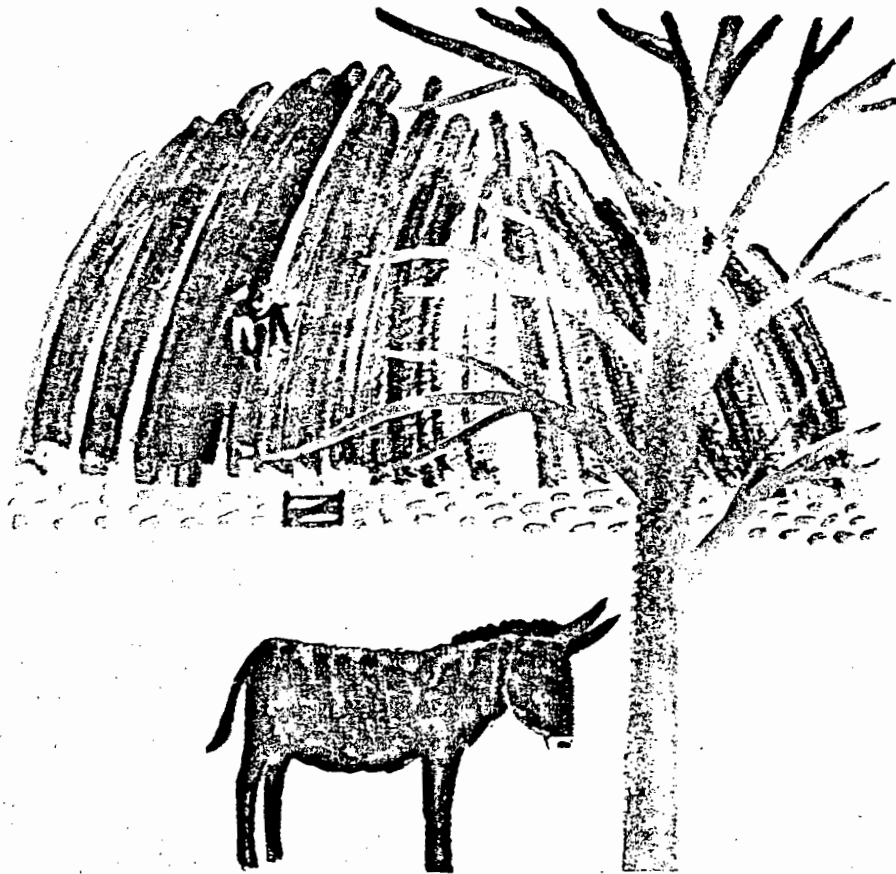
One day, Roger Red-hat and Johnny Yellow-hat had climbed up the hill near their home. Johnny had brought his kite. There was a strong wind blowing. Johnny wanted to fly his kite.



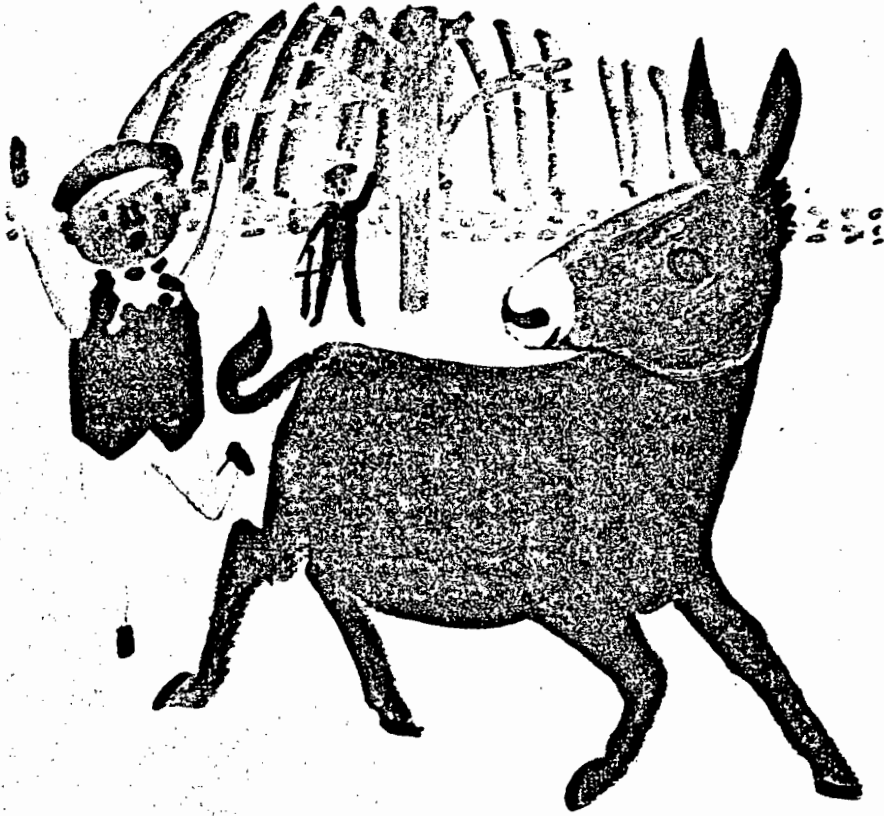
The wind blew the kite higher.
Johnny held the string and ran
after it.

"Don't let go, Johnny," cried Roger.
"The wind is too strong," called
Johnny.

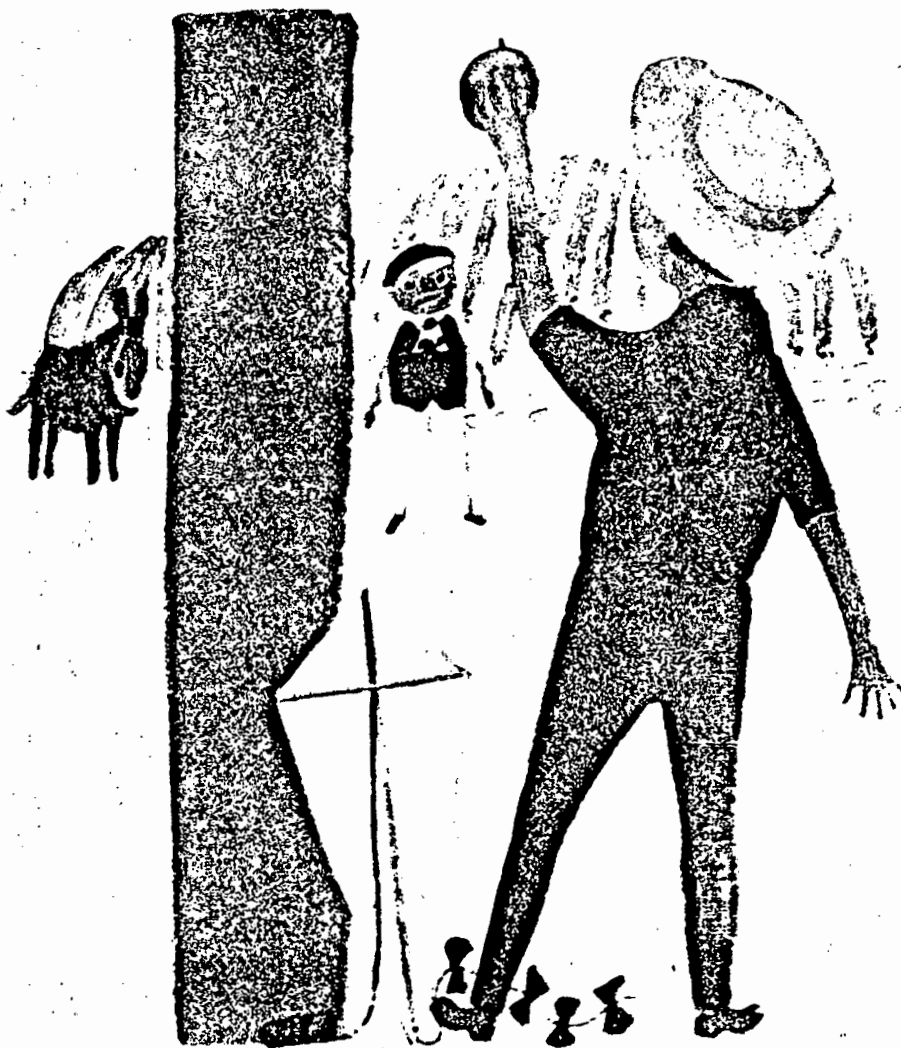
"Come and help me to pull it in."
Roger ran over to Johnny. They
pulled the kite in together.



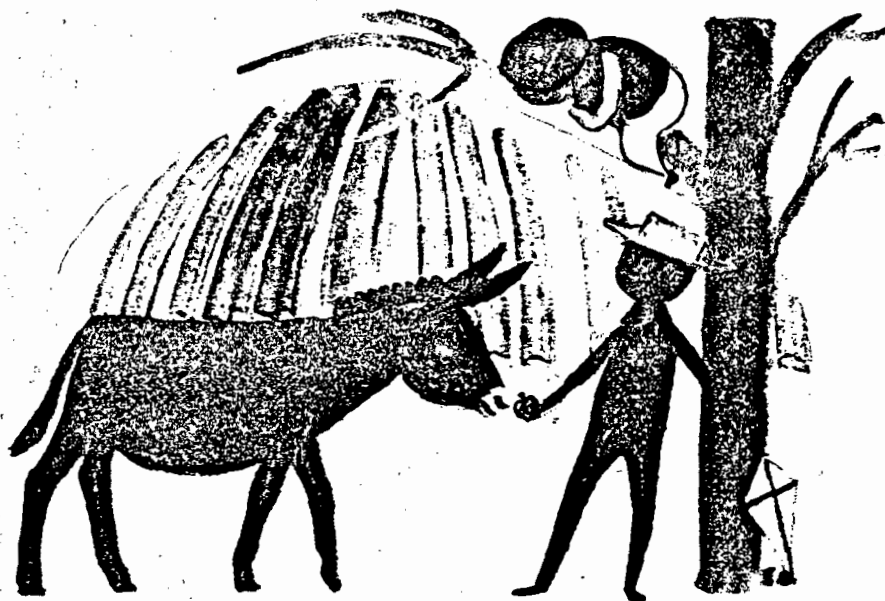
They were going back down the hill, when Roger stopped near the gate. He pointed and cried, "Look, Johnny!" There, under the tree, was a donkey. "Come on, Johnny," he said. "Let's have a ride on the donkey."



Quickly, they ran down the hill and into the field. Roger dashed over towards the donkey but the donkey saw him coming and ran off, wagging its tail. Roger waved his arms. "Stop!" he cried. But the donkey would not listen.



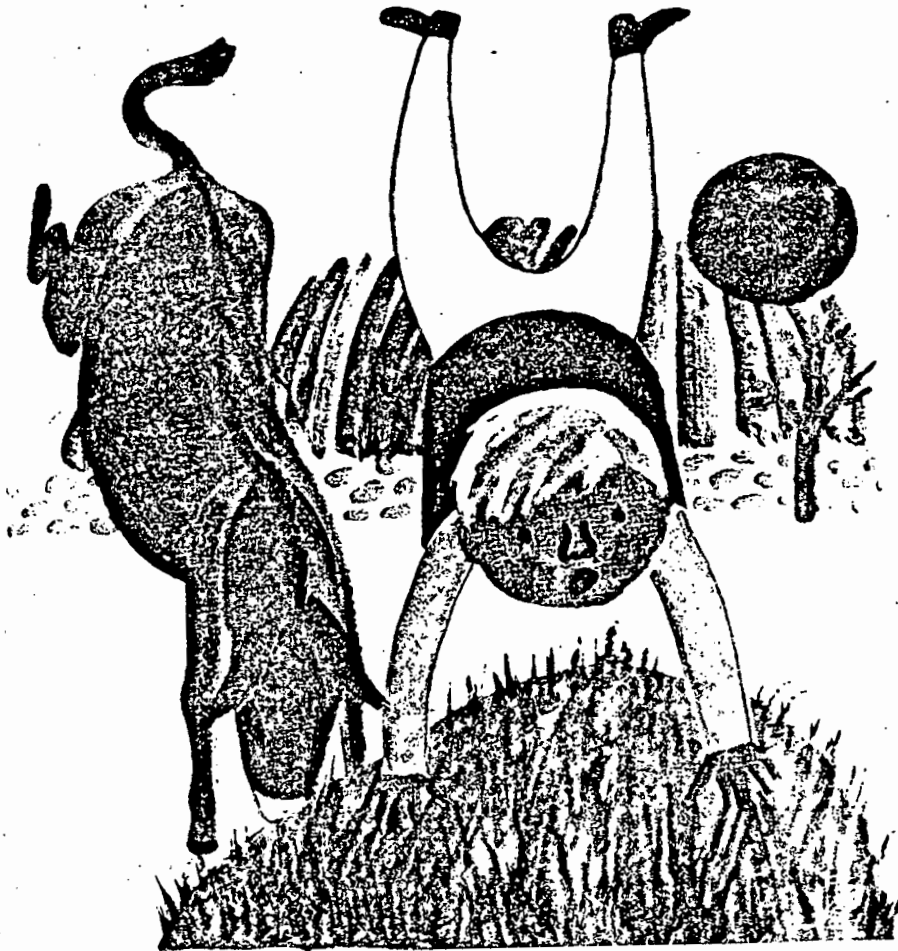
"Wait!" shouted Johnny. "Come back, Roger. I brought an apple with me. The donkey will come for an apple." Roger turned to look. "Right," he said.



"You climb up onto that branch," said Johnny. "I will hold out the apple. Then, when the donkey comes nearer, you can slip off the branch!" Roger climbed up the tree while Johnny stretched out his hand with the apple. The donkey sniffed at the apple. "Come on, donkey," said Johnny. The donkey came forward and took the apple from Johnny's hand.



Roger slipped down onto the donkey's back. Off he galloped. "Hold on tightly, Roger!" shouted Johnny.



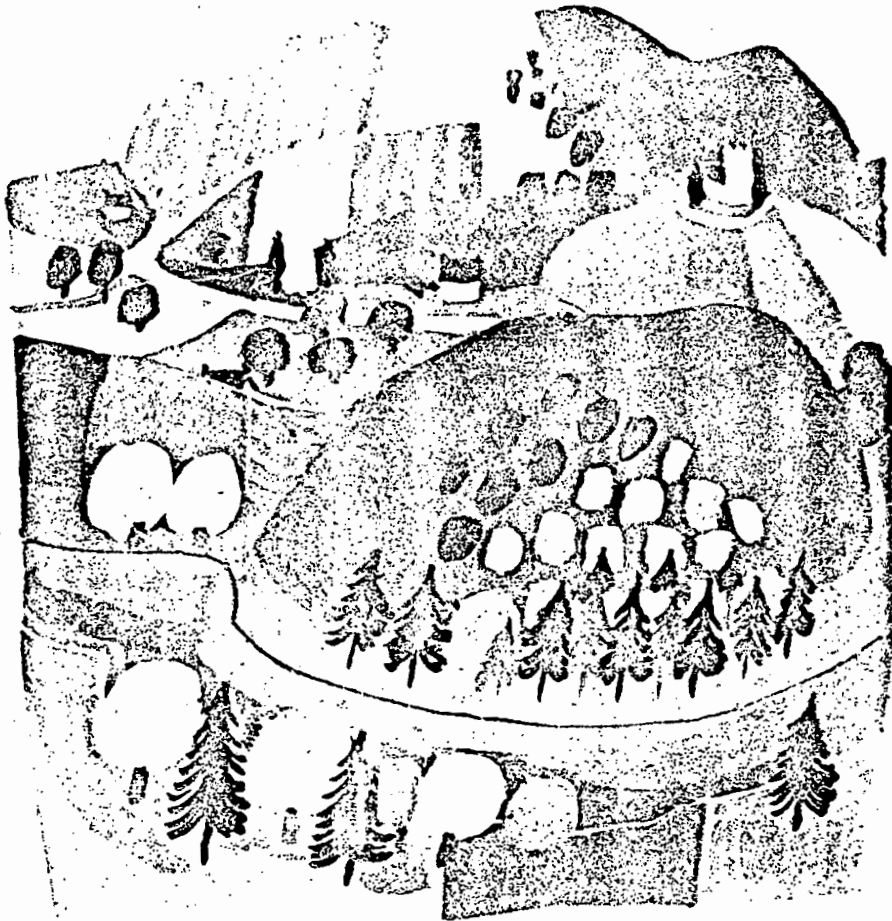
Then the donkey put his head down and kicked his hooves high in the air. Roger went flying. He landed in a haystack.



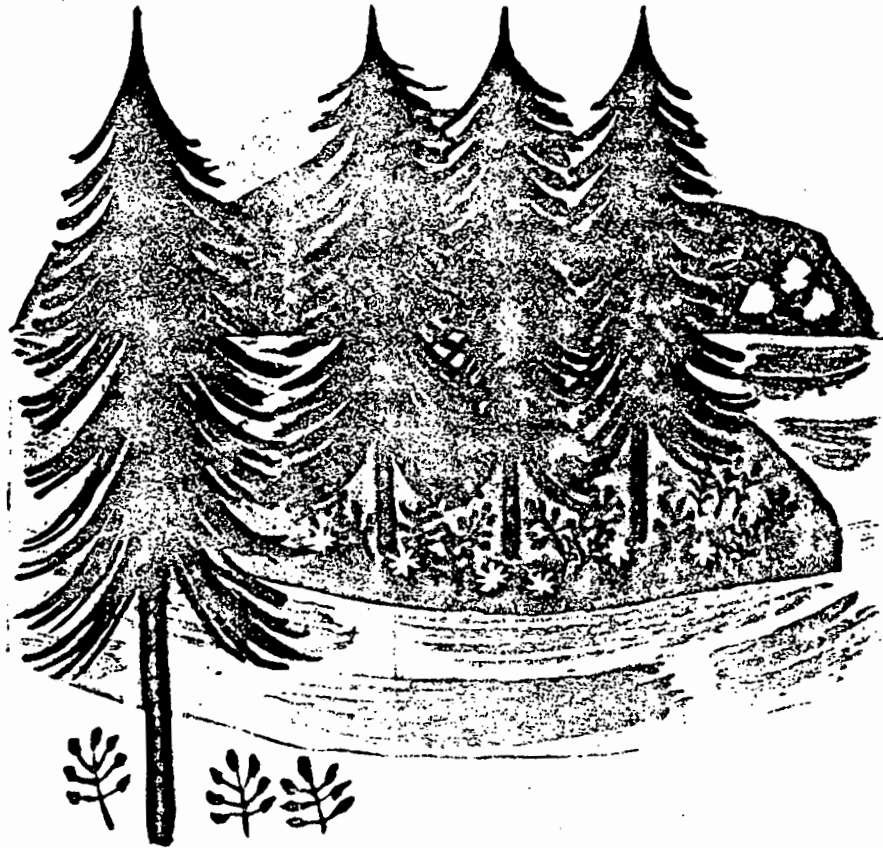
Johnny dashed across the field to Roger who was rubbing his back. "I'm not going to try that again," he said, looking angry.

STORY 2:

'THE ISLAND'



In between some rounded hills, lay
a little village. Not very far
away there was a wide, blue river.
It was called Deep River.



In the middle of Deep River there was a small island. Three large fir trees grew there. There were also many pretty flowers growing below the trees.



One day the sun was shining brightly. Johnny Yellow-hat fetched a bucket of bait and his fishing rod. He set off through the forest for Deep River.

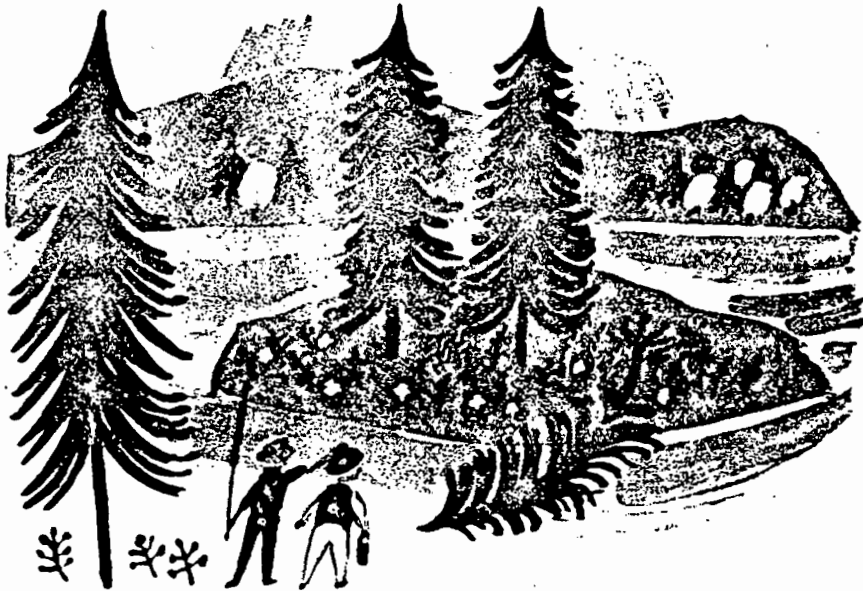


As he was walking along, he met Roger Red-hat. Roger waved a hand.

"Hallo! Where are you going to catch fish, Johnny?" he asked.

"Oh, I am going to the island," said Johnny. "There are always fish down there. They live under the bank."

"Then I will come too," said Roger.



The two boys walked along the river's edge until they came to the island. "Look!" cried Johnny. "That tree has crashed right over." One of the large fir trees had been blown over by the wind. It stretched across from the island to the bank.



"We could very easily cross over because the tree makes a bridge," said Roger. "We could reach the island. Come on!" The two friends climbed onto the tree trunk. They balanced carefully over the branches. Soon they had crossed over.



Roger started looking round for some firewood. He found some dry twigs. Soon he had a fire burning brightly. Then he looked up. Billy Blue-hat was standing on the river's edge. He was waving and calling to Roger.



Roger stood up. "What is the matter, Billy?" he shouted.

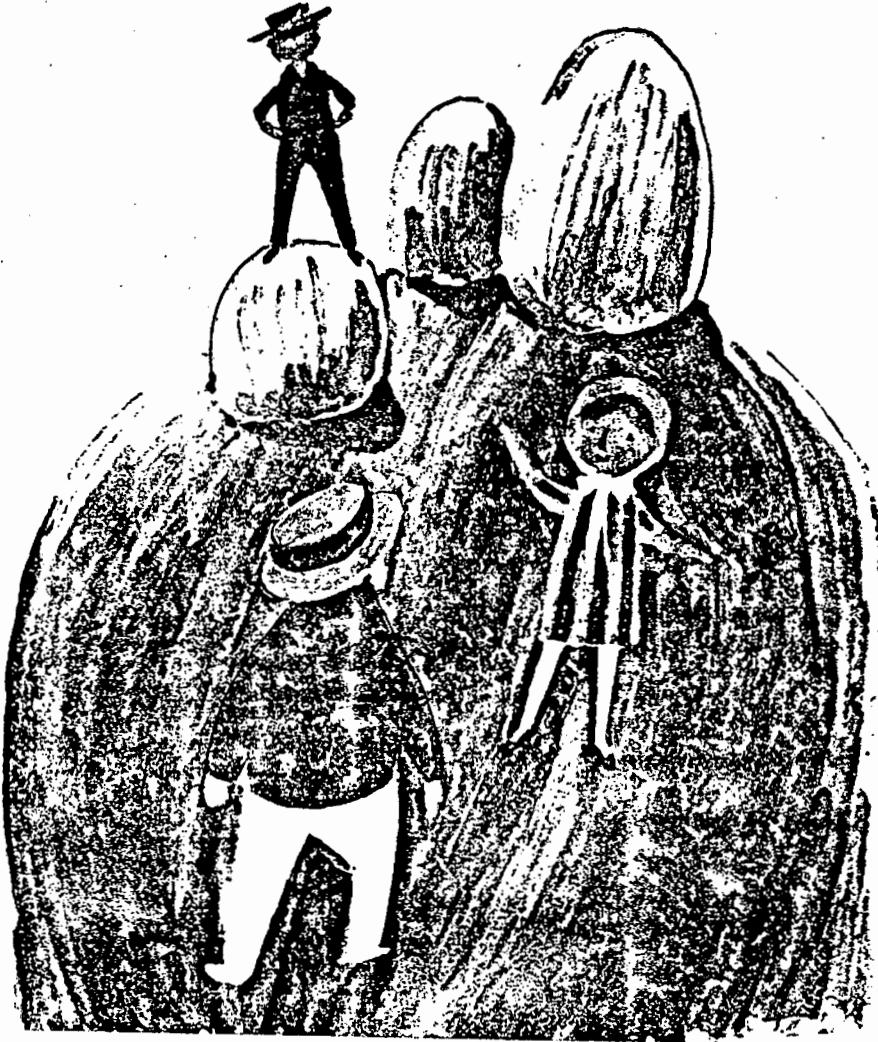
"Can I come onto the island too?" called Billy. "I also want to cross over the fallen tree trunk."

"No you can't!" cried Roger. "You will only fall into the river."

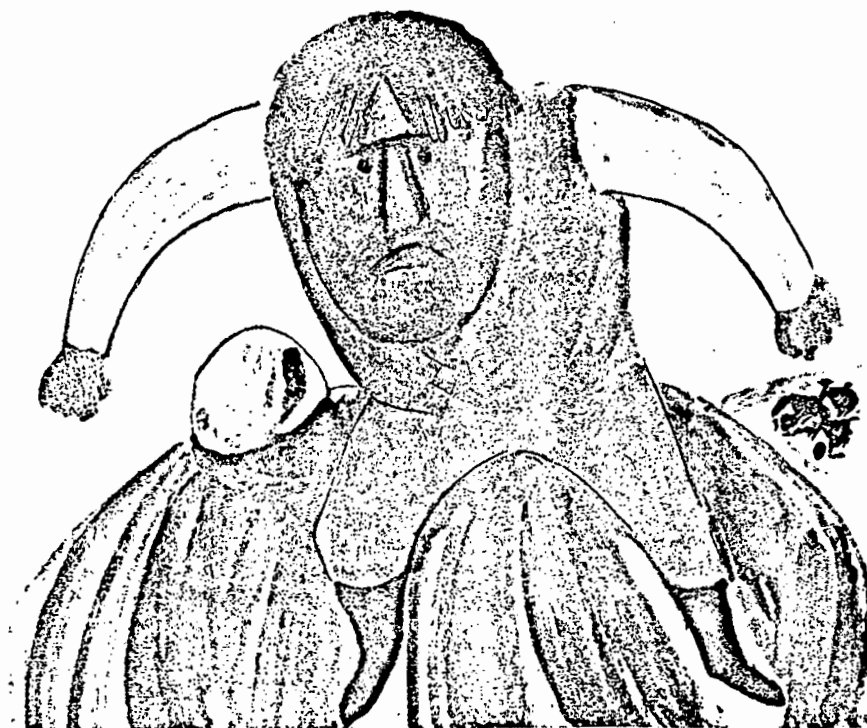
Johnny, who had been fishing, turned round. "Oh, let him try," he called to Roger.

STORY 3:

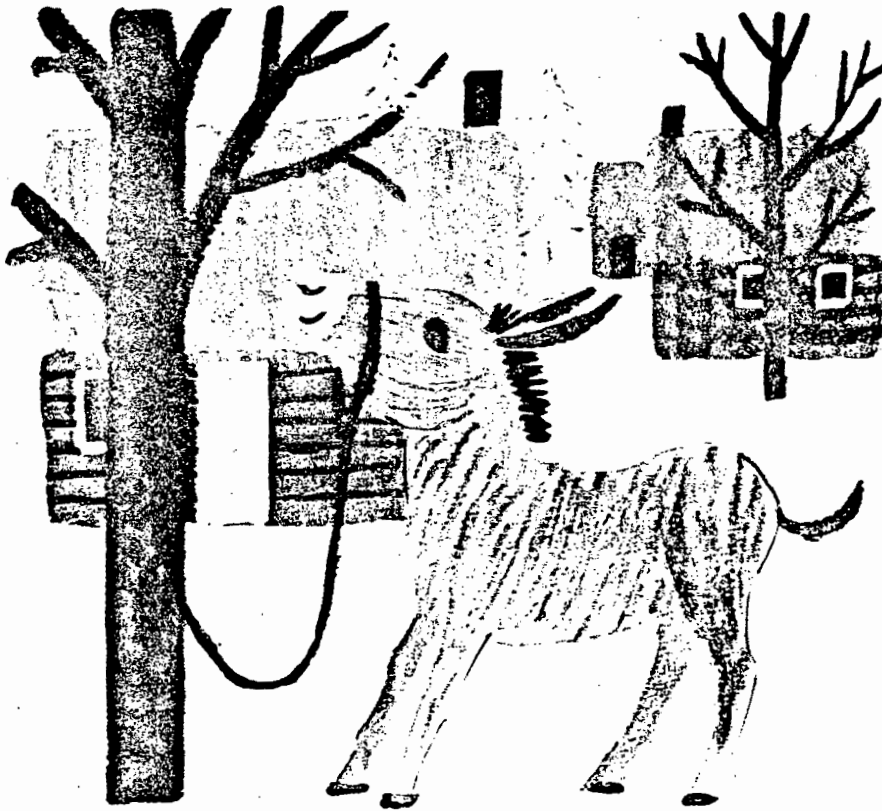
'THE GIANT'



One day Johnny and Jenny
Yellow-hat climbed a high
hill. Grandfather came too.



Long ago a very fierce giant lived on this hill. He was huge and strong. The giant was so fierce that he would not let anyone climb up his hill. He roared at them if they tried.



There was a village near the hill. A boy called Tom lived there. He owned a donkey. Tom knew that the giant liked eating donkeys. So he fastened his donkey to a tree near his cottage. One day the donkey slipped through the rope. He escaped.



When Tom saw that his donkey had escaped, he was very worried. He started to look for him. He looked all through the village. He could not find the donkey anywhere.

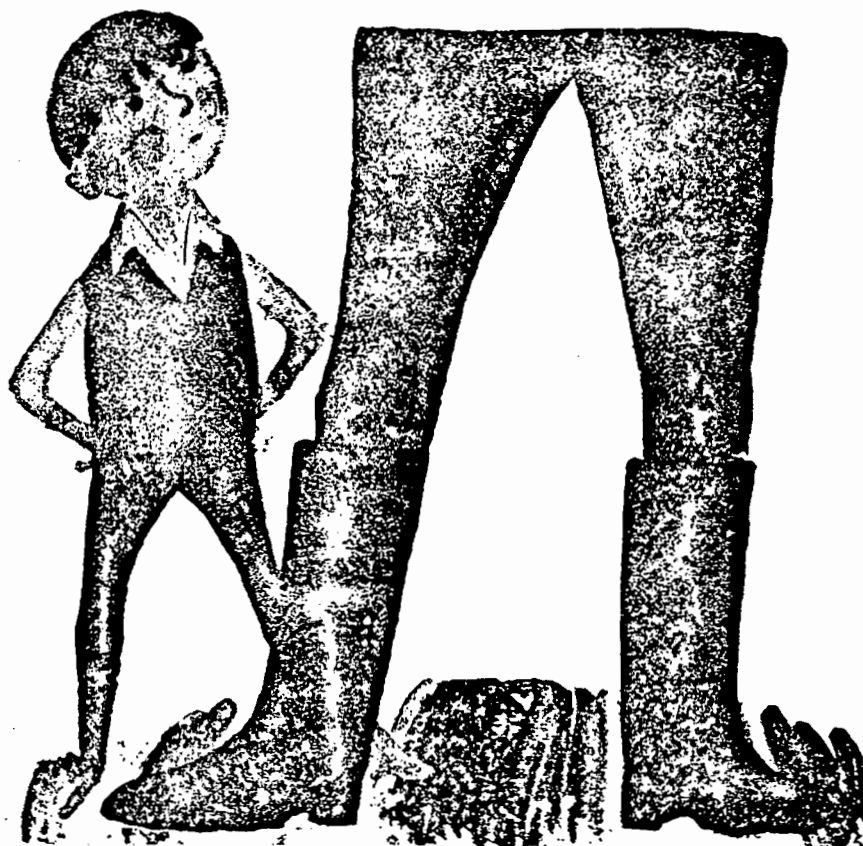


Then he started climbing the hill.
He had not climbed far, when he
saw the giant ahead of him. The
giant was leading Tom's donkey
away. He was taking him up the
hill.



Tom was so angry that he ran up to the giant. "You give me back my donkey!" he cried. He tried to snatch the rope away. The giant just laughed.

"I am going to eat him," he said.



Tom looked up at the giant.
He was very tall and strong.
But Tom was not frightened.
"Is that all you have for supper?"
he asked. "Only one little donkey?"
"You cannot be a very strong giant."



The giant was very angry. He lifted his fist. "I am strong," he roared. "I am fierce too. I shall eat you both for supper!"

B. STORIES FOR READING AGE NINE¹

STORY 1:

'THE DONKEY'

¹ Stories for reading age nine are presented in their no-illustration format. In the illustration format, the same illustrations as appear for reading age seven were used. Sections of text which appear in paragraphs were presented, in the illustration format, on separate pages below the relevant illustration as above.

In order to attempt flying their kite Roger Red-hat and Johnny Yellow-hat decided one bright but blustery morning, to climb the ridge outside their village. On the summit of the ridge the wind was blowing really strongly.

As the kite was swept higher, Johnny felt the string slipping rapidly through his grasp. Over the ridge he raced, chasing the kite. "Hold on tightly," shouted Roger. "Shall I help?" Roger joined Johnny and eventually they managed to pull it in together.

Half-way down the hillside, when they had almost reached the boundary wall surrounding the paddock at the bottom, Roger suddenly stopped.

"There's a new donkey!" he exclaimed, pointing excitedly. "He's dozing quietly over there underneath that tree."

They both rushed into the paddock. Roger, who was most eager to catch the donkey to ride him, dashed ahead waving his arms and yelling with excitement. However, becoming frightened with all the commotion, the animal galloped away, whisking its tail cheekily.

"Wait!" shouted Johnny. "Let's not terrify the poor donkey. How about tempting him nearer with this juicy apple which I brought along?"

Presently the donkey decided that he had had enough. Plunging his head forward and flicking his hooves high, he bucked his rider flying into a spikey mound of straw.

Johnny hurried across to his friend. Roger was looking distinctly annoyed and clutching a bruised leg.

"That was certainly no fun," he grumbled.

STORY 2:

'THE ISLAND'

In a sheltered valley amongst gently rolling hills nestled a small country village. A short distance away, a wide river swept swiftly past.

In the centre of the rapidly flowing current an island was situated on which three enormous fir trees had grown. They clustered close together, near the shore.

One bright, summer's day Johnny Yellow-hat decided to go fishing. Gathering together his fishing equipment and a container of bait, he set off eagerly through the forest.

As he was strolling along Johnny happened to meet Roger Red-hat.

"Hallo," Roger greeted his friend, raising a hand.

"Where are you intending to fish?" he enquired brightly.

"Alongside the island," replied Johnny.

"They usually feed beside the shore there," he explained.

"That sounds interesting," said Roger. "I'll keep you company."

As the two friends approached the river's edge, they suddenly stopped, staring in amazement. "Look there!" exclaimed Johnny, pointing with excitement.

One of the enormous fir trees had been completely uprooted so that, having collapsed right across the narrow channel between the island and the shore, it made a perfect bridge.

Without hesitation, the boys decided to cross. It was too good an opportunity to miss. Balancing carefully with their arms outstretched, they stepped boldly along the tree trunk. Although the river rushed past threateningly underneath, they weren't frightened and presently they had arrived safely on the island.

Roger immediately set about collecting some firewood while Johnny settled down to fish. No sooner had he started a campfire brightly blazing away, than he happened to glance up.

To his surprise, he noticed Billy Blue-hat shouting and waving frantically from the shore. However, he couldn't hear distinctly because of the river's noise.

"What's wrong, Billy?" yelled Roger, straightening up to make himself heard more clearly. Billy moved closer to the tree, indicating that he wanted to cross. "No!" bellowed Roger, shaking his arms excitedly, "you'll never manage!" Johnny, who had been busily occupied with his fishing overheard this. Turning around, he suggested to Roger that they let Billy attempt the crossing.

STORY 3

'THE GIANT'

One morning Jennifer, Johnny and Grandfather Yellow-hat had all trudged wearily up a high ridge. Three strange, grey boulders stood on the summit of the ridge.

"How did these three strange boulders come to be here?" enquired Jennifer.

Grandfather slowly lowered himself onto the smallest boulder. Filling his pipe, he stared quietly into the distance. The children, guessing that a story was coming, settled themselves down comfortably.

"It happened many years ago," began Grandfather.

"Only a solitary boulder stood on this ridge then.

A cruel and fearsome giant commanded the ridge and surrounding country and wouldn't allow anyone to climb his ridge.

In a neighbouring village lived a boy named Tom who owned a donkey. In order to prevent his donkey from wandering into the cruel clutches of the giant, Tom tethered him to a convenient tree outside his cottage. However, the donkey eventually managed to loosen the rope and escape.

When Tom discovered that his donkey had escaped, he became desperately concerned. He searched throughout the village. He frantically hunted everywhere.

Finally he decided to risk climbing the hill. Presently, as he climbed, he noticed the enormous figure of the giant striding out ahead of him. There he was leading away Tom's favourite friend. Tom decided he must be rescued.

Without hesitation Tom boldly approached the giant and demanded the return of his donkey. He was so furious that he forgot to be frightened. "You have no right to be stealing my donkey!" he exclaimed trying to snatch the rope away. The giant only roared with laughter.

Although the giant's enormous figure towered threateningly over him, Tom stood his ground firmly glaring angrily up at his opponent. "I don't believe in your strength anyway," he challenged.

"Surely a really strong giant could manage to eat more than one miserable little donkey!"

At this insulting challenge from such a little fellow, the giant became thoroughly enraged. Raising his fist menacingly, he bellowed, "I'm the strongest and fiercest of all giants. I'll show you!"

APPENDIX 3COMPREHENSION QUESTIONS AND EXAMPLES OFANSWERS: ALL STORIES, ALL LEVELS.A. READING AGE SEVEN1. THE DONKEY

DETAIL: Q: Where was the donkey standing when the boys first saw him?

TEST-BASED ANSWER (T):

Under a tree, in a field.

E.G.s

2: 'In a field, by a tree'

1: 'Near a tree'

0: 'In a stable'

CAUSE- Q: Why did the donkey run away from Roger?

EFFECT

T: Because he dashed over towards the donkey: he waved his arms: he cried, 'Stop!'

E.G.s

2: 'Because he shouted and waved at him.'

1: 'Because he got a fright.'

0: 'Because he was playing'

MAIN Q: How did Roger manage to get onto the donkey to ride him?

IDEA:

T: Johnny held out an apple. Roger

climbed up the tree (or, onto the branch).
 the donkey came nearer (or, took the
 apple from Johnny's hand). Roger
 slipped down onto the donkey's back.

E.G's

- 3: 'Johnny held an apple. Then Roger
 climbed a tree and jumped onto the don-
 key's back when he came over.'
- 2: 'The donkey came for the apple and
 Roger jumped on his back.'
- 1: 'He jumped off a branch.'
- 0: 'He caught the donkey.'

RELATED
 IDEAS:

- Q: What happened in the first part of
 the story?
- T: Roger and Johnny had climbed a hill.
 Johnny was flying his kite. The wind
 blew the kite higher. Johnny had to
 run after it. He asked Roger to help
 pull it in. They pulled it in together.

E.G's

- 3: They were flying their kite on the top
 of a hill and the wind blew it so they
 had to chase it.'
- 2: 'The wind blew their kite and they were
 chasing after it.'
- 1: 'They were flying a kite.'
- 0: 'They saw the donkey.'

INFER-
ENCE:

Q: Who do you think is the cleverest of
the two boys?

T: Johnny. Because it was he who thought
of the idea of holding out the apple so
that the donkey would come nearer and
Roger could jump from the tree onto his
back.

E.G's

3: 'Johnny. Because he held out the apple
so that Roger could get on the donkey.
He thought of it.'

2: 'Johnny. Because he held out an apple
to the donkey.'

1: 'Johnny.'

0: 'The boy who rode the donkey.'

2. THE ISLAND

DETAIL: Q: What was Johnny carrying with him?

T: A bucket (of bait) and his fishing rod.

E.G's

2: 'A bucket and a fishing rod.'

1: 'His fishing things.'

0: 'A bag.'

CAUSE-
EFFECT:

Q: What made Roger notice Billy at first?

T: He was waving and calling to Roger.

E.G's

2: 'He was shouting at him.'

1: 'He was on the other side of the river.'

0: 'He saw him.'

MAIN Q: How did the boys manage to reach the
IDEA: island?

T: A large fir tree had been blown over
by the wind. It made a bridge from
the island to the bank. They climbed
across the tree trunk.

E.G's

3: 'They climbed across a tree that the wind
had blown down. It made a sort of
bridge for them.'

2: 'They crossed the tree that had fallen
over.'

1: 'By climbing on the tree trunk.'

0: 'By crossing over the river.'

RELATED Q: What sort of place did the boys live in?

IDEAS: T: A village in between rounded hills.
A river not far away. Trees. Flowers.
A forest.

E.G's

3: 'A village in valley with trees and
things around.'

2: 'Where there were lots of trees and
a river.'

1: 'In a forest.'

0: 'In a house.'

INFER- Q: Who do you think is the friendliest of
ENCE:

Roger and Johnny? Why?

T: Johnny. Because he wanted to let
Billy cross over (while Roger did not).

E.G's

3: 'Johnny. Because he said to Roger,
'Let Billy come across.'

2: 'Johnny. Because he let Roger come
with him to his special fishing place.'

1: 'Johnny.'

0: 'Roger. Because he asked to come too.'

3. THE GIANT

DETAIL: Q: Where did Tom keep his donkey?

T: Fastened to tree, near his cottage.

E.G's

2: 'Tied on a tree outside his house.'

1: 'Tied up to a tree.'

0: 'In his cottage.'

CAUSE- Q: How did the donkey escape,

EFFECT: T: He slipped through the rope.

E.G's

2: 'He pulled his head through the rope.'

1: 'He pulled away from the tree.'

0: 'He ran away.'

MAIN Q: What did Tom do when he found that his
IDEA: donkey has escaped?

T: He was worried so he started to look
for him. He looked all through the
village. He started climbing the
giant's hill. He saw the giant. He
told the giant to give his donkey back.

E.G's

3: 'He started to look all over for him.
Then he climbed the hill and saw the
giant had him.'

2: 'He started to look for him and then he
went to the giant.'

1: 'He looked for him.'

0: 'He called him.'

RELATED Q: What made Tom so worried when his donkey
IDEAS: escaped?

T: There was a giant. He was very fierce.
They knew that he liked eating donkeys.
The giant might catch the donkey. Tom
liked his donkey.

E.G's

3: 'Because he knew about the fierce giant
who liked eating donkeys - he might eat
his donkey.'

2: 'Because there was a giant and he might
catch his donkey.'

1: 'He liked his donkey and didn't want to
lose it.'

0: 'He couldn't find it.'

INFER- Q: What made Tom so brave with the giant?

ENCE: T: He was so angry and he liked his donkey
so much that he wasn't frightened of him.

E.G's

3: 'He wanted his donkey back so badly
that he wasn't frightened.'

2: 'He just thought of his donkey.'

1: 'He was cross.'

0: 'Because he was a boy' (Answer given
by a girl!)

B. READING AGE NINE

1. THE DONKEY

DETAIL: Q. What was the donkey doing when the boys
first saw him?

T: He was dozing quietly underneath a tree.

E.G's

2: 'He was sleeping under a tree.'

1: 'Standing under a tree.'

0: 'Eating.'

CAUSE- Q: Why did Roger not succeed in catching
EFFECT: the donkey at first?

T: Because Roger dashed ahead waving his
arms and yelling with excitement. The
donkey became frightened with the commotion,
galloped away.

E.G's

2: 'Because he frightened the donkey with his running and shouting.'

1: 'Because he frightened him.'

0: 'The donkey was too clever.'

MAIN
IDEA:

Q: How did the boys eventually manage to ride the donkey?

T: Johnny offered an apple to the donkey while Roger scrambled up onto an overhanging branch. The donkey approached the bait, sniffing at Johnny's outstretched hand. Roger slipped quickly off the branch and landed on the donkey's back.

E.G's

3: 'Roger climbed onto a branch and Johnny offered the donkey an apple. When he came nearer, Roger jumped down.'

2: 'Johnny tempted the donkey with an apple and then Roger jumped down onto his back.'

1: 'They offered the donkey an apple.'

0: 'By catching him.'

RELATED
IDEAS:

Q: The story had two parts to it. What was the first part?

T: Roger and Johnny had climbed a ridge outside their village. They were trying to fly their kite. The wind swept the

kite higher and Johnny had to chase after it. Roger joined him and they managed to pull it in together.

E.G's

3: 'They were flying their kite on a windy hill. They kite nearly blew away and they had to run after it.'

2: 'They'd climbed up a hill to fly their kite'

1: 'They were playing with a kite.'

0: 'About the two boys.'

INFER-
ENCE:

Q: Which of the two boys do you think is the cleverest? Why?

T: Johnny. Because it was he who suggested that they stop terrifying the donkey and tempt him nearer with an apple so that Roger could slip off the branch and onto his back.

E.G's

3: 'Johnny. Because he was the one who persuaded the donkey under the tree with an apple. Roger only chased the donkey.'

2: 'Johnny. Because he thought of the plan.'

1: 'Johnny.'

0: 'Roger. He rode the donkey.'

2. THE ISLAND

DETAIL: Q: What did Johnny gather together to take with him?

T: His fishing equipment and a container of bait.

E.G.'s

2: 'His fishing things and a bucket of bait.'

1: 'A tin of bait.'

0: 'Some dry twigs.'

CAUSE Q: How did Roger first come to notice Billy?

EFFECT: T: He glanced up from the campfire.
Billy was shouting and waving frantically from the shore.

E.G's

2: 'He looked up and saw him shouting and waving.'

1: 'He glanced up.'

0: 'He was in the forest.'

MAIN Q: How was it that the boys managed

IDEA: to reach the island?

T: An enormous fir tree had collapsed right across the narrow channel between the island and the shore. It made a perfect bridge. Balancing carefully, the boys crossed over the tree trunk.

E.G's

- 3: 'A big fir tree had fallen over and made a bridge so that boys could cross over to the island.'
- 2: 'They walked across a tree trunk that made a bridge.'
- 1: 'A big fir tree had been blown over.'
- 0: 'Across the river.'

RELATED
IDEAS:

- Q: Can you describe the sort of place that the boys lived in?
- T: A small country village. in a sheltered valley; amongst gently rolling hills; a river a short distance away; a forest, fir trees.

E.G's

- 3: 'A sort of little village with trees and hills around.'
- 2: 'In a forest with a river.'
- 1: 'In the country.'
- 0: 'In a town.'

INFER-
ENCE:

- Q: Which, of Roger and Johnny, do you think is the kindest? Why?
- T: Johnny. Because he suggested to Roger that they let Billy attempt the crossing while Roger had said, 'No' to Billy.

E.G's

- 3: 'Johnny. Because he wanted to let Billy come across but Roger didn't.'

2: 'Johnny. Because he was nicer to Billy.'
 1: 'Johnny.'
 0: 'Roger.'

3. THE GIANT

DETAIL: Q: Where did Tom tether his donkey?
 T: To a convenient tree outside his cottage.

E.G's

2: 'To a tree near his house.'
 1: 'To a tree.'
 0: 'With a rope.'

CAUSE- Q: How did the donkey manage to escape?
 EFFECT: T: He loosened the rope.'

E.G's

2: 'He got the rope loose.'
 1: 'He got away from the rope.'
 0: 'He ran away in the night.'

MAIN IDEA: Q: What did Tom do when he discovered
 that his donkey had escaped?
 T: He became desperately concerned and he
 searched throughout the village; he
 hunted everywhere. Finally he decided
 to risk climbing the giant's hill.
 He noticed the giant leading away his
 donkey. He approached the giant and
 demanded the return of his donkey.

E.G's

- 3: 'He looked everywhere for him and then he saw the giant had him so he said, 'You give my donkey back!''
- 2: 'He climbed up the hill after the giant.'
- 1: 'He went to rescue him.'
- 0: 'He didn't know where to look.'

RELATED
IDEAS:

- Q: What made Tom so desperately worried about his donkey's escape?
- T: A fierce giant lived on the hill near his village. He thought the donkey might get into the cruel clutches of the giant.

E.G's

- 3: 'There was a fierce giant nearby and he was worried that the giant would catch his donkey and hurt him.'
- 2: 'Because the giant might capture him.'
- 1: 'He didn't want his donkey to get hurt.'
- 0: 'Because he didn't know where he was.'

INFER-
ENCE:

- Q: What made Tom able to be so brave?
- T: The donkey was his favourite friend. He was so worried about his donkey and so furious with the giant for stealing him that he forgot to be frightened.

E.G's

- 3: 'He loved his donkey so much that he just wanted to rescue him. He wasn't

scared of the giant.'

2: 'Because the donkey was his favourite
friend.'

1: 'He was angry.'

0: 'He knew the giant wouldn't hurt him.'

APPENDIX 4EXAMPLE OF TRANSCRIPTION AND SCORING
OF ORAL READING ERRORS

The example presented is based on the actual protocol of a high progress reader at reading age seven reading under the illustration condition. A few additional errors taken from other protocols have, however, been included in order to demonstrate some of the more exceptional aspects of error transcription and scoring.

STORY 1: R.A.7.

One day, Roger Red-hat and
 John¹
Johnny Yellow-hat had climbed up the

hill near their home. Johnny had

brought his kite. There was a

^cstring
strong wind blowing. Johnny wanted

to fly his kite.

The wind blew the kite ^{high}²
higher.

Johnny held the string and ran

after it.

"Don't let go, Johnny", cried Roger.

"The wind is too strong", called

Johnny.

"Come and help me ^(to)³ pull it in",

the⁴ hill⁵
 Roger ran over to Johnny.

They pulled the kite in together.

They were going back down the
 Hill, when Roger ^{stopping}⁶ ^a⁷
stopped near the

gate. He pointed ^c
at and cried, "Look,

Johnny!" There, under the tree, was

<u>ERROR</u> <u>NO:</u>	<u>ORT</u>	<u>SEM</u>	<u>SYN</u>
1	3	4	4
C	4	1	2
2	3	3	4
3	0	4	4
4	3	3	4
5	0	3	4
6	2	3	2
7	0	4	4
C	3	1	2

a donkey. "Come on Johnny",
 he said. "Let's have a ride
 on the donkey".
 Quick⁸
Quickly, they ran down the hill
 and into the field. Roger ^{'dishde'}dashed
 over towards the donkey but the
 donkey ^Csaw him coming and ran off,
 wagging ^{his}its tail.
 Roger waved his ^Chand arms. "Stop!"
 he cried. But the donkey would not
 listen.
 "Wait!" shouted Johnny.
 "Come back, Roger. I brought an
 apple with me. The donkey will
 come ^{on} for an apple." Roger turned
 and ^{12^}looked ¹³to look. "Right" he said.
 "You ^{climbed}climb ¹⁴ up onto that
 branch," said Johnny. "I will
 hold out the apple. Then, when

<u>ERROR</u> <u>NO</u>	<u>ORT</u>	<u>SEM</u>	<u>SYN</u>
8	3	3	2
9	3	0	0
C	4	1	2
10	3	4	4
C	0	3	4
11	0	3	4
12	0	4	4
13	3	4	4
14	3	3	3

the donkey comes nearer, you can

slip off the ^C bench branch."

Roger climbed up the tree

while Johnny ¹⁵ started ¹⁶ to hold ¹⁷ stretched out his

hand with the apple ¹⁸ and [^] The donkey

sniffed at the apple. "Come on,

donkey" said Johnny. The donkey

came ¹⁹ forward and took the apple

from Johnny's hand.

Roger slipped down onto the

donkey's back. Off he ^{'goppled'} ²⁰ galloped.

"Hold on ²¹ tightly, Roger!" shouted

Johnny.

Then the donkey put his head

down and kicked his ^C hooves ^{'hov-es'} high ^{into} ²² in

the air. Roger went flying. He

landed ²³ in a ^C haystack.

Johnny ^{"dashted"} ²⁴ dashed across the

field to Roger ^C who was ^C running rubbing

ERROR NO	ORT	SEM	SYN
C	2	2	4
15	2	2	2
16	3	2	2
17	1	2	2
18	0	4	4
19	4	4	4
20	2	0	4
21	3	3	4
C	3	0	0
22	3	4	4
23	4	3	4
C	2	0	0
24	3	3	4
C	3	1	2
C	3	1	4

his back. "I'm not going to try
that again," he said, looking
angrily²⁵
angry.

ERROR	ORT	SEM	SYN
<u>NO</u>			
25	2	3	2
Σ	77	85	103

C = 9

SCORES:

- a (total number of words) = 309
- b (number of uncorrected errors) = 25
- c (self-corrected errors) = 9
- d (b + c) = 34
1. Word identification accuracy = $100 \left(\frac{a-b}{a} \right) = 91,91$
 2. Semantic acceptability of errors = $100 \left(\frac{\Sigma SEM}{4d} \right) = 62,50$
 3. Syntactic acceptability of errors = $100 \left(\frac{\Sigma SYN}{4d} \right) = 75,74$
 4. Orthographic acceptability of errors = $100 \left(\frac{\Sigma ORT}{5d} \right) = 45,29$
 - 5 Self-corrections = $100 \left(\frac{c}{b+c} \right) = 26,47$

APPENDIX 5TABLES OF INDIVIDUAL VARIATES7, H, P1

ACC	SEM	SYN	ORT	COR	LIT	INF
95,28	47,62	70,24	47,62	38,10	70,00	33,33
96,12	55,43	76,09	36,52	47,83	50,00	0,00
97,43	64,06	81,25	38,75	50,00	40,00	100,00
88,42	29,79	61,17	51,49	23,40	60,00	33,33
95,18	55,00	63,75	42,00	25,00	60,00	33,33
94,82	45,00	59,00	44,80	36,00	60,00	66,67
97,41	58,82	79,41	52,94	52,94	60,00	33,33
98,06	59,09	75,00	40,00	45,45	50,00	33,33
99,35	58,33	83,33	40,00	33,33	60,00	0,00
94,50	51,92	87,50	47,69	34,62	40,00	0,00
97,40	47,73	54,55	36,36	27,27	80,00	0,00
97,40	44,64	50,00	48,57	42,86	20,00	0,00
94,16	28,85	35,58	45,38	30,77	70,00	66,67
97,40	38,24	63,24	42,35	52,94	50,00	0,00
95,78	31,25	63,75	54,00	35,00	80,00	66,67
MEANS						
95,91	47,72	66,92	44,57	38,37	56,67	31,11
STANDARD DEVIATIONS						
2,55	11,40	14,05	5,75	9,86	15,89	32,04

7, H, PO

ACC	SEM	SYN	ORT	COR	LIT	INF
98,38	46,88	78,13	70,00	37,50	40,00	66,67
95,15	29,72	81,94	43,33	28,57	30,00	66,67
96,76	56,82	93,18	61,82	16,67	40,00	100,00
85,44	37,50	50,00	49,17	10,00	30,00	0,00
93,20	22,58	52,42	63,23	38,24	20,00	0,00
93,57	33,00	44,00	60,00	28,57	50,00	100,00
95,13	36,76	67,65	63,53	16,67	50,00	0,00
86,50	32,39	40,34	43,64	4,55	30,00	0,00
66,88	17,76	27,10	43,36	4,67	0,00	0,00
89,39	11,03	59,56	84,71	5,71	20,00	0,00
77,60	12,50	35,14	66,22	6,76	0,00	0,00
91,00	25,00	58,87	58,06	9,68	30,33	0,00
90,91	22,79	35,29	58,24	20,00	50,00	0,00
90,03	22,92	64,58	63,89	16,22	20,00	0,00
89,61	12,50	28,68	54,12	8,57	30,00	33,33
MEANS						
89,30	28,01	54,46	59,89	16,82	29,36	24,44
STANDARD DEVIATIONS						
8,05	13,00	20,04	11,19	11,53	15,80	38,76

7. L, P1.

ACC	SEM	SYN	ORT	COR	LIT	INF
98,71	81,25	65,63	32,50	50,00	60,00	100,00
93,20	51,72	71,55	53,10	27,59	50,00	0,00
92,56	56,90	68,97	45,52	20,69	80,00	66,67
95,15	59,09	75,00	46,36	31,82	80,00	100,00
96,12	41,07	60,71	57,14	14,29	70,00	66,67
88,42	46,88	75,00	45,00	10,00	50,00	66,67
92,93	34,48	75,86	58,62	24,14	50,00	100,00
93,25	53,45	81,90	48,28	27,59	60,00	66,67
95,50	46,25	68,75	43,00	30,00	80,00	100,00
93,57	40,00	64,17	41,33	33,33	70,00	0,00
97,08	35,29	69,12	45,88	47,06	70,00	0,00
90,91	49,35	60,53	32,11	26,32	50,00	33,33
96,10	44,12	72,06	37,65	29,41	40,00	66,67
95,78	39,77	62,50	42,73	40,91	70,00	100,00
97,73	53,57	83,93	58,57	50,00	60,00	0,00
MEANS						
94,47	48,88	70,38	45,85	30,88	62,67	57,78
STANDARD DEVIATIONS						
2,72	11,70	7,14	8,39	11,97	12,80	40,76

7, L, PO

ACC	SEM	SYN	ORT	COR	LIT	INF
96,76	47,37	61,84	33,68	47,37	40,00	0,00
96,75	38,46	61,54	67,69	23,08	80,00	0,00
95,15	51,61	75,81	34,19	51,61	40,00	100,00
81,23	39,75	53,69	39,02	4,92	50,00	66,67
93,57	51,43	52,68	70,00	28,57	40,00	100,00
86,82	46,35	59,90	44,17	14,58	60,00	100,00
88,42	41,30	58,15	55,22	21,74	20,00	33,33
94,50	40,28	76,39	48,89	5,56	50,00	66,67
87,78	31,91	52,13	53,19	19,15	20,00	100,00
96,43	57,95	64,77	45,45	50,00	70,00	66,67
82,47	41,42	60,45	47,76	19,40	30,00	66,67
96,43	53,85	84,62	35,38	15,38	80,00	0,00
94,17	44,00	67,00	48,80	28,00	60,00	0,00
93,25	30,43	41,30	57,39	8,70	30,00	0,00
87,01	28,92	65,69	52,16	21,57	50,00	0,00
MEANS						
91,38	43,00	62,40	48,87	23,98	48,00	46,67
STANDARD DEVIATIONS						
5,30	8,62	10,88	11,03	15,07	19,35	43,28

9, H, P1

ACC	SEM	SYN	ORT	COR	LIT	INF
94,64	26,25	78,75	52,00	10,00	70,00	66,67
95,54	50,00	86,00	51,20	40,00	60,00	66,67
95,54	44,32	57,95	51,82	31,82	70,00	33,33
95,83	38,06	87,50	31,25	12,50	60,00	66,67
95,83	40,63	96,88	52,50	12,50	20,00	66,67
95,60	53,95	86,84	50,53	21,05	70,00	100,00
92,67	44,59	66,89	46,65	32,43	70,00	100,00
98,53	47,50	82,50	28,00	50,00	50,00	66,67
97,65	45,83	47,92	40,00	33,33	70,00	66,67
99,41	56,25	75,00	40,00	50,00	70,00	100,00
94,31	46,25	80,00	49,00	5,00	30,00	66,67
98,20	35,71	100,00	60,00	14,29	30,00	66,67
93,71	50,00	66,67	45,83	12,50	90,00	66,67
97,60	70,83	72,22	20,00	55,56	70,00	66,67
96,11	54,17	80,56	56,67	27,78	50,00	33,33
MEANS						
96,08	47,02	77,71	45,16	27,25	58,67	68,89
STANDARD DEVIATIONS						
1,89	10,17	13,90	11,23	16,34	19,22	19,79

9, H, PO

ACC	SEM	SYN	ORT	COR	LIT	INF
93,75	36,11	69,44	48,89	22,22	60,00	66,67
94,64	50,00	84,72	34,44	0,00	90,00	100,00
99,11	50,00	42,86	40,00	57,14	60,00	100,00
99,11	46,88	68,75	47,50	62,50	60,00	66,67
95,54	42,11	77,63	40,00	21,05	20,00	0,00
96,48	40,38	75,00	52,31	7,69	50,00	33,33
97,95	44,44	77,78	66,67	22,22	50,00	33,33
95,01	28,13	88,54	55,83	29,17	40,00	100,00
96,71	35,00	75,00	41,33	26,67	30,00	33,33
95,51	38,64	81,82	50,91	31,82	70,00	33,33
94,31	50,00	72,83	40,00	17,39	40,00	0,00
88,02	45,41	73,47	45,31	18,37	60,00	0,00
96,19	41,18	57,35	51,76	23,53	40,00	66,67
94,72	26,14	62,50	63,64	18,18	60,00	0,00
97,01	32,69	76,92	53,85	23,08	40,00	66,67
MEANS						
95,60	40,47	72,31	48,83	25,40	51,33	46,67
STANDARD DEVIATIONS						
2,66	7,69	11,38	9,06	16,07	17,27	37,37

9, L, P1

ACC	SEM	SYN	ORT	COR	LIT	INF
98,21	67,50	70,00	34,00	40,00	80,00	66,67
98,81	63,64	72,73	43,46	63,64	80,00	66,67
92,56	52,42	87,10	47,74	19,35	70,00	66,67
95,24	59,52	75,00	46,67	23,81	70,00	66,67
97,62	72,92	95,83	51,67	33,33	90,00	66,67
97,95	60,42	72,92	33,33	41,67	100,00	66,67
92,96	37,50	70,83	54,00	20,00	90,00	66,67
92,96	72,50	78,33	35,33	20,00	70,00	100,00
92,08	42,76	78,29	53,68	28,95	70,00	66,67
89,52	62,20	75,61	30,73	14,63	100,00	66,67
90,42	50,74	76,47	44,71	5,88	60,00	66,67
89,82	53,41	73,30	47,73	22,73	70,00	66,67
96,11	41,18	75,00	44,71	23,53	70,00	0,00
96,77	50,00	80,00	48,00	26,67	90,00	0,00
92,51	59,17	75,83	32,00	16,67	60,00	0,00
MEANS						
94,24	56,39	77,15	43,18	26,72	78,00	55,56
STANDARD DEVIATIONS						
3,20	10,85	6,63	8,04	13,82	13,20	29,99

9, L, PO

ACC	SEM	SYN	ORT	COR	LIT	INF
91,37	59,09	81,06	30,30	12,12	80,00	66,67
98,81	65,00	55,00	20,00	20,00	40,00	33,33
91,37	44,70	76,52	48,48	12,12	40,00	66,67
89,58	45,73	60,37	49,27	14,63	20,00	0,00
92,56	47,41	74,14	43,45	13,79	60,00	66,67
94,43	54,55	78,41	47,27	13,64	80,00	100,00
92,08	48,39	62,10	41,29	12,90	50,00	100,00
96,77	54,17	88,33	49,33	26,67	70,00	100,00
82,99	44,84	54,76	53,97	7,94	30,00	0,00
92,08	25,86	67,24	63,45	6,90	20,00	0,00
88,92	37,18	98,59	52,31	5,13	40,00	0,00
87,13	35,80	64,20	43,18	2,27	40,00	66,67
96,41	31,67	76,67	58,67	20,00	80,00	66,67
85,63	41,67	61,27	44,71	5,88	40,00	66,67
90,12	25,00	69,51	45,85	19,51	70,00	66,67
MEANS						
91,35	44,07	71,21	46,10	12,90	50,67	53,33
STANDARD DEVIATIONS						
4,26	11,59	12,42	10,55	6,67	21,20	37,37

6. MEANS/STANDARD DEVIATIONS (A) AND ANALYSIS
OF VARIANCE SUMMARY TABLES (B) FOR EXPERIMENT I

i Word Identification Accuracy

A			
	P1	P0	COMBINED
GROUP I	84,70 (8,85)	84,39 (10,86)	84,55
GROUP II	88,64 (7,10)	81,06 (9,24)	84,85
COMBINED	86,67	82,73	

B				
SOURCE	SS	df	MS	F
<u>BETWEEN</u>				
GROUPS	0,96	1	0,96	0,01
ERROR	2397,16	18	133,18	
<u>WITHIN</u>				
PICTURE	154,45	1	154,45	6,14+
P x GR	133,22	1	133,22	5,30+
ERROR	452,53	18	25,14	

+ = $p < ,05$

++ = $p < ,01$

+++ = $p < ,001$

ii Semantic Acceptability of Errors

A

	P1	P0	COMBINED
GROUP I	40,11	26,86	33,49
	(24,98)	(15,00)	
GROUP II	48,37	27,80	38,09
	(18,88)	(11,22)	
COMBINED	44,24	27,33	

B

SOURCE	SS	df	MS	F
<u>BETWEEN</u>				
GROUPS	211,14	1	211,14	0,43
ERROR	3771,42	18	487,30	
<u>WITHIN</u>				
PICTURE	2857,78	1	2857,78	16,02+++
P x GR	134,32	1	134,32	0,75
ERROR	3210,25	18	178,34	

iii SYNTACTIC ACCEPTABILITY OF ERRORS

A			
	P1	P0	COMBINED
GROUP I	48,62	34,23	41,42
	(28,79)	(19,53)	
GROUP II	54,36	40,70	47,53
	(21,75)	(15,84)	
COMBINED	51,49	37,46	

B				
SOURCE	SS	df	MS	F
<u>BETWEEN</u>				
GROUPS	372,10	1	372,10	0,53
ERROR	12579,29	18	698,63	
<u>WITHIN</u>				
Picture	1965,60	1	1965,60	7,33+
P x GR	1,37	1	1,37	0,01
ERROR	4825,87	18	268,10	

iv Orthographic Acceptability of Errors.

A

	P1	P0	COMBINED
GROUP I	32,91	41,41	37,16
	(15,90)	(18,00)	
GROUP II	37,59	38,49	38,04
	(12,76)	(7,02)	
COMBINED	35,25	39,95	

SOURCE	MS	df	SS	F
<u>BETWEEN</u>				
GROUPS	7,57	1	7,57	0,02
ERROR	6202,76	18	344,60	
<u>WITHIN</u>				
PICTURE	221,84	1	221,84	4,46+
P x GR	144,40	1	144,40	2,90
ERROR	895,18	18	49,73	

v. Self-Corrections

A

	P1	P0	COMBINED
GROUP I	22,23	10,43	16,33
	(16,64)	(12,98)	
GROUP II	22,22	10,55	16,38
	(13,13)	(11,02)	
COMBINED	22,22	10,49	

B

SOURCE	SS	df	MS	F
<u>BETWEEN</u>				
GROUPS	0,03	1	0,03	0,00
ERROR	4758,73	18	264,37	
<u>WITHIN</u>				
PICTURE	1375,93	1	1375,93	13,07++
P x GR	0,04	1	0,04	0,00
ERROR	1895,03	18	105,30	

vi Comprehension - Idea Recall

A

	P1	PO	COMBINED
GROUP I	22,50	18,33	20,42
	(10,59)	(12,97)	
GROUP II	30,27	19,17	24,72
	(10,35)	(11,38)	
COMBINED	26,39	18,75	

B

SOURCE	SS	df	MS	F
<u>BETWEEN</u>				
GROUPS	186,19	1	186,19	0,94
ERROR	3585,87	18	199,22	
<u>WITHIN</u>				
PICTURE	582,93	1	582,93	9,85++
P x GR	120,76	1	120,76	2,04
ERROR	1064,94	18	59,16	

vii Inferential Comprehension

A

	P1	PO	COMBINED
GROUP I	70,00	64,44	67,22
	(27,24)	(26,60)	
GROUP II	66,66	42,22	54,44
	(26,19)	(33,05)	
COMBINED	68,33	53,33	

B

<u>SOURCE</u>	SS	df	MS	F
<u>BETWEEN</u>				
GROUPS	1628,17	1	1628,17	1,76
ERROR	16682,34	18	926,80	
<u>WITHIN</u>				
PICTURE	2247,00	1	2247,00	3,28
P x GR	891,13	1	891,13	1,30
ERROR	12346,92	18	685,94	

APPENDIX 7

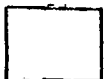
RESULTS OF SCREENING ON THE D. YOUNG
GROUP READING TEST

FREQUENCY TABLE FOR D. YOUNG RAW SCORES AND READING AGES:

GRADES I, III and V

RAW SCORE	READING AGE	F R E Q U E N C Y		
		GRADE I	GRADE III	GRADE V
0		8		
1		4		
2] <6,0	6		
3		6		
4		8		
5		10		
6		10		
7	6,0	9		
8	6,1	25		
9	6,2	21		
10	6,3	18		
11	6,4	14		
12	6,5	22		
13	6,6	35		
14	6,7	49		
15	6,8	51		
16	6,8	53	1	
17	6,9	40	2	
18	7,0	28	3	
19	7,0	38	6	
20	7,1	31	6	
21	7,1	19	8	
22	7,2	10	9	1
23	7,2	13	13	0
24	7,3	15 H	16 L	2
25	7,3	10	10	1
26	7,4	5	15	3
27	7,4	8	24	1
28	7,5	6	13	0
29	7,6	5	19	0
30	7,6	1	20	1
31	7,7	1	25	7
32	7,8	3	22	4
33	7,9	2	29	3
34	8,0	1	23	10
35	8,1	1	-26-	4
36	8,2	1	26	8
37	8,4	2	41	14
38	8,6	2	35	22
39	8,8		35	25
40	9,1		36	29
41	9,4		39 H	39 L
42	9,7		44	67
43	10,1		32	-91-
44	> 10,1 (10,5)		22	119
45	> 10,1 (11,0)		18	208
		N 591	618	659
	MEAN RAW SCORE	15,52	34,80	42,27
	SD	6,61	6,87	3,61

= mean



= Sample pools selected (chapter IX) at respective high and low progress levels

12 JUL 1982