

Some parts of this thesis may have been removed for copyright restrictions.

If you have discovered material in AURA which is unlawful e.g. breaches copyright, (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please read our [Takedown Policy](#) and [contact the service](#) immediately

IMAGERY AND CHILDREN'S STRATEGIES
IN READING COMPREHENSION AND FRENCH VOCABULARY LEARNING

Thesis submitted towards the degree of PhD at the University
of Aston in Birmingham.

Imagery and children's strategies in reading comprehension and French vocabulary learning.

ROGER MERRY

1978 PhD

SUMMARY

Within cognitive psychology, recent research into strategies in general, and a return of interest in imagery in particular, have suggested techniques that might be useful in classroom tasks like comprehension work and foreign language vocabulary learning, but which have so far usually been examined only under rather artificial conditions. An exploratory questionnaire given to a small number of pupils from each of 39 schools showed that, in spite of the views expressed in the specialist subject journals, many children reported that they were often expected to perform both tasks, often with little guidance.

In the vocabulary experiment, analysis of the reports of control group subjects left to their own devices to learn "real" French vocabulary items suggested significant differences in the spontaneous strategies of successful and poor learners, which were not due to differences in their ability to remember or describe what they had done. A technique derived from research into mnemonics proved highly effective in improving 11-year-olds' vocabulary learning on both immediate and delayed tests, giving them scores ranging from 60% to 300% better than those of control group subjects.

In the second experiment, simple imagery instructions were compared with the provision of adjunct questions in improving passage comprehension and learning, but under conditions probably more like those of the classroom, the facilitation reported in previous research was not upheld, the only significant variable being text availability. This, and analysis of the children's reported strategies, suggested important cognitive differences between comprehension tasks often used by researchers and those used in schools. The analysis also showed that the children spontaneously used several strategies that were both differently effective and were themselves differently affected by test conditions. On this basis, re-analysis of the group data suggested that imagery did have significant effects.

Key words: imagery, strategies, comprehension, vocabulary.

ACKNOWLEDGEMENTS

I should particularly like to thank Dr Norman Graham, Reader in the Department of Educational Enquiry, for all his encouragement and advice throughout, and Dr Peter Coxhead (and the Aston computer) for their great help in analysing the data. Thanks are also due not only to the pupils of Archbishop Grimshaw School who acted as subjects, but also to the staff for their interest and help in collecting the data, and to the various group organisers and children who were involved in the questionnaire survey. Finally, I should like to thank Mrs Denise Ratcliffe who took on the unenviable task of typing the whole thing out.

CONTENTS

List of Tables	(i)	
List of Figures	(iv)	
List of Appendices	(v)	
Ch.I	Summary	1
	1. Introduction	2
	2. The Psychological Background	9
	3. Imagery - Introduction	14
	4. Definitions and the Nature of Imagery	18
	5. The Functional Importance of Imagery	25
	6. Imagery, Language and the Dual Coding Hypothesis	30
	7. Imagery and Cognitive Development	38
	8. Some Educational Implications	42
Ch.II	Summary	52
	1. Some Current Issues in Modern Languages	53
	2. A Survey of Classroom Practice	71
	3. Mnemonics	99
Ch.III	Summary	137
	1. The "Keyword" Method	138
	2. Experiment I - Pilot Study	151
	3. Experiment II - Main Study	161
Ch.IV	Summary	224
	1. Introduction	225
	2. Some Current Issues in English Teaching	226
	3. The Survey - Part Two	252
	4. Psychological Studies of Comprehension	267
Ch.V	Summary	300
	1. Two Techniques for Improving Passage Comprehension	301
	2. An Experiment in Children's Passage Comprehension and Learning	331
Ch.VI	Conclusions	383
	References	399
	Appendices	APPENDIX A MISSING

LIST OF TABLES

<u>Table 1</u>	Number of responses in each category, questions 4 and 5.
<u>Table 2</u>	Chi-Square table - use of slides and tapes.
<u>Table 3</u>	Number of responses in each category, questions 4 and 6.
<u>Table 4</u>	Number of responses in each category, questions 5 and 6.
<u>Table 5</u>	Number of responses in each category, questions 10 and 11.
<u>Table 6</u>	Results of six studies comparing experimenter - provided and subject - generated mediators on both immediate and delayed recall.
<u>Table 7</u>	Mean percentage recall, by sentence type and recall condition, children and student subjects.
<u>Table 8</u>	Pilot study vocabulary list.
<u>Table 9</u>	Mean scores by group condition, pilot study.
<u>Table 10</u>	Number of subjects with maximum scores, control and key-word groups, pilot study.
<u>Table 11</u>	The pool of 50 vocabulary items.
<u>Table 12</u>	Main study vocabulary list.
<u>Table 13</u>	Control group (n = 51). Correlations between the four vocabulary tests.
<u>Table 14</u>	Reported use of strategies and their success in recall, control group.
<u>Table 15</u>	Frequency of correct and wrong test answers, according to reported strategy, control group.
<u>Table 16</u>	Control group, high and low scorers. Comparison of mean total scores, reading ages and reported frequency of each strategy.

<u>Table 17</u>	Control group, boys and girls. Comparison of mean total scores, reading ages and reported frequency of each strategy.
<u>Table 18</u>	Rote group (n = 23). Correlations between the four vocabulary tests.
<u>Table 19</u>	Image group (n = 46). Correlations between the four vocabulary tests.
<u>Table 20</u>	Picture group (n = 46.) Correlations between the four vocabulary tests.
<u>Table 21</u>	Reading age by condition, with analysis of variance.
<u>Table 22</u>	Scores on Test I (FE IMM) by condition, with analysis of variance.
<u>Table 23</u>	Scores on Test II (EF IMM) by condition, with analysis of variance.
<u>Table 24</u>	Scores on Test III (FE DEL) by condition, with analysis of variance.
<u>Table 25</u>	Scores on Test IV (EF DEL) by condition, with analysis of variance.
<u>Table 26</u>	Rating by condition, with analysis of variance.
<u>Table 27</u>	Analysis of variance - interactions between covariates (sex and R.A.) and treatment conditions.
<u>Table 28</u>	Mean total scores (out of 68) by reading age and treatment.
<u>Table 29</u>	Number of responses in each category, questions 10 and 17.
<u>Table 30</u>	Number of responses in each category, questions 17 and 18.
<u>Table 31</u>	Number of spontaneous comments on French and English lessons.
<u>Table 32</u>	Examples of variables affecting the value of adjunct questions.

<u>Table 33</u>	Scores on immediate literal comprehension questions by condition, with analysis of variance, all groups.
<u>Table 34</u>	Scores on immediate inference comprehension questions by condition, with analysis of variance, all groups.
<u>Table 35</u>	Scores on delayed adjunct questions by condition, with analysis of variance, all groups.
<u>Table 36</u>	Scores on delayed literal comprehension questions by condition, with analysis of variance all groups.
<u>Table 37</u>	Scores on delayed inference comprehension questions by condition with analysis of variance, all groups.
<u>Table 38</u>	Further analysis of variance, main groups, immediate literal comprehension questions.
<u>Table 39</u>	Further analysis of variance, main groups, immediate inference comprehension questions.
<u>Table 40</u>	Further analysis of variance, main groups, delayed adjunct questions.
<u>Table 41</u>	Further analysis of variance, main groups, delayed literal comprehension questions.
<u>Table 42</u>	Further analysis of variance, main groups, delayed inference comprehension questions.
<u>Table 43</u>	Analysis of variance, reading times of main groups.
<u>Table 44</u>	Reported use of imagery, and test scores, control groups.
<u>Table 45</u>	Reported guessing at questions, and test scores, control groups.
<u>Table 46</u>	Reported re-reading, and test scores, control groups.
<u>Table 47</u>	Differences between control and imagery groups' mean scores, according to reported strategies.

LIST OF FIGURES

- Figure 1 Annual number of references to Imagery in Psychological Abstracts, 1955-1977.
- Figure 2 "Non-verbal communication" (From Gomcrich, 1972).
- Figure 3 An interactive picture presented to group 1. From Ott, Butler, Blake and Ball (1973).
- Figure 4 Distribution of the sample's reading ages (Schonell "B").
- Figure 5 Practice items for "picture" group, main vocabulary study.
- Figure 6 Recall scores of control and experimental groups on Test I (FE IMM) and Test II (FE DEL).
- Figure 7 Recognition scores of control and experimental groups on Test II (EF IMM) and Test IV (EF DEL).
- Figure 8 Mean scores for each group, immediate literal questions.
- Figure 9 Mean scores for each group, immediate inference questions.
- Figure 10 Mean scores for each group, delayed literal questions.
- Figure 11 Mean scores for each group, delayed inference questions.
- Figure 12 Mean scores for each group, adjunct questions on delayed test.
- Figure 13 Mean reading times in minutes for each group.

LIST OF APPENDICES

- Appendix A The questionnaire.
- Appendix B Connecting images supplied by E (Pilot vocabulary study), in order of presentation.
- Appendix C Connecting images and pictures supplied by E (Main vocabulary study), in order of presentation.
- Appendix D The comprehension passage.
- Appendix E The test question sheet.
- Appendix F Marking schemes.
- Appendix G Easiest and hardest words, vocabulary experiment, keyword ("image") group.

CHAPTER I

SUMMARY

The first chapter begins with a brief discussion of pure and applied research, and outlines the attempt to use both approaches in the present study. Although the main emphasis will be on two experiments, some current issues in the two school subjects involved - French and English - need to be mentioned, and a small-scale exploratory questionnaire survey of some current teaching practices in these two areas is also proposed.

The rest of the chapter reviews some of the psychological literature on mental imagery, which has recently re-emerged as a respectable subject and constitutes an important part of a growing interest in learning strategies. Recent issues in imagery research form a major part of the theoretical background to both experiments, and five areas are discussed - definitions and the nature of imagery, language and the dual coding hypothesis, imagery and cognitive development, and some educational implications.

II INTRODUCTION

(a) Approach and Methods of the Study

Psychologists themselves are often the first to admit that the results of laboratory experiments cannot be directly or easily applied to the classroom. The gulf between pure and applied research still looms depressingly large, though Melton (1963) argued that the two merely lie at opposite ends of a continuum, the only difference being the relative freedom of the researcher to manipulate independent variables. Applied research, which takes existing situations as its starting-point, tends to be descriptive rather than prescriptive, relying on survey and observation techniques rather than experiments, and the relative advantages and disadvantages of each approach are well-established (e.g. Borg, 1963; Simon, 1969; Nisbet and Entwistle, 1970). Borg claims that "the experiment is the ultimate form of research design", and many psychologists have long regarded the experimental approach as an essential part of psychology's development as a natural science (e.g. Miller, 1962; Evans, 1977), in spite of the criticisms of their more "radical" colleagues (e.g. Joynson, 1970; Bannister, 1977).

However, this is not the place for arguments such as those which sometimes threaten the pages of the Bulletin of the British Psychological Society. It is true that an experimental approach in educational research can probe into multiple causes, examine extreme cases and even suggest improvements, but its very advantages of control and manipulation have meant that experimental studies naturally tended for many years to emphasise certain sorts of learning at the expense of others - paired associates rather than connected discourse for example (cf Palermo, 1970) - leaving them open to the frequent criticism that their results were largely irrelevant to the classroom (e.g. Carroll, 1963; Kintsch, 1974; Smith, 1975; Phillips, 1976). "Educational implications" often appear in a paragraph or two tacked on the end, even when the researchers themselves feel that such implications are highly important (e.g. Rohwer, 1970; Paivio, 1971; Kulhavy and Swenson, 1975). Conversely, survey approaches, while immediately relevant and "real", suffer from all the disadvantages of trying to describe the complexities of what goes on in schools simply by observing or questioning those involved.

The same sorts of issues appear if one looks not just at the methods, but at the whole basis of "applied" and "pure" research. Hilgard and Bower (1975) sum up the two approaches:

Basic research is necessarily patient, working on such problems as are manageable at the present. Applied research, however, must move ahead and solve as best it can the pressing problems. Hence basic research looks for weaknesses in present conceptions that need re-examination; applied research looks to what we know well enough to provide hints upon which we can make decisions affecting the present. It turns out that many of the quarrels of the theorists are over matters of a firmer foundation for their theories, but not very important in relation to problems of training.

What emerges is simply that the teacher and the experimenter are often interested in different things - the teacher being "achievement - orientated" (Williams, 1960), so that perhaps not all the blame for lack of contact should rest on the latter. Perhaps teachers themselves should become more concerned with "processes" as well as "products" - certainly those trying to apply cognitive psychology to education would agree (e.g. Farnham-Diggory, 1972; Broadbent, 1975, Smith, 1975). McDonald (1963) went one stage further and suggested that, while experimenters should try to relate their work to the classroom, it was also up to teachers to become more like experimental psychologists themselves, carrying out small experiments rather than simply relying on common-sense.

The value of such small-scale experimental research in the classroom has received indirect support from Bruner (1966), who saw experimental evidence as the basis of an eventual theory of instruction, as well as more direct support from writers like Carroll (1968), Heien (1973), Wambold and Hayden (1975) and Kintsch and Bates (1977). The rationale behind small-scale experimental research is that there is no one "best method" of learning for all occasions, and that the complexities of the classroom can best be examined by controlled experiments on specific situations, especially if the researcher takes actual classroom practice, rather than theoretical issues from the research literature, as his starting-point (Carroll, 1963).

The main emphasis in the research to be reported here is on such an experimental approach, but in view of what has been said already, care was taken to make the experiments as "real" as possible, without sacrificing too much of the control which gives them their value.

This was attempted in two ways:

- (a) An exploratory survey was carried out, using a questionnaire given to children from a number of schools, to give an idea of current teaching practices in the areas chosen for study, and to ensure that the experimental conditions were actually related to what goes on in classrooms.
- (b) The materials used in both experiments were not artificially created to examine current issues in the research literature, but were drawn from the existing curriculum, and teachers were closely consulted whenever possible on selection, presentation and testing.

(b) Content and Outline of the Study

There appears to be a growing interest in language in education, ranging from the language used by teachers and pupils to the linguistic demands of various school subjects, perhaps partly due to developments in psychology, sociology and linguistics, and culminating in the Bullock report (1975) and in various projects funded by the Schools Council and the N.F.E.R.

The two areas chosen for study here both emphasise language skills, but at two very different levels and in two normally separate school subjects - French and English. The first topic was chosen largely because of its close relationship with a well-established field of psychological research, and involved the learning of individual French words and their meanings. The second topic is one that psychologists have only recently begun to examine, in spite of its widespread educational implications - children's comprehension of long passages. In each case, a review of the literature on teaching each subject showed that these are currently important and even controversial topics, while the questionnaire survey suggested that they are both tasks which many teachers do expect their pupils to perform, often with a minimum of guidance. Moreover, an examination of the psychological research literature suggested various techniques which appear useful in related laboratory tasks, but which have not yet been assessed properly in the classroom, in spite of their authors' claims of educational relevance.

Consequently, the background to each experiment involves a discussion of two areas which often remain unrelated - the theoretical issues and experimental findings of "pure" psychological research, and the ideas and practices of specialists in each subject.

The rest of chapter one, therefore, looks briefly at some areas of cognitive psychology, particularly research on imagery, which contain theoretical basis for both studies. Chapter two examines some current curricular issues in French teaching and presents the results of the questionnaire survey, before focusing on a specific technique derived from "pure" research into mnemonics. Chapter three consists of a pilot study and the first main experiment, on strategies for learning vocabulary. The next two chapters can be seen as parallel to chapter two and three. Chapter four contains a review of current issues in English teaching, particularly comprehension, the results of the second half of the questionnaire survey, and some ideas on comprehension drawn from an examination of the psychological literature. Chapter five consists of the second main experiment, on comprehension strategies, while chapter six attempts to summarise and discuss the main conclusions.

12 THE PSYCHOLOGICAL BACKGROUND

The decline of traditional Behaviourism saw a corresponding re-awakening of interest in the mental processes that had interested the early psychologists. Miller, Galanter and Pribram (1960) tried to synthesise the two approaches in "subjective behaviorism"; a further influential landmark was Neisser's Cognitive Psychology (1967), and by 1972, Dodwell could claim that "psychology, in short, has re-discovered its mind". In memory studies, for example, instead of trying to minimise the subject's contributions in the Ebbinghaus tradition, more psychologists became interested in these contributions in their own right; the research on Natural Language Mediators (e.g. Prytulak, 1971) is one example of an attempt to combine two very different traditions in psychology, though it may be open to the criticisms that its data are largely only correlational (Underwood, 1972), or are simply responses to the experimenter's questions (Montague, 1972).

In spite of problems of research design, there has thus been a great renewal of interest in both the verbal and non-verbal strategies used by subjects in a wide variety of tasks.

In general, the spontaneous use of strategies appears to increase with age (e.g. Flavell, 1970; Hagen, 1972; Rohwer, 1973; Hunter, 1974; Reese and Porges, 1976), but a growing number of studies have shown that even quite young children can improve their performance in a variety of tasks if they are given a suitable strategy (e.g. Mowbray and Luria, 1973; Brown, 1973; Kulhavy and Swenson, 1975; Paris and Lundauer, 1976).

Parallel findings have been made by those working with retarded subjects rather than young children, and learning failure can sometimes be seen in terms of poor strategies (e.g. Yarmey and Bowen, 1972; Paris, Mahoney and Buckhalt, 1974; Odom-Brooks and Arnold, 1976). Similarly, the provision of suitable strategies by the experimenter has often led to large immediate improvements (e.g. Yarmey and Bowen, 1972; Riding and Shore, 1974; Hunter, 1976). On the other hand, the facilitation produced by training may not last, and results have often been disappointing when retarded subjects are tested again after a delay (e.g. Jensen and Rohwer, 1963; Milgram, 1967 (a); Wambold and Hayden, 1975). More optimistically Zupnick and Meyer (1975), while agreeing that facilitation does not last, did find that retarded subjects could transfer what they had learned to a new but similar task, provided that it was given immediately after initial learning.

Long-term effects were claimed by Yuille and Catchpole (1974), whose six year old normal subjects showed evidence of permanent strategy changes after training.

The exact nature of learning failure seen in terms of deficient strategies is still not completely settled. Several researchers have suggested that retarded subjects do have some spontaneous strategies; Gordon and Baumeister (1971) claim that their subjects could generate sentence mediators for themselves; Paris, Mahoney and Buckhalt (1974) found evidence of a certain amount of spontaneous semantic integration in retardates; Flavell (1970) points out that, although normal and older subjects are more likely to use various mediation strategies spontaneously, it is not an "all-or-nothing" affair, and failure may be due to inefficient or inappropriate strategies rather than a complete absence of them. This suggestion that mediators are available but not spontaneously used gave rise to the proposal of "mediation deficiency" as opposed to "production deficiency". The former suggests that the subject may well be able to produce a suitable mediator (e.g. a linking image or a sentence), but will not spontaneously use it to improve performance. "Production deficiency", on the other hand, claims that the subject is unable even to produce a suitable mediator.

These alternatives have been discussed by several writers (e.g. Reese, 1970; Flavell, 1970; Jablonski, 1972; Pressley, 1977), and either may be the cause of learning failure, depending on the circumstances of the study. Paris and Lindauer (1976) add a word of caution about the dangers of circular arguments, and prefer to emphasise "the processing demands of the task as perceived by the children". Finally, it must be pointed out that facilitation brought about by strategy instructions, though widespread, is not automatic, and depends on the interaction of subjects, materials, and instructions. (Levin, Horvitz and Kaplan, 1972). For example the opportunity to act out relationships or manipulate objects (Paris and Lindauer, 1976; Wolff and Levin, 1972) to illustrate the situations (Lesgold, McCormick and Golinkoff, 1975) or to verbalise about the mediators (Taylor, Josberger and Whitely, 1973) may be crucial, depending in particular on the age of the subjects. Such factors would appear to be related to Bruner's theory of the development of representation, and need to be discussed later (17).

What emerges, even from this very brief outline, is that there is a growing body of research which is trying to examine and improve learning by studying subjects' strategies.

Almost all the work so far has tended to be "pure" experimental research. in the sense suggested in the introductory section, but its attempts both to relate subjects' strategies to learning success and to teach strategies in the hope of improving learning form the basis for the experimental parts of the present study. Consequently, it is important to look in more detail at some of the general issues behind this research before examining the two experimental topics themselves.

I3 IMAGERY - INTRODUCTION

In the general field of cognitive psychology, one area has shown a particularly rapid growth over the last few years - mental imagery, long considered a subject unfit for serious study, has once again become respectable. Some idea of this growth is given in Figure 1, which shows the number of references to Imagery in Psychological Abstracts, 1955 - 1977. Even allowing for an increase in psychological research generally, the development is obvious. Moreover, the status of imagery seems to have changed. In 1960, the odd reader who was interested in that sort of thing was advised to look also at "autism and fantasy", while those currently interested in images are now also advised to look under "conceptual imagery". The importance of images in cognition has thus perhaps been recognised, even if not yet fully understood. Research on imagery forms a major part of the theoretical background to both the experiments to be reported here, and some of the major issues in recent research need to be briefly examined before looking at the two more specific topics of mnemonics (Chapter Two) and imagery in comprehension (Chapter Four).

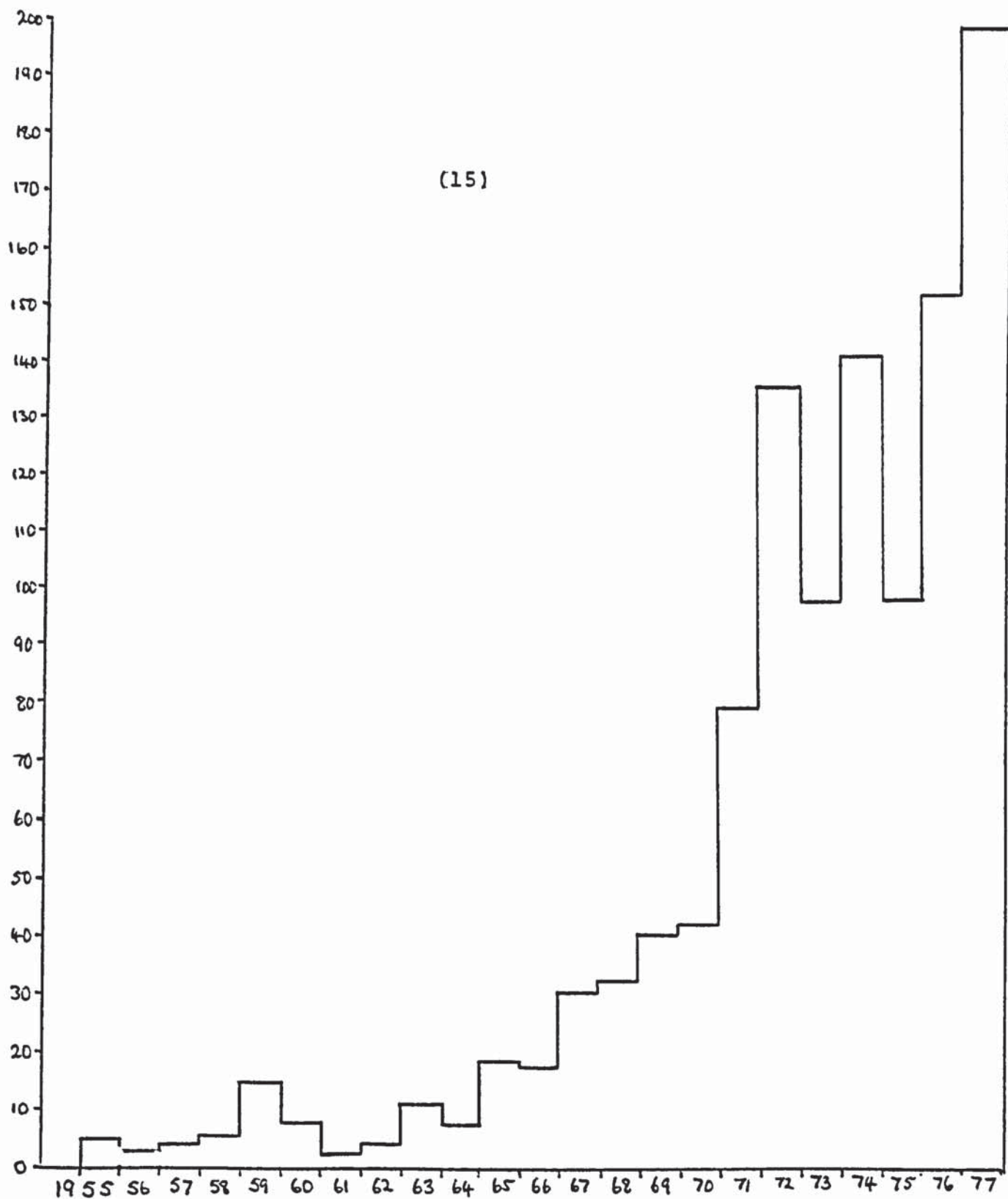


Figure 1 Annual number of References to Imagery in
Psychological Abstracts, 1955-1977.

Interest in mental images has a turbulent history which can be traced back to the ancient Greeks; the pioneering psychological work is usually accepted as Galton's questionnaire study of 1880 (reprinted in Semeonoff, 1966), though Fechner had produced a list of types of imagery twenty years earlier. Galton was interested in individual differences, and was surprised to find that many of his learned colleagues reported little or no imagery, while younger subjects often described very vivid and detailed images. Galton relied on introspection, concluding that:

It is a much easier matter than I had anticipated to obtain trustworthy replies to psychological questions. Many persons, especially women and intelligent children, take pleasure in introspection, and strive their very best to explain their mental processes.

This enthusiasm was not shared by Behaviourists, however; Watson dismissed images as "mere ghosts", unavailable and unsuitable for proper psychological study.

Moreover, the apparent findings of "imageless thought" and the failure to establish the functional significance of imagery by studying individual differences meant that verbal behaviour rather than visual images was to dominate American psychology for the next forty years (Paivio, 1970). However, the decline of Behaviourism itself, and its inability to avoid inferring "mental variables" and subjects' covert contributions would eventually allow a return of interest in imagery and the development of techniques of manipulating stimulus attributes and experimental procedures as well as studying individual differences.

Perhaps the best known statement of this re-emergence is Holt's paper "Imagery: the return of the ostracised" (1964), which justifies the study of images, and describes various types ranging from dreaming to "phantom limb" experiences. Richardson (1969), Paivio (1969) and Palermo (1970) all trace the decline and return of imagery, and an important landmark in the field is Paivio's Imagery and Verbal Processes (1971), though the vague disquiet that some of his colleagues might still feel was summed up a year later in Bower's half serious comment:

Many experimental psychologists cannot entertain thoughts about imagery without some deep sense of guilt associated with forbidden taboos.

(Bower, 1972)

The early psychologists had turned to introspection as an alternative to the reliance on the Authority of the Classical writers; introspection was rejected by the Behaviourists in favour of the experimental study of overt behaviour; traditional Behaviourism has in turn been rejected by many psychologists who are concerned with "mental events" without wishing to lose experimental objectivity. By 1974, Starker could claim that, "once abandoned as non-measurable and hopelessly subjective, the mental image has begun to yield its secrets to the tools of cognitive psychology". Some of the current issues and findings will now be discussed, divided for convenience into five sections - definitions and the nature of imagery, the functional importance of imagery, imagery, language and the dual coding hypothesis, imagery and cognitive development, and some educational implications.

I4 DEFINITIONS AND THE NATURE OF IMAGERY

It is perhaps not surprising that several rather different views of the nature of imagery have arisen since its return to popularity. The vast majority of work has involved visual imagery, and only a few writers seem to feel that this emphasis could be possibly damaging (e.g. Lindauer, 1972).

A visual image would appear to have at least some perceptual attributes, and some writers have stressed the visual nature of mental images; conversely, the fact that an image can be formed in the absence of appropriate sensory stimulation implies that it should also be at least partly conceptual and schematic in nature. Each approach will be briefly considered.

(a) The Visual Nature of Imagery

The traditional "picture metaphor" - the idea that an image is like a picture "in the mind's eye" has come under strong attack from a variety of sources, including those who place a great emphasis on the functional importance of imagery. Indeed, Kosslyn and Pomerantz (1977) claim that attacks on imagery like that of Pylyshyn (1973) are invalid because they are simply attacks on a view that "no serious student of imagery" holds. For them, the picture metaphor is a "straw-man". Others who have worked in the field are also at pains to dismiss the picture metaphor, though not always for quite the same reasons. Like Pylyshyn, Anderson and Bower (1973) prefer conceptualization in terms of underlying propositions. A similar point is made by Bransford and McCarrell (1974), who say that knowledge derived from perception is abstract and cannot be accounted for in a "storehouse of images" view.

A rather different emphasis has been suggested several times by Paivio (e.g. 1974, 1975) who feels that the picture metaphor implies a static form of representation that cannot do justice to the rich, dynamic properties of imagery. Neisser's view (1972) is similar - images are not "mental contents awaiting descriptions", but "imaging is a kind of mental activity". The picture metaphor is thus unacceptable to writers adopting a variety of theoretical positions, and its rejection does not appear to weaken the arguments of those who support the functional significance of images (see section five below).

And yet it often appears that mental images do seem to have many of the visual properties of actual pictures - in spite of the problems posed by a "mind's eye" (presumably directed by a "mind's brain")? Kosslyn (1974), and Kosslyn and Alper (1977) describe studies in which the imagined size of part of an image is related to its ease of recall; Kosslyn and Pomerantz (1977) claim to show the effects of "scanning" visual images; Beech (1978) found that the time needed to scan an image of a set of digits increased as a function of the "distance" between the digits-a result which is difficult to explain in terms of abstract propositional processes. Similar support is given by the studies carried out on mental rotation by Shepard and Metzler (1971), where reaction time increased with the angle of rotation, and Shepard and Feng (1972) who studied "mental paper folding".

Another approach has involved showing that performance of a visual task interferes with imaging, suggesting that the two involve at least some of the same processes. The best-known study is that of Brooks (1968), though Janssen (1976) found differential interference effects in the recall of words with high and low imagery value, and words studied under verbal and imagery instructions, claiming that "the visual qualities of a visual image are retained all the way from acquisition to retrieval". Such evidence of selective interference has been reviewed by Bower (1970), Alwood (1971), Montague (1972) and Turvey (1974), though Anderson and Bower (1974) do point out that not all replication studies have supported the idea. Although many psychologists reject the picture metaphor, therefore, a variety of experimental approaches have been used to emphasise the visual nature of images.

(b) The Symbolic Nature of Imagery

The alternative view, stressing the symbolic and schematic aspects of images, has already been briefly mentioned, two of its strongest proponents being Pylyshyn (1973) and Anderson and Bower (1974), while Smith (1975) suggests a similar argument. Stated in its simplest form, the idea is that images are the products of memory and thought, rather than their basis, and a parallel can perhaps be traced with Piaget's position (Piaget and Inhelder, 1971):-

As the image is an internalised imitation, it develops under the influence of new external models provided by the intelligence as its overall functioning develops. The..... image thus forms as a result of the symbolic invitation of the operations.

Similarly, Bower (1970) states that "the information represented in the ordinary memory images is largely conceptual, generic, schematic", and Levin, Davidson, Wolff and Citron (1973) likewise conclude that imagery, and language, "are both manifestations of the child's symbolic function, being derived from the same process". Paivio too is careful to stress the symbolic nature of imagery, seeing images and verbal processes as alternative coding systems, or models of symbolic representation (e.g. Paivio 1969, 1971). Finally, mention should perhaps also be made of the rather different way in which Miller, Galanter and Pribram (1960) used the term "image" to mean "all the accumulated, organised knowledge that the organism has about itself and the world". Here, the term is obviously schematic, and need not involve any visual elements at all.

(c) Conclusion - The Nature of Imagery

It is perhaps not surprising that psychologists disagree about the nature of mental images, especially when one considers that they also disagree about the psychological nature of external, explicit "images" in the form of pictures, which would appear much more open to examination.

The Gibsons are perhaps the best-known exponents of a veridical or "registration" theory (e.g. Gibson, 1967), and a Gibsonian view of picture perception has been proposed by Kennedy (1974). (Interestingly, Kennedy concludes that "we do not need captions to help us identify the pictured items", yet provides a detailed caption for every single illustration in his book.....). Conversely, writers like Neisser and Gombrich offer a "constructive" theory of picture perception, where the meaning of the image is largely constructed by the viewer, and there are thus some parallels with the opposing views of mental images briefly described above. (cf Bruner's comment that "images 'stand for' perceptual events in the close but conventionally selective way that a picture stands for the object pictured" - Bruner, 1964).

There are perhaps two different ways of reconciling these apparent contradictions. Firstly, it has long been accepted that there are not only very large individual differences in clarity, control and use of images, but that there are also several different kinds of image. At one extreme is the rare "eidetic" image, which can be considered as a prolongation of the stimulus, and is thus highly visual in nature (e.g. Richardson, 1969). Such images, involving very little "construction", are apparently much more common in retarded or brain-damaged subjects (e.g. Siipola and Hayden, 1965; Van Tharp, 1972).

At the other extreme, one might consider the deliberate, controlled formation of novel "imagination" images, and the numerous accounts in the literature of images being used in problem-solving (e.g. Arnheim, 1970). These different types are described by McKellar (1972) for example, who points out that images are involved in both "A" and "R" thinking. The wide range of individual differences and types of images would thus be one way of trying to reconcile the disagreements about the nature of imagery.

A second approach would be to adopt a "psychological metaphor", rather than metaphors based on technology for recording information. Thus Paivio's (1974) conclusion is worth quoting at some length:-

The mind is not a wax tablet, or a tape-recorder, or computer. It does not "contain" pictures, words or propositions. Such metaphors.....are limited and potentially misleading. Their limitations should not be imposed on the organism..... there is good reason to believe that visual imagery is, in some respects at least, rather like active visual perception. This is a far cry from saying that images are like static pictures or abstract propositions... ..Pictures and propositions tend to divert attention away from the psychological phenomena to the properties of photographs and logical machines. Mind becomes an imitation of the model.

To the extent that perception itself depends both on stimulation and schemas, the dispute can be resolved by defining imagery as using at least some of the mechanisms of perception, in the absence of appropriate stimulation, and such a definition is roughly in line with those proposed by Hunter (1957), Richardson (1969, 1977), Neisser (1967, 1972) or Anderson (1975).

It may not even matter if total agreement on a definition cannot be reached; Kosslyn and Pomerantz (1977) point out that imagery is not the only "fuzzy" concept in psychology and that "unitary operational definitions should be avoided when dealing with inferred constructs, since such definitions confound the measurement" (cf Paivio, above). Similarly, Johnson-Laird and Wason (1977) appear to have little time for problems of definition - "instead of trying to define imagery, it may be more profitable to consider what can be done with it". Consequently, the next section looks briefly at the functional importance of imagery.

15 THE FUNCTIONAL IMPORTANCE OF IMAGERY

It will already be apparent that those theorists who define imagery as a product of deeper thought processes will tend to regard its functional significance as slight. Certainly there are problems with "an image theory of meaning" which go back to the rationalist - empiricist debates, and which have been more recently discussed by writers like Ogden and Richards (1923), Slobin (1971), and Brewer (1974).

Ogden and Richards defined images as "mental luxuries" and "mere signs" probably serving no important function, and others working from a variety of theoretical standpoints have also denied the functional significance of imagery. Thus Kintsch (1974) insists that images cannot form the basis of pre-propositional thought, and Riding (1977) suggests that images may merely be "an optional extra". Franks (1974) defines images as:-

Consciously experienced aspects of particular derivations generated from tacit knowledge:
The structure of an image is determined by tacit knowledge structures. Particular meanings are generated from tacit knowledge, and sometimes these derivations are manifested in imagery. But all images have underlying tacit meanings. Images are not units of storage: They are products.

Experimental support has been offered by several researchers. Clarkson, Haggith, Tierney and Kobasigawa (1973), for example, suggest that any facilitation apparently produced by imaging could really be a result of the "mental activities" required to produce the images rather than the images themselves. Similarly Young, Overby and Powell (1976) conclude - "it may be well to wonder what, if any, information is provided by a mental image. It may be that imagery is simply a by-product of cognitive processing, containing little information".

In spite of all this opinion there is a growing body of work which supports the functional importance of imagery. Early research trying to relate individual differences in imaging to task performance proved disappointing, but more recent studies, manipulating experimental conditions and materials, suggest that imagery can have important functions in a variety of tasks. Paivio's paper "on the functional significance of imagery" (1970) reviews a number of studies supporting a common-sense view "which psychologists perversely rejected in their pursuit of pseudo-objectivity". Certainly, material high in imagery value is generally easier to comprehend and recall under various conditions, and instructions to form images have often greatly facilitated the performance of children, adults and sub-normal subjects. Postman (1975) sums up rather scathingly:-

There has been a large number of new experiments demonstrating positive effects of rated imagery value and imagery instructions on performance in a great variety of tasks, ranging from the Brown-Peterson short-term memory situation to the recall of meaningful sentences. There would be no point in enumerating them. Concreteness is better than abstractness and pictures are better than words.

Two specific aspects of imagery facilitation - in mnemonics and reading comprehension - will be considered in chapter two and four respectively, but though the general body of research is widely accepted, there are one or two more controversial areas where the active role of imagery is contested.

For example the anticipatory nature of imagery has been investigated in detail by Piaget and Inhelder (1971) and forms the basis for Neisser's approach to images as "perceptual anticipations, preparations for picking up certain kinds of information". (Neisser, 1976). Conversely, Bower (1970 a), cites evidence to suggest that an image reduces the likelihood of noticing a similar stimulus because of interference effects, and Morris (1978) feels that Neisser's view has little to offer. Another problem involves the degree to which images need be conscious. Paivio (1974) says that they "can mediate overt responses without necessarily being consciously experienced as visual imagery", yet the ratings of images on which so many of the experiments depend imply that subjects must be conscious of their images - a criticism made by Richardson (1977). A third area of apparent disagreement concerns the relationship between imagery and "creativity"; Paivio (1975) hypothesizes that imagery's richness and flexibility "may underlie the intuitive leaps of imagination that often characterise creative thinking" (with verbal processes supplying "logical direction"). Conversely, Gordon (1949) and Richardson (1972) point out that vivid, autonomous images established early in life could be a factor in stereotyped thought, and even racial prejudice.

Again the disagreements could be at least partly resolved by citing the varied nature of different types of imagery, individual differences in image control, and vividness, and different situations in which images might be used.

Empirical research into mental imagery presents many problems; it is very difficult, to say the least, to prove whether images themselves have functional significance or whether there is some underlying process, with images as a mere side-effect. What is perhaps important is that the "emergent properties of images do in fact play a functional role in cognition". (Kosslyn and Pomerantz, 1977). This role may be even wider than has already been suggested. For example, the use of imagery has been developed in a number of therapeutic techniques, ranging from systematic desensitization (Bower, 1970) and self-administered aversion (Berecz, 1976) to improving the client's self-concept (Shorr, 1975). King (1973) has even suggested "an image theory of classical conditioning", proposing that the conditional response is in fact a response made to an image of the unconditioned stimulus. With regard to the present study, however, one of the most important areas is the relationship between imagery and language, and this will now be discussed, with special reference to Paivio's dual coding hypothesis.

I6 IMAGERY, LANGUAGE AND THE DUAL CODING HYPOTHESIS

The dual coding hypothesis is very much associated with the work of Paivio, who has proposed it on a number of occasions (e.g. Paivio, 1969, 1970, 1972, 1974, 1975a, 1975b). The most detailed version is that described in his book Imagery and Verbal Processes (1971), where he says:-

The dual coding theory was not intended to be a complete theory of memory, and a more formal and detailed theoretical statement therefore remains to be developed.....The greatest deficiency of contemporary theoretical and empirical approaches to memory is that they have been concerned almost entirely with auditory - motor verbal memory.

In view of this apparent deficiency, the dual coding hypothesis proposes two functionally independent but interacting symbolic systems which are differentially available according to the nature of the information, the preferences of the subject, and the demands of the task. The image system specialises in processing synchronous or "parallel" information, and is more likely to be evoked by highly concrete materials; conversely, the verbal system specialises in processing sequential, linguistic and abstract information. Many studies have shown that pictures are generally better recalled than words, and that concrete words are better recalled than abstract ones, and the dual coding hypothesis deals with such findings in terms of "coding redundancy".

A picture can be processed by both image coding and verbal coding (in the form of labels); a concrete word may well also evoke an image, but an abstract word can normally be processed only by the verbal code. (In fact, some research does suggest that abstract words can evoke images of concrete exemplars, e.g. Bugelski, 1971; Anderson and McGaw, 1973). The superior recall of concrete items is thus generally explained by the dual coding hypothesis in quantitative rather than qualitative terms (e.g. Paivio and Csapo, 1973).

Several studies have supported the dual coding hypothesis. An often quoted finding is that of Begg and Paivio (1969), where subjects performing a recognition task tended to notice wording changes more easily than meaning changes in abstract sentences, but vice-versa in concrete sentences, suggesting that the two are stored in different forms. However, their results were contested by Johnson, Bransford, Nyberg and Cleary (1972) and Bransford and McCarrell (1974) who stress the factor of comprehensibility. The issue has also been discussed by Moeser (1974, 1975), who gives partial support to dual coding but suggests that Begg and Paivio's result could have been due to experimental design.

On the other hand there has been considerable experimental support for dual coding from a variety of other sources. Brain research has established that the two hemispheres appear to be specialised for functions very similar to the two systems proposed by dual coding. Although they are not entirely exclusive, the left hemisphere deals generally with verbal and sequential information, and the right with non-verbal and spatial information (e.g. Seamon and Gazzaniga, 1973; Hicks and Young, 1973; Van Tharp, 1972; Blakemore, 1976). The importance of cerebral laterality effects in thinking has been recognised to the extent that the UCLA Educator devoted a recent issue (1975, 17, number 1) to research on brain hemispheres and its implications for education. A second line of support for dual coding comes from interference studies like those already briefly mentioned (see 14 above). The implication here is that a perceptual task and a memory task interfere with each other if they both involve the same system, but not if the two different systems are used. Sasson and Fraisse (1972), for example, showed both retroactive interference and positive transfer effects occurring between pictures and concrete sentences, with abstract sentences being affected only by other abstract sentences, while Klee and Eysenck (1973) found differential effects on comprehension latencies for concrete and abstract sentences presented with visual and verbal interference.

Research on individual differences also offers slight support for dual coding, to the extent that some subjects appear to make preferential use of one system rather than the other. Earlier attempts to relate individual differences to psychological indicants like alpha rhythm blocking, or even breathing (e.g. Golla, Hutton and Grey Walter, 1943; Short, 1952) do not appear to have been taken up very often in recent research, but modality preferences and the independence of verbal and non-verbal abilities do offer at least indirect support for the dual coding hypothesis (Levin, Divine-Hawkins, Kerst and Guttman, 1974; Pressley, 1977).

Finally some studies which have deliberately manipulated stimulus attributes (pictures and words high and low in imagery value) and experimental instructions have also tended to support the idea of the additive effects of two relatively independent systems (e.g. Paivio and Csapo, 1973; Snodgrass, Wasser and Finkelstein, 1974; Odom and Nesbitt, 1974). Kosslyn, Holyoak and Huffman (1976), for example, report clustering effects according to processing mode rather than semantic content.

Apart from such experimental evidence, a few writers have proposed theories which appear very similar to dual coding. Gombrich (1972), for example, suggests that, although language is superior at dealing with abstracts and specific information, images are often better at conveying emotional reactions. He concludes that the "mutual support of language and image facilitates memorizing. The use of two independent channels, as it were, guarantees the ease of reconstruction". A more detailed theoretical approach is taken by Seymour (1973, 1976), who elaborated Morton's original "logogen" model to incorporate a separate non-verbal "inconogen" system. There is an obvious similarity here with dual coding, though Seymour also proposes "an overlying semantic system" to relate the lexical and pictorial elements. Finally, a theory involving a central distinction between simultaneous and successive syntheses has been proposed by Das and his associates (e.g. Kirby and Das, 1976), who maintain that their approach has wider explanatory powers than dual coding, and can be related to other theories, particularly Jensen's proposal of Level I and Level II abilities (Jarman, 1978), while at the same time remaining compatible with research supporting dual coding.

It would be misleading, however, to suggest that there has been no opposition to the dual coding hypothesis. The alternative proposed by Pylyshyn (1973), that propositional knowledge underlies both verbal and visual systems has already been mentioned. Anderson and Bower (1974) also agree that we may create images when confronted with certain problems, but claim that we operate on these images and derive information from them in a non-visual way. They conclude that "the perceptual descriptions stored are inferential conclusions; the raw texture of appearance which initially supported those conclusions are rarely stored in any detail". Experimental support for this view could be offered by Rosenberg and Simon (1977) who found evidence of cross-modal integration in a "false recognition" paradigm involving pictures and words, suggesting that there is a single underlying semantic structure (cf Seymour, above). Another experimental approach has involved comparing the performance of subjects instructed to use both verbal and non-verbal mediation with that of subjects instructed to use only one.

Mueller and Jablonski (1970) and Levin, Davidson, Wolf and Citron (1973) failed to find the superior recall predicted by the dual coding hypothesis. However, Wittrock and Goldberg (1975) point out that such studies are invalid because they rely on instructions and ignore the established mediation strategies of the subjects; the "elusive interaction" between word attributes and instructional set is hard to demonstrate because instructions can only enhance and not change, the coding process determined by the word attributes.

Emphasis on stimulus attributes has also been developed by some writers who regard the dual coding hypothesis as misleading, however. Rohwer's "elaboration hypothesis", for example, "explicitly eschews assertions about the modality of underlying processes, relegating the modality distinction to the domains of stimulus materials and peripheral processes". (Rohwer, 1973). The position adopted by Friedman and Bourne (1976) is in some ways similar, since they claim that different latencies for pictures and words are due simply to stimulus discriminability, and not to encoding differences. They prefer a "levels" approach like that of Craik and Lockhart (1972), "in which multiple codes or representations do not exist to be activated by the appropriate stimuli, but rather the stimuli themselves embody levels of information which are encoded and used as needed".

It thus appears that there are both theoretical positions and experimental results in opposition to dual coding. On the empirical level, however, the dual coding hypothesis continues to offer testable predictions which have been verified by research using a variety of approaches. On the theoretical level, it may not be possible to distinguish between imagery and verbal codes as functional in their own right, or as surface manifestations of some underlying propositional representation. The evidence outlined in I4 and I5 above does suggest that images have distinct properties not found in verbal mediation; moreover, these images (or the processes which cause them) appear to have important functions in a variety of tasks, including many that might appear to be purely verbal. The two experiments in the present study both involved children's language skills, though at two very different levels, and some of the experimental groups in each experiment used techniques involving the generation of images. While not resting entirely on the dual coding hypothesis, therefore, the study is very much concerned with the roles of imagery in children's performance of language tasks, and the developmental significance of imagery now needs to be discussed.

I7 IMAGERY AND COGNITIVE DEVELOPMENT

The widely accepted developmental theories of Piaget and Bruner both propose that imagery is vital in cognitive development, though they differ as to the nature and priority of images. Piaget's position with regard to images as internalised imitations dependent on the development of the operations, has already been mentioned (I4) but it would be wrong to imply that Piaget consequently minimises the importance of imagery:-

If one wishes to evoke in thought and some past perception, it is necessary to supplement the verbal system with a system of imaginal symbols. Without some semiotic means it would be impossible to think at all.

(Piaget and Inhelder, 1971)

Bruner (1964) and Bruner, Olver and Greenfield (1966) propose a theory of "instrumental conceptualism" involving "iconic representation" which develops from "enactive representation" and itself forms the basis for the final stage of "symbolic representation". The three may well all persist into adult life, but their order of appearance is constant. It will thus be apparent that, although they both accept the developmental importance of imagery, Piaget and Bruner have slightly different views as to its nature and functions. This difference has received experimental support from an interesting study by Harris (1976), who correlated scores on two of Piaget's tests of imagery with scores on two tests developed by Kuhlman (under Bruner's supervision).

Harris reported only a very low correlation with seven-year-olds, and none at all with ten-year-olds, suggesting that the two types of tests are measuring different things.

There has also been some controversy over the proposed priority of iconic representation due to several findings that young children seemed to prefer verbal mediation. Thus Milgram concluded that "children preferentially utilize the verbal mode of representation as early as age four and are more likely to give covert verbal descriptions to visual presentations than covert visualising responses to verbal representations". (Milgram, 1967 b). Similar conclusions are reached by Rohwer (1970), Davidson and Adams (1970) and Horvitz and Levin (1972), though Reese (1970), while accepting the experimental evidence, appears unwilling to consider the possibility that verbal memory could be better than visual memory in young children because it would conflict with Bruner's theory.

Conversely, there is experimental evidence that young children can make effective use of images, and that this ability emerges before the ability to use verbal mediation (e.g. Yuille and Catchpole, 1973, 1974). Similarly, Holyoak, Hogeterp and Yuille (1972) conclude that:-

Imagery is a preferred and effective mode of mediation in the paired-associate learning of pictures, and may be elicited by both interactive pictures and by concrete sentences. The capacity to benefit from both types of mnemonic is well developed in children by the age of five or older.

In view of such discrepancies, the exact role of imagery in cognitive development is not clear. Doob (1972) points out that many of the studies have used paired-associate learning, which is "obviously sequential rather than spatial in character" as well as being far removed from the everyday world of young children. A different line is suggested by Wolff and Levin (1972) who distinguish between the child's ability to use an image, provided perhaps by a picture, and the later development of his ability to generate an image for himself. Similarly, Reese (1970) considers several possible ways of reconciling the findings, including production deficiency and mediation deficiency (see I2 above), though he prefers the idea that children fail to "read" the relationships shown in a picture, which are explicitly stated in a corresponding sentence. A detailed survey of imagery research from a development perspective has recently been provided by Pressley (1977) who suggests that imagery effects may be task specific; for example, age has been shown to interact with stimulus presentation mode in paired associate learning, but not in recognition memory or free recall. He concludes:-

The ability to increase learning by self-produced internal visual elaboration is truly a developmental phenomenon. Nursery School children cannot produce such internal elaborations when simply told to do so. By six or seven years of age, children can produce internal elaborations for paired-associates. Children a few years older (eight years old) can gainfully apply the visual imagery strategy to more complex tasks, such as prose learning.

In fact, several writers agree that there is apparently a drastic change in the child's use of imagery between about five and eight years of age. For Piaget, the appearance of anticipatory imagery is of course dependent on the development of concrete operations. This change is supported by Paivio (1970) who describes the "quantum leap" in the child's ability to perform symbolic transformations, while Wolff and Levin (1972) describe a similar rapid development in five to eight year olds' use of imagery. But whether one accepts the idea of "stages" of development or not it is clear that imagery is very much involved in children's cognitive development.

As with several previous sections, a general idea of the importance of imagery thus emerges, in spite of some discrepancies in experimental findings which can be at least partly reconciled by the variety of conditions and subjects. It would, therefore seem reasonable to expect some researchers in the field to suggest "educational implications" even if, as has already been suggested, they often appear almost as an afterthought, and some of these implications can now be considered.

I8 SOME EDUCATIONAL IMPLICATIONS

Several writers have in fact proposed such implications, in very general terms. The very last page or so of Paivio's influential book (1971), for example, says that "probably the most important practical outcome of research on imagery and verbal processes will be in relation to problems of education", and goes on to suggest, for example, that it is generally better to provide a concrete example before a general rule, rather than vice-versa, while admitting that research on paired-associate learning may well have little relevance for the classroom. Rohwer (1970) also describes some very general implications, roughly divided into two sorts:-

The first class of recommendations concerns the manner in which information can best be presented to children in order to foster acquisition and retention.....The second class of recommendations has to do with the kinds of activities children should be taught to engage in so as to increase their own powers of learning. In brief, they should be taught the use of both verbal and visual kinds of elaborative strategies.

In the most general terms, then, research into strategies, of which imagery research is an increasingly important part, offers a way of looking at learning failure and its improvement (as suggested in I2 above) through the optimum presentation of materials and the actual teaching of strategies.

At a descriptive level, the research does suggest that imagery plays an important part in cognitive development. The Bullock report asked all teachers to consider the linguistic demands of their subjects; perhaps they should also consider the role of non-verbal processes in the curriculum. Kuhlman, for example (in Bruner, Olver and Greenfield, 1966) suggested that the ability to form images, though an advantage in the early years of schooling, becomes a disadvantage as the Secondary School comes to demand more and more symbolic skills. Luria (1969) compares the plight of "S" with that of a child struggling to cope with abstract concepts:-

How odd and yet how familiar these experiences are. They are inevitable for any adolescent who, having grown used to thinking in terms of graphic images, suddenly finds there is a world of abstract ideas to be mastered.....These concepts exist and are taught to us in school. Yet how can we picture them in our minds? And if it is impossible to imagine them, what do they mean?

It may even be possible to relate the ability to form images to success in various school subjects, in a more specific way, though the ability to form vivid images must not be equated with the ability to use images to solve problems. Paivio (1971) describes an unpublished study carried out with Rogers, which applied factor analysis to students' responses to a questionnaire and their success in various subjects, concluding that Geography, for example, demanded the highest "spatial ability".

On the other hand, Taylor (1960) found no evidence that spatial ability was important in Geography, when success in the subject was assessed by a written school examination. This is clearly an interesting approach, but very little work has been done yet, perhaps in part because the use of individual differences as the independent variable has been generally less successful than the manipulation of materials and instructional tasks (e.g. Paivio, 1971; Marks, 1972).

The recognition of individual differences and their consideration in teaching methods have, however, been described by several writers. Hollenberg (1970), for example concludes that "research is needed to determine the interactions of the child's mode of thought with types of learning materials". Levin and his associates (e.g. Levin, Rohwer and Cleary, 1971; Levin, Divine-Hawkins, Kerst and Guttman, 1974) have developed this idea into "learner types" based on individuals' visual or verbal preferences, and conclude that imagery instructions are not universally helpful, but interact with the capabilities of the learner. Their finding that, generally speaking, subjects who are able to use both verbal and visual mediation perform better than subjects who tend to use only one is also supported by Potter (1977).

Other writers have suggested other "types" of their own. Goldman (1972), for example, talks about "logical" and "mnemonic" subjects; Pask (1976) describes "wholist" and "serialist" learners, and points out that each learns much better if the learning programme matches his style. Reinert (1976) has developed ELSIE (Edmonds Learning Style Identification Exercise) to identify four learning styles, two of which involve mental imagery, and he goes on to suggest that teachers could use the exercise to identify the predominant style of a whole group of learners and teach them accordingly. On a more specific level, an unpublished study by Hartnett (described by Krashen, 1975) even managed to relate hemispheric dominance to subjects' preferred methods of learning a foreign language.

There are obvious parallels here with the growing body of research on cognitive styles, and perhaps equally obvious dangers in the approach. It is almost easier now for a researcher to decide on his own "learner styles" than to try to master and relate the large number of them proposed by others. For example, how far are Pask's "wholist" and "serialist" styles related to the "synchronous" and "sequential" processing proposed by the dual coding hypothesis?

The other danger is that it becomes very tempting, given the dichotomies that most writers propose, to label one as "better" than the other - there is perhaps already the feeling in research on cognitive style that "divergers" are rather superior to "convergers" or that "field independence" is a desirable trait, regardless of the circumstances.

On the positive side, however, imagery research has been one factor in a general trend to get teachers to recognise and take account of individual differences in learning strategies. Thus Yanoff (1972) for example proposes "multisensory stimulation in the classroom to allow each pupil to use his strongest perceptual mode - auditory strength might suggest a "phonics" approach in learning to read, while the visually - orientated child might prefer a "whole word" method. The teacher should also be aware that his own preferred strategies, which have presumably brought him success in his chosen subject, may be very different from those of his struggling pupils. Broadbent (1975) sums up:-

Each kind of individual learns better with material designed on his own principles than he does on material arranged in accordance with the opposite principle.....Different teaching methods are appropriate for different people: it is particularly important to realise this, because each of us finds it intuitively unreasonable that a method of presentation he himself finds hard to understand may be positively easier for somebody else.

Study technique is even more widely neglected than teaching technique. The point of higher education, in contrast to schooling, is supposed to be that it involves independent study. Yet few students are given any preparation in its techniques. Using a book, for instance.....Students need to learn how to skim, how to select, how to use an index.

There is some opposition to the idea of actually teaching strategies, however. Smith (1975), for example, points out that children's strategies are inadequate only when they do not understand what they are doing, and concludes that "the solution may be not to try to inculcate different learning strategies in a struggling child, but to make the situation more meaningful to him by providing more relevant information." A more unorthodox solution is offered by Rohwer, (1971), whose own research has emphasised the lack of spontaneous strategies in young children. Rohwer concludes that early childhood is an inefficient time to teach skills that an adolescent could learn quickly, and suggests that systematic schooling could be postponed, or even dropped altogether.....

Finally, a particular implication of imagery research has been the proposal, made by several writers, that there should be more emphasis on "iconic education". It does seem that formal education tends to emphasise "left hemisphere" skills at the expense of the visual and intuitive thought processes which are the special concern of the right hemisphere,

(Van Tharp, 1972; Blakemore, 1976). Bogen (1975) points out that society's demands exert pressure on schools so that "life success" is judged in terms of the functions of the left hemisphere, which may seem irrelevant to those in whom it is less dominant, and he concludes that "the entire student body is being educated lopsidedly". Consequently, when students do use imagery, it may almost be with a sense of embarrassment! Huttenlocher (1977) claims that her subjects "regard their spatial imagery as an intellectual crutch, unseemly for the solution of formal reasoning problems". (There are often similar reactions to the use of mnemonics - a point raised in Chapter II).

Similarly, the use of pictures in schools, particularly at the Secondary level, tends to be limited to "enrichment" or illustration of a point made verbally, in spite of the claims of writers like Huggins and Entwisle (1974) that "iconic communication" can actually teach rather than simply illustrate. I. A. Richards (1968) has also emphasised the role of actions and pictures in education, pointing out the "neglected problem of pictorial literacy". There are potential dangers in this enthusiasm, however.

Actions and pictures may convey some ideas with great success, and have considerable emotional impact, but they lack language's ability to deal with abstract ideas, and to transcend time and space, while at the same time they may be open to ambiguous interpretation.



Figure 2 "Non-Verbal Communication" (From Gombrich, 1972)

Simon (1972) attacks some of I. A. Richards' educational ideas for these very reasons, and a similar point is nicely made in the illustration from Gombrich (1972). It is perhaps interesting that a picture is here apparently conveying quite a subtle and abstract idea very successfully; but it can do so only by employing the accepted symbolic "bubbles" to represent thought. Moreover, there are many different types of picture, just as there are many different kinds of image, ranging from the exact recordings of photographs to the schematic representations of line drawings, and not all are equally beneficial in learning. Dwyer (1976), for example, found that a very simple line drawing resulted in better understanding and transfer than a photograph. The previous emphasis on individual differences, and the parallels between "visual representations" in the sense of mental images and "visual representations" in the sense of external pictures, surely imply that the optimum use of either in education will almost always be in conjunction with language rather than instead of it.

In experimental psychology, mental imagery has re-emerged as a respectable subject for study and, in spite of problems about its exact nature and controversy about its functions, the recent flood of research does seem to support its importance in cognitive development and learning. The existence of imagery may well be universal, but its potential value in education probably involves a need to train children in its use. (Doob, 1972; Marks, 1972). For Bruner, all three modes of representation develop by the internalisation of the ways of acting, imaging or symbolising that exist as amplifiers of man's own powers in his culture; yet at the moment, education as a powerful agent of cultural transmission tends to neglect and even suppress imagery in favour of symbolic verbal representation. Bruner (1966) concludes - "I do not think we have begun to scratch the surface of training in visualisation - whether related to the arts, to science, or simply to the pleasures of viewing our environments more richly". An attempt to show that strategies involving visual imagery can improve children's performance in some language tasks which they are actually expected to perform at school is perhaps a small step in this direction.

CHAPTER II

SUMMARY

Chapter two falls fairly clearly into three parts, the first two being concerned with modern language teaching and the third with the possible application of a technique derived from experimental psychological research.

The first part discusses some of the more important issues involved in the aims, content and methods of teaching modern languages in schools, with particular emphasis on one controversial issue - vocabulary learning.

The second part presents an exploratory questionnaire survey of actual classroom practices in this area, suggesting that many teachers still do expect their pupils to learn vocabulary, usually with little advice on how to do it, and in spite of the theoretical opposition to it from the currently popular audio - visual approaches.

The third part looks at some of the issues in one branch of research on learning strategies and imagery - mnemonics, and suggests that a technique derived from this research could be applied to children's vocabulary learning in the classroom.

II 1 SOME CURRENT ISSUES IN MODERN LANGUAGES

(a) Introduction

"There is something rotten in the state of language teaching". These were Professor Harry Ree's opening words to the 1972 A.V.L.A. conference, and even a brief examination of recent writings on foreign language teaching suggests that there is growing concern about the subject's aims, content, methods and status in schools. Modern languages still tend to be seen as a "respectable" academic subject area for various reasons - important educationally because of the Grammar School tradition and the development, albeit controversial, of Primary School French, and important outside the educational system because of growing links with Europe in particular (Whitfield, 1975). On the other hand, although a strong case can be made for the retention of modern languages in the curriculum, at least for some pupils, this enthusiasm does not appear to be shared by the pupils themselves. In 1974, Duckworth and Entwistle published a survey of the attitudes of six hundred boys and girls to various school subjects, using a repertory grid technique. On the basis of interviews, they isolated four dimensions - Interest, Difficulty, Freedom and Social Benefit. When twelve-year-olds were asked to rate nine school subjects, French emerged very badly, particularly if results for Latin, now very much a minority subject, are not counted.

Both boys and girls rated French as the least interesting and least "free" subject; boys found it the most difficult and girls the second most difficult (after physics); both boys and girls saw it as very low in social benefit, with only history as less useful.

Bearing this in mind, it is perhaps not surprising that a recent H.M.I. report on modern languages (1977) expressed some concern about what they found in a survey of eighty-three Comprehensive Schools. They report that the quality of the work they saw was often poor, and that many children seemed to be under-achieving; in spite of the growing stress on oral work, many pupils "were unable to produce other than inadequate or largely unusable statements in French".

This somewhat gloomy picture suggests at the very least, that modern languages represent a worthwhile area of research for educationalists in various fields, including the psychology of education. Hirst (1967) offers a useful way of approaching the complexities of a subject by dividing it into three broad areas - educational objectives, content and methods. Although this distinction may appear rather arbitrary, and not all writers would agree with him, Hirst in fact differentiates quite sharply between the three, and thus does at least suggest a useful framework for looking at curriculum problems which might well include some of those in modern languages.

For example, Hirst describes the "traditional" curriculum's failure to distinguish between objectives and content, contrasting it with the "progressive" curriculum's failure to distinguish between objectives and methods. Psychologists have traditionally been concerned only marginally with curricular objectives and content, limiting their contributions here largely to the implications of cognitive development. Thus Bruner (1966) saw psychology's role in education as outlining "the possible" in terms of aims and content, while being centrally concerned with methods of learning and teaching through the eventual development of a theory of instruction. However, since objectives and content logically precede methods, it is proposed to look very briefly at some current issues in the objectives and content of modern languages in schools before examining in more detail the possible contributions of psychology to the actual methods of learning and teaching.

(b) Objectives and Content

Although objectives should logically be formulated first, this is not always the case in practice, for a variety of reasons. The 1977 H.M.I. report suggests that the best schools visited had clear objectives and expectations of their pupils, depending on their ability, whereas the poor schools appeared to have formulated few aims at all.

However, they point out that objectives have almost certainly changed since modern languages were first introduced "to provide the pupil with mental training and to develop habits of accuracy". This traditional objective in foreign language teaching involved the development of skills that would be of value in other areas; although there is no hard evidence that learning a second language even develops native language ability (Whitfield, 1975), let alone anything else. Burstall (1974) makes this point strongly, and also distinguishes between "instrumental" and "integrative" objectives, the former using French largely as a qualification to go on and do something else, and the latter trying to involve the pupil in the French way of life. In the most general terms, objectives thus probably involve widening pupils' general interests as well as developing language proficiency. For example, the three stated aims of the Nuffield Introductory French Course En Avant (1972) are:-

- (a) To teach the pupils to understand, speak, read and write French rather than to teach them about French.
- (b) To provide a simple introduction to French life.
- (c) To contribute to the pupils' general educational experience.

Similarly, the Longman's French course, by far the most widely used in Secondary Schools, aims "to enable the students to understand and use the French language effectively and introduce them to the people of France and their way of life", and "to introduce simple idiomatic French in lively situations related to the interests of the pupils". (Moore and Antrobus, 1973).

But which pupils? One controversial issue in this area has been the introduction of a foreign language for almost all pupils with the development of Comprehensive Schools. According to the H.M.I. survey, about 81% of all first formers now study French, with 5% and 4% studying Spanish and German respectively. This development has been of central importance to foreign language teachers, and objectives are having to be reformulated accordingly. Williams (1976) for example, feels that many teachers have clung to aims now far beyond the reach of the majority of their pupils. Similarly, although both the GCE and CSE examinations now seem to have changed their objectives, with more stress on oral work and comprehension, and much less on traditional grammar and translation exercises, there is still some feeling that they are unrealistic in both aims and content. Wilby (1978) for example, describes the experimental introduction of "survival - French", pointing out that many of the terms expected at 'O' level would be useless in practical situations in France.

Similarly Page (1978) has developed a series of eight graded tests, not geared to age, ranging from basics to beyond 'A' level. Level two, for example, would involve "survival abilities", and most pupils should be able to reach this or the next level.

Another aspect of this question "which pupils?" is the controversial area of Primary School French, where much discussion was sparked off by an influential N.F.E.R. report, (Burstall et al, 1974). After a ten year survey of 18,000 pupils in 125 schools, Burstall concluded that "pupils taught French from the age of eight do not show any substantial gains compared with those who were taught French from the age of eleven". She also found that girls scored significantly higher than boys on all tests measuring achievement in French, and that high achievement was related to other factors like socio-economic status, the size and location of the school and the attitude of the Head. In relation to the traditional objective of developing general skills mentioned earlier, it must also be noted that "the introduction of French into the Primary School curriculum did not exert any significant influence on children's other attainments", and even appeared to produce a negative attitude towards other languages in Secondary School. Though there is still a certain amount of controversy, the report appears to have been very influential, and has obviously cast serious doubts on the value of Primary School French.

Finally, it has been suggested that foreign languages are fundamentally different from many other subjects in the curriculum, to the extent that the new student has virtually no knowledge of the subject to begin with, (Ree, 1972). The most obvious contrast is with English, where almost all the pupils already have enough experience to enable them at least to survive in the outside world. Ree goes on to argue that the French teacher thus remains much more a figure of Authority on his subject and that "new methods" are not suited to foreign language learning. Burstall found that socio-economic status was an important factor in achievement in French, and Ree agrees that the subject may seem completely alien, particularly to working-class children. He concludes that "the modern linguist is seen by his colleagues as a pedlar of a skill which chiefly benefits the middle-class child". The development of "survival" French and the growth of European studies might both be seen as attempts to make modern languages more relevant to pupils, particularly those who would not have been expected to learn a second language in the days before Comprehensive Schools.

(c) Methods

According to the 1977 H.M.I. report, the post-war years saw an end to the controversy between traditional "grammar-translation" method, which directly related the second language to the native language, and the "direct" method, which completely avoided native language.

The compromise "oral" method which resulted formed the basis for the "audio-visual" approach which appeared in a large number of schools during the 1960's. These newer methods owed a lot, not only to the decline of the influence of S-R theory, but to exciting developments in our knowledge of how children first acquire language and to new approaches in linguistics, particularly the very influential work of Chomsky. The parallel with native language learning is often drawn (e.g. Mueller, 1974; Hakatu and Cancino, 1977), and the learner is seen as actively formulating rules and hypotheses rather than passively receiving individual S-R units. The four skills of listening, speaking, reading and writing are developed in a rich context of real objects, filmstrips and pictures, and an emphasis on the spoken word, often using a tape or a language laboratory. New material is introduced in this visual and verbal context of what is already familiar rather than being paired with native language translations.

However, according to some writers, there has been a recent change of emphasis with a growth of interest in pupils' strategies, probably related to developments in psychology noted in chapter one. There appears to be a growing concern, not just with materials and presentation, but with what pupils actually do when trying to learn a second language.

Pimsleur and Quinn (1971), for example, discuss the realisation that the supposedly passive "receptive" skills of listening and reading are in fact very complex, and demands a considerable contribution from the learner. In the same volume, Selinker stresses the importance of the student's strategies, and describes the "interlanguage" - his attempts, based on his native language, to produce utterances in the new or "target" language. Anderson (1970) has shown how different presentations of new material in a foreign language can affect the attention and learning strategies of students. Ott, Blake and Butler (1976) point out that statements such as "Vocabulary must be learned in context" may have implications for presentation of materials, but they say little about what the learner might actually do with the presented material.

This interest in learner strategies can be seen as having important implications for the audio-visual method, and its reluctance to use native language except as a last resort. Those who had supported the "direct" method had claimed that it avoided the unnecessary and interfering effects of the learner's native language (e.g. Lambert, 1973; Rivers, 1968). Rivers describes how "even in reading, students were encouraged to forge this direct bond between the printed word and their understanding of it without passing through an intermediate stage of translation".

Such writers frequently drew a parallel between the acquisition of native language and learning a second language, emphasizing the importance of context and the direct link between language and concepts, and to some extent, these views have been taken up by more recent supporters of audio-visual methods. Thus Winitz and Reeds (1973) and Mueller (1974), for example, strongly oppose the use of native language as a mediator and prefer methods in which the learner is exposed to pictures or objects to be directly related to the new language, just as he was when he first learned his native language. Similarly the Nuffield French Course En Avant (1972) states that:-

Particularly in the early stages, all new sentence patterns and items of vocabulary are presented in a meaningful context using the recorded voices of native French speakers. Visual aids are used to convey the meaning, thus avoiding the need for translation.

This approach may have encouraged lively teaching methods and rich, interesting materials, but it has two important weaknesses - the parallel between learning native and second languages, and the neglect of pupil strategies.

Firstly, the parallel between learning one's native language and learning a second language is highly misleading, both in social and cognitive terms. Social factors have been recognised as vital in language learning, yet the conditions in which a baby learns to talk are obviously entirely different from those in a French class. Enthusiasts might argue in favour of "total immersion" methods, where everything on the timetable is taught in French, but this is only a very partial solution. What about the one-to-one warm loving relationship between mother and baby, the use of "baby-talk" and total lack of any formal teaching or knowledge from the native language speaker on whom the baby depends for everything throughout its waking hours?

Even if one argues in favour of re-creating just some of the social conditions in a context appropriate to the age of the learner, the cognitive conditions under which native and second languages are learned are entirely different. In the baby, language develops along with concepts, and the two are closely bound together, whatever view one takes of the exact relationship between them; the eleven-year-old beginning a second language already has a large store of concepts and strategies inextricably tied in with his native language, and it is pointless to expect this language not to intrude, especially if teachers are dealing with ideas and situations that are already familiar.

Even Lambert (1963), who generally favoured "direct" links between concepts and the new language, had to admit "that students studying under a direct method utilize the semantic features of both their languages and permit the two to interact and that this tendency towards linguistic interdependency apparently assists students in acquiring their second language".

Similarly Carroll (1963), Preibusch and Zander (1971) and Vielau (1976) emphasise the unavoidable influence of native language, while Roger Brown (1973) took the unusual step of joining a "total immersion" Berlitz class in Japanese, to look at the whole thing from the learner's point of view. The Berlitz method avoids native language completely, and Brown's experiences are worth quoting at some length.

One long morning my teacher tried to put across three verbs, Kimasu, Yukimasu and Kaerimasu, with the aid of paper and pencil drawings of pathways and persons and loci, and by much moving of herself and me - uncomprehendingly passive as a patient in a hospital. But I could not grasp the concepts. I feel Mr. Berlitz would have suffered no great dishonour if she had said to me that the concepts in question sometimes go by the names come, go, and return.

In view of what many psychologists say about learning, Brown's words "uncomprehendingly passive" are perhaps particularly significant, and if a Harvard Professor and world authority on language learning had problems with such methods, it is perhaps not surprising that eleven-year-olds of average ability run into difficulties!

It would also seem that, apart from causing problems (Brown mentions "reference" and "segmentation"), attempts to avoid native language by using actions and pictures are doomed to fail anyway, since most learners spontaneously supply their own verbal interpretations. Spontaneous labelling is more likely as the age of the subjects increases (Wilder and Levin, 1973; Pressley, 1977), though quite young children may well label pictures for themselves (Milgram, 1967; Murray and House, 1976). Similarly, the research discussed in the first chapter in support of the dual coding hypothesis would predict this, and there is experimental evidence that most subjects label pictures which are provided in foreign language learning materials. Jenkins, Neale and Deno (1967), for example, describe the futility of presenting a picture only, in the hope that the learner will not encode it verbally, but will form a direct link with the new word or construction to be learned. Hammerly (1974) found that, when presented with a new Russian or German word with an appropriate picture, about 70% of his subjects thought immediately of a native language label. The ambiguity of pictures has already been mentioned, and some of the subjects' labels were either rather tentative or actually wrong - about 75% were both confident and correct with "bird", but only 40% with "deer".

Hammerly concludes by describing "the inadvisability and the futility of relying only on pictures to convey meanings to second language students: (1) unless the native language is used, the meaning may not be clear and (2) the learner at this stage tends to think in native language anyway". Similarly, Cole (1976), discusses the use of pictures in language instruction and concludes that "it is naive to assume that in structured audio-visual courses a given utterance can be matched with a single composite non-coded picture that conveys its 'meaning' more adequately and effectively than direct translation". In a broader sense, the attempt to avoid native language by the use of pictures involves a failure to distinguish between the "nominal stimulus" presented by the teacher and the "functional stimulus" as perceived by the learner (cf Montague, 1972).

A strong case can thus be made for the explicit use of native language in at least the early stages of second language learning, but this obviously does not imply that pictures (or real objects) should never be used. The generally superior recall of pictures and their advantages and disadvantages compared with language, the dual coding hypothesis and the importance of individual differences were all discussed in chapter one, and all point to the value of using both pictures and native language in conjunction, each contributing something different to the learning situation.

Vocabulary Learning

One aspect of the "traditional" or "grammar-translation" curriculum which took a particularly heavy beating from the supporters of the "direct" and audio-visual approaches was vocabulary learning, and much of what has been said already refers directly or indirectly to it. The task of learning by heart individual words and their meanings, isolated from any context, appeared to be one of the most direct applications of S-R theory, and the only way it could be done appeared to be by the much criticised use of rote repetition. Rivers (1968) describes a typical group of students who have been given this "boring chore" (usually for homework) and are now in class, "babbling over to themselves" long lists of words without enthusiasm or comprehension, and he goes on to discuss how audio-visual methods tend in contrast to minimise the importance of vocabulary, especially in the early stages.



Unfortunately, it has to be admitted that vocabulary is a vital part of learning a second language. Steinführer (1973) found that successful students compiled their own vocabulary lists, whereas poor ones did not, and although this evidence is only correlational, he concludes that "intensive learning by heart of isolated lexical items is a first step which cannot be omitted. Without it, the second step, the use of vocabulary in a context, is not possible". Rapid development of vocabulary may also be important in enabling the learner to go beyond the course of materials. Atkinson and Raugh (1975) comment that:-

Many foreign language instructors believe that the major obstacle to successful instruction is not learning the grammar of a language, but in acquiring a vocabulary sufficient to engage in spontaneous conversation and read materials other than text books.

Similarly, Harrington (1972) claims that most students are keen to widen their vocabulary because "communication is better achieved with a good knowledge of vocabulary and a less sound knowledge of structures. There has been insufficient scientific investigation on the teaching of vocabulary". Moreover, even if the teacher tries to avoid native language, it seems likely that most students will allow it to intrude and will covertly attempt to learn vocabulary in a disgracefully "traditional" way. (e.g. Preibusch and Zander, 1971; Lubke, 1972).

At least two recent reports (Williams, 1976; H.M.I. report, 1977) have noted a possible swing away from audio-visual methods and the first suggestions of a return to "traditional" language teaching, but such a return is certainly not being advocated here. What has been attempted so far is to show that there are several controversial areas in modern language teaching, not only in its objectives and content, but also in its methods, which are particularly of interest to those involved in the psychology of education. Some recent developments in psychology suggest that the reluctance of the audio-visual method's supporters to use native language except as a last resort, and their emphasis on similarities between first and second language learning, are misplaced. Two further steps are now proposed: The gap between teachers and those doing research in education remains wide, as the first chapter suggested. In theory audio-visual methods minimise the importance of vocabulary learning, though an equally strong theoretical case can be made for its use. However, little appears to be known about what teachers actually do in practice. If they never expect their pupils to "learn vocabulary" anyway, there is little point in examining methods that would help them to do it more successfully.

If, on the other hand, children still are expected to learn lists of vocabulary, in spite of what the advocates of audio-visual methods say, then any help that psychology could provide would at least be relevant, if not actually welcome. Consequently, the next part of the chapter describes an exploratory survey of some current practices in French, particularly the conditions under which it takes place, if at all.

The final section of the chapter consists of a brief survey of the main issues in one branch of research on learning strategies and imagery - the use of mnemonics. It appears that a technique derived from experimental research here could perhaps be applied to children's vocabulary learning, though, this possibility has never been examined. This section forms the specific theoretical background for the experiments to be described in the third chapter.

II 2 A SURVEY OF CLASSROOM PRACTICES

(a) Preparation and Pilot Version of a Questionnaire

Given that the survey was to be of an exploratory nature only, and was to form a relatively minor part of the total data, it was felt that a short questionnaire would be adequate. Observation of a large number of lessons would obviously have yielded much more information, but would also have involved the well documented problems of observer bias and non-typical lessons. The latter point was particularly relevant because of suspicions that tasks like learning vocabulary and doing comprehension exercises might be seen by some teachers as rather "traditional", and perhaps not the sort of activity to be included in an observed lesson. A short, simple questionnaire administered to children from several schools, and not to their teachers, could avoid this problem. Such a survey obviously samples children's perceptions of the classroom rather than what "really" goes on, but bias also enters into the reports of researchers or teachers, and the children's perceptions are perhaps the most relevant here.

The first task was to decide on the topics to be looked at, and to formulate suitable questions, largely in multiple-choice form for ease and objectivity of scoring. The first part was to cover French lessons, and the second English lessons, (the latter being discussed in Chapter Four). Taking French separately, the following topics emerged:-

- (i) Roughly what proportion of the children in the sample were expected to learn French vocabulary regularly, or occasionally?
- (ii) How was such learning carried out - how many words were given, were the children tested and, in particular, how much direct help was given in methods of learning the words?
- (iii) Audio and audio-visual courses do not encourage vocabulary learning in this manner. Roughly how many children in the sample were taught at least partly by these methods, and what proportion of them were still expected to learn vocabulary?
- (iv) The original intention had also been to ask about the strategies used by the children, but the pilot version of the questionnaire (see below) suggested that this would be better covered by the more detailed experimental study, relating strategies to specific items.

A number of questions were then formulated and discussed with three members of the University staff who were experienced in questionnaire work, resulting in several changes in wording and layout. Eleven questions (along with fourteen on English) were next piloted on a small group of twenty-seven girls and thirteen boys, (see below for details on the selection of this and the final sample).

Their responses suggested that the detailed questions on strategies had been rather difficult to answer without actual examples, and those questions were withdrawn. The pilot study also showed that one or two points needed clarifying, and that it might be the best policy to read the questions aloud one at a time as the subjects filled in their answers. The pilot study also suggested the potential value of a section at the end where the children were free to comment. With these points in mind, a final version was printed with a very short introduction, nine questions on French, ten on English, and a space for comments at the end. (Appendix A shows the final version).

(b) Subjects

From the outset, it was felt very important that the children should not be asked the questions in school, where the classroom setting and the presence of the teacher might well affect the answers, and the whole thing would almost certainly be seen by the children as directly associated with the school itself.

Confidentiality was much more likely outside the classroom. Moreover, the exercise of asking children to comment on their lessons could understandably be seen as a potential embarrassment, or even a threat, by the staff, and it was felt important to avoid any possible naming of either the pupils or their teachers.

It was therefore proposed to draw both the pilot and the main sample from children meeting in various groups outside school - youth clubs, Scouts and Guides. An immediate objection is thus that the sample is not representative of the population as a whole - children who join such groups may not be typical. Three points can be made in the face of such an objection:

- (1) The practical advantages were considerable, and meant that groups of children could be given the questionnaire outside school, but still under reasonably controlled conditions. The alternative of trying to question individual children (at home? in the street?) would involve many problems, and the sample could end up being equally unrepresentative unless great care was taken.

- (ii) Particular objections to the use of Scouts and Guides may well be founded on an outdated "image", since these organisations have become more aware of their position in recent years, and now offer a much wider range of activities and hobbies in the hope of attracting children from a variety of backgrounds (Brooks, 1972). Binyon (1970) concluded that "such a large voluntary organisation as the Scouts is bound to vary from district to district. The image might be cosy and middle class in Bracknell or aggressively working class in Hull". Care was therefore taken to use groups from a variety of areas - rural and village, overspill housing estate, town and city.
- (iii) The aim of the questionnaire was to give an impression of children's views of current teaching practices in two areas of the curriculum. Only three questions, asking how difficult the children found certain tasks, and the final section asking for comments, might reasonably be expected to be strongly affected by personality or social background.

In a sense, the true "sample" was made up of the teachers whose methods were being described, although it is always possible that children who join groups are more likely to under-estimate the amount of time spent listening to tape recorders, in contrast with children who do not join groups (or vice-versa).

The final version of the questionnaire was given to a small number of pupils from thirty-nine different schools. 155 children answered in all. Apart from 10 ten-year-olds and 3 fourteen-year-olds, the children were all in the 11 - 13 age group. (Mean age 11.6, modal age 12).

(c) Procedure and Results

The questionnaire was administered to the children in small groups, after a short informal introduction in which E explained that he was carrying out a survey, and that the questions were in no way a "test".

The subjects were asked to fill in their answers without looking at those of their neighbours and it was pointed out that there were no "right" or "wrong" answers. E also asked the children to say if any of them were taught French or English by the same teacher as somebody else in the group, and their papers were marked accordingly, to be compared later. They were also asked to write their age on the top of the first sheet, but it was stressed that their names and the names of their teachers were not required.

E then read out the short introduction:-

The questions are about things you did at school this year. If you don't know what to put, just write 'don't know' by the question. If there's a choice of answers, tick the best one.

Questions were related to the present year only, in order to avoid problems of recall and possible confusions due to having been taught by several teachers over the years. The questionnaires were all administered during April and May, so that the children had at least six months' experience to draw on. It was also pointed out that "this year" meant since October, not January.

E then read through each question slowly, and the children filled in their answers. Any problems were explained as indicated next to each question below, and the proportion of children giving each answer is also given, to the nearest whole percent.

- (1) Name of school. (E asked for details of the type of school here, and the sample broke down into five categories).

Junior	10%	(5 schools)
Middle	19%	(9 schools)
High	2%	(3 schools)
Comprehensive	67%	(19 schools)
Grammar/Independent	2%	(3 schools)

The sample thus has a heavy bias towards Comprehensive Schools, and cannot really be used to make comparisons between, say, Grammar and High Schools. However, the proportions are reasonably representative of children in the 11 - 13 age range since a few of them would be in their final year at Junior School, but most would have gone on to Comprehensives.

Similarly, where a Middle School system operates, most children in this age range would be in the Middle School, with a few of the older ones having moved on to High School.

(2) Did you do any French at school? (i.e. this year, since September).

Yes	87%
No	13%

Of the twenty subjects who did not take French, five were at Junior School, one was at a Middle School, and fourteen were at Comprehensives. All the subjects attending Grammar or High Schools did French. In relation to the discussion on Primary French in the previous section, it is perhaps interesting that only one of the five Primary Schools appeared to have French on the curriculum, though the sample is obviously far too small to make any generalizations. The percentages given in the following questions refer only to those children studying French.

- (3) How many years' French have you done? (E instructed the children to count this year as one full year, e.g. if they'd started last September, they should put "one year").

One year	39%
Two years	28%
Three years	10%
Four years	8%
Five years	8%
Six years	4%
Seven years	2%

About two thirds of the subjects had thus been studying French for less than two full years. Of those with more than four years' experience, all but one were at Comprehensive Schools, and the remaining one was at High School. All eight of those with six or seven years' experience must have been taught French at Junior School, since none of the sample was over fifteen years old.

- (4) In French lessons, did you ever have film strips or slides? (This was an attempt to assess the "visual" element of the lessons, and four alternative answers were given).

Often	(Say every few weeks, or more often)	10%
Sometimes	(Say once or twice per term)	10%
Rarely	(Say only once or twice ever)	18%
Never		63%

The descriptions of each category were given in an attempt to minimise individual differences in interpreting words like "often" and "sometimes". Bearing in mind that these answers may not be completely accurate, it would still appear that the slides or filmstrips which ideally form a central part of audio-visual courses were not in common use in many of the schools. Interestingly, all the Grammar School pupils reported that they used slides "often", though the sample is obviously too small to make generalizations, and a comparison of different practices in different types of school is beyond the scope of this exploratory survey.

Many other questions could also have been asked about the use of visual aids, including devices like flash cards or flannel graphs, but the main interest here was in vocabulary learning, since this was the area to be examined in the experimental study. The two courses most widely used (Longman's and En Avant) both provide filmstrips whose use is strongly recommended, though pupils could well be following one of these audio-visual courses and still never use filmstrips or slides.

- (5) Did you ever listen to tapes in French, or use a language laboratory? (Answers as question four).

Often	39%
Sometimes	21%
Rarely	23%
Never	17%

Responses here suggest that use of audio equipment is rather more common than use of slides or filmstrips, though the difference could well be due to practical difficulties (e.g. blacking out a room or shortage of equipment) as much as to deliberate policy.

Table one shows the two sets of answers broken down into the four categories for each.

Use of Slides (Question 4)	Use of Tapes (Question 5)			
	Often	Sometimes	Rarely	Never
Often	10	2	1	0
Sometimes	5	4	3	1
Rarely	12	4	4	4
Never	25	19	23	18

Table 1 Number of responses in each category, questions 4 and 5.

Because of the very low frequencies in some cells, the table was collapsed by combining "often" and "sometimes" together and "rarely" and "never" together, to meet the requirements of a chi square test, giving a value for chi square of 4.8 ($df = 1, p < .05$). This suggests that use of slides is less common than use of tapes, but that teachers who do use slides are likely to use tapes as well.

Table 2 shows the number of responses in each of the collapsed categories.

Slides	Tapes		Totals
	Often/Some	Rare/Never	
Often/Some	21	5	26
Rare/Never	60	49	109
Totals	81	54	135

Table 2 Chi-square table - use of slides and tapes.

- (6) Did you ever have to learn vocabulary? (This was defined for the children as "lists of French words and what they mean in English", and the same four answer choices were given as in the previous two questions).

Often	67%
Sometimes	21%
Rarely	8%
Never	4%

This question was of particular interest, and the result does seem clear, even allowing for bias due to the sample selection or in the responses. Almost 90% of the sample said they had to learn vocabulary at least sometimes, and two thirds reported that it was often set, so that vocabulary learning appears here to be more common than use of slides or tapes. It might also be interesting to see if there was any apparent relationship between this "traditional" task and the use of audio-visual methods. For example, are teachers who set vocabulary learning less likely to use slides and tapes? Tables 3 and 4 show the breakdown of responses to question 6, related to the two previous questions.

Learn Vocabulary (Question 6)	Use of Slides (Question 4)			
	Often	Sometimes	Rarely	Never
	Often	11	7	14
Sometimes	2	4	7	15
Rarely	1	2	2	6
Never	0	0	0	5

Table 3 Number of responses in each category, questions 4 and 6.

Learn Vocabulary (Question 6)	Use of Tapes (Question 5)			
	Often	Sometimes	Rarely	Never
	Often	31	23	23
Sometimes	13	4	6	5
Rarely	5	2	2	2
Never	3	0	0	2

Table 4 Number of responses in each category, questions 5 and 6.

In each case, the four categories were collapsed into two as before, but in neither case was chi-square significant. (For questions 4 and 6, chi-square = .04, d.f. = 1, $p > .05$. For questions 5 and 6, chi-square = .003, d.f. = 1, $p > .05$). These results suggest that, in this sample, there was no apparent relationship between frequency of vocabulary work and use of either slides or tapes, and that the vast majority of the teachers do set vocabulary learning at least sometimes, regardless of what other methods they use.

(7)	If you did have to learn vocabulary, was it usually
Homework	28%
Classwork	27%
Both	45%

This question was asked to see if teachers perhaps regarded vocabulary learning as the sort of task which children could be left to do on their own, at home. In fact, the results are very evenly divided, suggesting that vocabulary work was equally likely to be given in class or as homework.

- (8) How many words, roughly, did you have to learn at a time? (The children were asked to write down a number, though a few gave a range - e.g. "10 - 15". When this happened, E simply entered the mean number. Responses have been grouped for display purposes).

Less than 10	10%
10 - 14	35%
15 - 19	18%
20 - 24	20%
25 - 29	10%
30 or over	7%

The mean was 16.9 with a standard deviation of 9.9, but the modal number was 10, with 26% of the children giving that answer.

(9) Were you usually given a test on the words?

Yes 80%

No 20%

This result suggests strongly that the "traditional" vocabulary test is still in use in many schools, in spite of the fact that it has no place at all in audio-visual courses, where vocabulary is supposed to be built up, and tested indirectly, in context.

(10) Did you find vocabulary learning generally

Very easy 6%

Quite easy 64%

Quite difficult 28%

Very difficult 2%

(The children were asked to "be honest" here and to choose one answer, though a few indicated a response somewhere between two of the possible answers. When this happened, E allocated the responses alternately to the lower and higher category).

As one might expect, the two extremes tended to be avoided, though a majority of children chose the "quite easy" category. However, this was one of the few questions where an opinion or judgement was explicitly requested, and the problems of a non-representative sample might have affected responses here more than on the more "factual" questions.

- (11) Did the teacher ever give you advice about how to learn vocabulary - what to do to help you remember the words? (E provided some examples here, like "Repeat the words over and over again" or "See if the French word reminds you of an English word").

Often	8%
Sometimes	27%
Rarely	34%
Never	30%

Here again the result is quite clear. Although vocabulary learning may be frequently set and tested for most children, there appears to be rather less advice on how the words might be learned. It would be interesting to see if there was any apparent connection between the subjects' reports of how easy they found vocabulary learning, and how much advice they were given. Table 5 presents the two sets of answers for questions 10 and 11. When categories were collapsed as before, a chi-square test gave a non-significant result (chi-square = .03, d.f. = 1, $p > .05$), suggesting that there was no significant relationship between the amount of help given and the difficulty of learning vocabulary. One interpretation could be that this is an easy task in which advice from the teacher would be of little use, but this possibility will be considered in the main experiment in the next chapter.

Ease of learning Vocabulary (Question 10)		Help With Vocabulary (Question 11)			
		Often	Sometimes	Rarely	Never
	Very Easy	1	3	3	1
	Quite Easy	6	23	33	19
	Quite Difficult	3	10	8	14
	Very Difficult	1	0	1	1

Table 5 Number of responses in each category, questions 10 and 11.

The next part of the questionnaire consisted of the ten questions on English lessons and comprehension work, and will be discussed in chapter four. The final section of the questionnaire had a space left for "Any extra comments about your school, French and English lessons". An examination of the responses suggested that most of them could be classified fairly easily into a clear expression of a positive or a negative attitude.

In all, 117 children made some comments on either French or English, or both, and thirty-eight wrote nothing. It had been hoped at first to try to classify their comments using the four categories found by Duckworth and Entwistle (1974) - interest, difficulty, freedom and social benefit, but this proved virtually impossible since about half the comments expressed feelings about the teacher rather than the subject, while many others were simply statements of liking or disliking the subject, with no real reason or details given.

Taking French separately, however, a few general trends can be observed, with examples to support them. In all, there were only fifteen definitely favourable comments, and all but three of these were really expressing a liking for the teacher rather than the subject. Only two children said they found French "interesting", while seventeen described it as "boring". In all, seventy-nine children made some sort of negative comment about French. A reason was occasionally given - "French is boring because we just learn words without games or fun" or "French is not very interesting, all we do is work from the French book", but most confined themselves to the simple statement "French is boring". The other clear complaint was that the subject was difficult, and twenty children made statements like "We get too much at a time to learn" or "The teacher goes on about something no-one understands".

Two particular criticisms were that the teacher went too quickly or expected too much - "In French lessons he says things too fast and he has to ask us it more slowly".

"Confusing. Talks in French and does not explain much. She asks you a question and does not say what she said".

"French is extremely hard because the teacher rushes a lot. He gives help, but sometimes does not fully explain himself. I hate him".

Dislike for the teacher was in fact the most common negative comment of all, and thirty-eight children clearly stated a dislike, as opposed to twelve who said they liked their French teacher. Attitudes to the subject and the teacher were very closely linked, and only very occasionally did a child distinguish between them - "French is boring because she goes on and on. In other words I hate it. But she's a good teacher".

Finally, a few children commented on specific aspects of their lessons - "I think we should not have so much vocabulary" or "I think French would be more interesting if they had slides". On the whole, however, comments tended to be simple statements about liking, or more commonly disliking, French and/or the teacher. When a reason was given, it was almost always that the subject was boring or was difficult, because the teacher went too quickly and did not explain enough.

(d) A Note on Reliability

The children had been asked to indicate if they had the same teacher, in the same class, as any other person present when the questionnaire had been administered, and their papers had been marked accordingly. It was thus possible to make a rough estimate of the reliability of the questionnaire, though it is quite probable that several children copied answers from each other in spite of requests not to do so. Of the 135 who did French, forty-one said they were in the same French class as somebody else present, and ideally, their answers should have agreed on seven of the eight questions on French lessons, the only exception being question ten, which asked how easy they found vocabulary learning. In fact, there were thirty-eight occasions when one person's answer differed from the other's. (When three or more people had the same teacher, the "majority decision" was taken. For example, if two children put that they "often" had slides, and a third put that they "sometimes" had slides, this was counted as one disagreement).

Further examination showed that disagreements on questions offering four choices tended to involve differences of only one category - only eight times did one child put "often" where the other or others had put "rarely" or "never", for example. Finally, disagreements were evenly spread over the seven questions, with no one obviously worse than the rest.

Agreement was thus reached 249 times out of 287, or 87% of the time. Though it must again be emphasised that figure could be an inflated one because of copying, it does at least suggest that the children were able to agree about their ratings most of the time.

(e) Summary and Conclusion

It was stressed from the outset that the questionnaire was of an exploratory nature only. It was not really the intention, for example, to compare different types of school, or possible differences in teaching methods used with children of various ages, and if this had been the intention, a much larger, stratified sample should have been used, and correlations between responses to different questions examined in much more detail.

In spite of its limitations, however, the questionnaire did at least suggest tentative answers to the questions asked at the beginning of this section, and tried to establish a link between classroom practices and the experiment to be described in the next chapter. Even accepting possible bias, it does seem likely that many children are still expected to learn lists of French vocabulary, whether they are taught by audio-visual methods or not, and the children in the sample in fact described vocabulary work as more common than use of either tapes or slides. The number of words learned at a time varied considerably, but the range ten - twenty would cover about three quarters of the sample. The vast majority of them were given tests, though advice on how to learn the words was not very common. The question on how difficult they found vocabulary learning, and their free comments are probably rather more open to sample bias, but French did not emerge as a very popular subject, and many children described it as difficult or boring (c.f. Duckworth and Entwistle, 1974), with a corresponding dislike of their teacher.

The first part of this chapter outlined several controversial issues concerning foreign languages in schools, and described a growing interest in pupils' learning strategies which has been expressed by psychologists and foreign language specialists alike.

Using evidence from this recent research a case can be made for retaining the "traditional" task of vocabulary learning, in spite of the recommendations of the widely used audio-visual courses. The second part of the chapter attempted to support this argument with evidence that, in practice, many teachers appear to set vocabulary learning anyway, though they often give little advice on how it should be carried out. The third part of the chapter describes one area of research on imagery - the study of mnemonics, which offers techniques that have produced convincing results in the laboratory, but have not yet been applied to children's classroom learning.

II 3 MNEMONICS

(a) Introduction, and Three Objections to Mnemonics

The recent research described in the first chapter suggested that mental imagery can be helpful in a variety of tasks, including many which might appear purely verbal. The idea of using images in mnemonic systems can be traced back to the ancient Greeks, whose advice has been modified and passed on over the centuries right up to the present day. Several writers have described this tradition (e.g. Hunter, 1956; Paivio, 1971; Wittrock, 1975), though the most detailed historical account is almost certainly that by Yates (1966) describing the various systems of mnemonics, in most of which the items to be remembered are related, often by an image, to some established pattern, which the learner has already memorised. Different types include the idea of "chaining" items in a list together, "story" mnemonics, various digit-letter systems, first-letter mnemonics, and variations on the method of loci ranging from the simple "one-bun" list to the complex "memory theatres" of the Renaissance. Although some of these may be purely verbal, with images as a useful addition for remembering the more concrete items, most writers agree that imagery forms a vital part of many mnemonic systems (e.g. Underwood, 1969; Paivio, 1969, 1975a; Van Tharp, 1972).

Mnemonics seem to be particularly good at conveying sequential information: a recent study by Morris and Cook (1978), for example, suggests that the first-letter mnemonics, which need involve no imagery at all, facilitate recall of the order of items in a list, rather than the retrieval of individual items. However, many mnemonic systems do use imagery, and it is interesting that the images are being used to help recall sequential information, even though the idea of sequence is conveyed by the previously learned system rather than the individual images it incorporates.

Images may have a key role in the artificial memory systems that have been used for centuries, but objections to mnemonics have an equally long tradition, going back to Quintillian. Three types of opposition can be discerned:

- (i) Mnemonics have always been considered as rather vulgar, not quite "proper", or even downright wicked. Mediaeval scholars held that the imagination was a lower faculty than the memory and that its use to aid recall was a concession to man's weakness. Yates (1966) makes this point and goes on to quote the Puritan preacher William Perkins' claim that:

The animation of the images which is the key of memory is impious: because it calls up absurd thoughts, insolent, prodigious and the like which stimulate and light up depraved carnal affections.....A thing faigned in the mind by imagination is an idol.

The controversy over mnemonics which raged for centuries in fact involved much larger issues like the Renaissance glorification of man's imagination on the one hand, and the Protestant opposition to any form of idolatry on the other, and though objections like those of Perkins may not persist today, there is still the feeling that mnemonics are merely amusing tricks, more appropriate to the music hall than the classroom or the laboratory. (Perensky and Senter, 1970(b); Bower, 1973; Hitch and Baddeley, 1977). Huttenlocher's description (1977) of subjects' apparent embarrassment at using imagery as "an intellectual crutch" in solving logic problems is rather similar, though mnemonics, like the study of imagery generally, have probably become more respectable over the last few years.

- (ii) Part of the explanation for this attitude to mnemonics may lie in a second objection to them - the idea that they are associated with rote learning. A mnemonic can help someone recite information without understanding a word of it, and such "learning" is totally opposed to the current models preferred by educationalists and psychologists alike, with the emphasis on comprehension, hypothesis - testing, discovery, and the importance of relating new information to what the learner already knows.

In a sense, there may be a parallel here between a preference for intrinsic rather than extrinsic structure and a similar preference for intrinsic rather than extrinsic motivation, both reflecting a generally "child-centred" approach. Certainly, educational demands have changed, so that children are not now expected to memorise lists of historical dates, for example, and systems like Brayshaw's "metrical mnemonics" (Hunter, 1957) are no longer necessary or desirable. Technological changes have also been important; without books, the ancient Greeks simply had to commit things to memory. Hitch and Baddeley (1977) conclude that "the long-practised art of memory is now dead, killed by the advent of literacy and cheap and convenient writing materials". Mnemonics can thus be objected to for reasons involving social and cultural changes which have rendered them no longer useful.

- (iii) A final type of objection to mnemonics is simply that they do not work. Common-sense alone would suggest that mnemonics actually increase the amount of information to be remembered (Montague, 1972). Similarly, Yates quotes Perkins' other objection to the method of loci on the grounds that:

It burdens the mind and memory because it imposes a triple task on memory instead of one; first (the remembering of) the places; then of the images; then of the thing to be spoken of.

Moreover, Yates is very careful to point out her own views on mnemonics, upholding:

The common-sense view, to which, personally I heartily subscribe - as explained earlier I am a historian only of the art, not a practitioner of it - that all these places and images would only bury under a heap of rubble whatever little one does remember naturally.

These objections to mnemonics, on the grounds that they are merely amusing and no longer useful tricks which do not work anyway, will be considered during the rest of this chapter. The next section looks at possible explanations for the effectiveness of mnemonics, and suggests that their main value lies in their ability to form mediating links to aid recall. Two current issues from the research are then discussed, the first involving the best source for these mediating links, and the second the best form that they should take. The final section discusses some possible educational applications of mnemonics.

(b) The Effectiveness of Mnemonics

As with many other topics, William James provides a good starting point for the psychological study of mnemonics. In his Psychology (1892), James distinguished between three ways of improving the memory. First were "mechanical methods" involving the "intensification, prolongation, and repetition of the impression to be remembered". Second were the "judicious methods" of remembering things which were "nothing but logical ways of conceiving them and working them into rational systems, classifying them, analysing them into parts etc., etc. All the sciences are such methods". James' definition of the third type, "ingenious methods", is worth quoting at some length:

Of ingenious methods many have been invented, under the name of technical memories. By means of these systems it is possible to retain entirely disconnected facts, lists of names, numbers and so forth, so multitudinous as to be entirely unrememberable in a natural way. The method consists usually of a framework learned mechanically, of which the mind is supposed to remain in secure and permanent possession. Then, whatever is to be remembered is deliberately associated by some fanciful analogy or connection with some part of this framework, and this connection thenceforward helps it recall.

One objection to these "ingenious methods" - the common sense view that they do not work because of the extra burden they impose - can be disposed of immediately. Research using a variety of tasks, conditions, subjects and types of mnemonics has shown that they do improve learning, often to a dramatic extent.

Evidence ranges from the anecdotal (e.g. Miller, Galanter and Pribram, 1960), through small-scale studies (e.g. Ross and Lawrence, 1968) to detailed examination of specific aspects of mnemonics (e.g. Wood, 1967), and several writers have provided reviews of the evidence or descriptions of various types of mnemonic (e.g. Hunter, 1957; Norman, 1969; Bower, 1970(a); Paivio, 1971). Bower (1973) appears particularly enthusiastic in his support for mnemonics: after stating that they really do work, he adds that "to doubt that is to doubt God, country and the value of motherhood"!

Accepting the fact that mnemonic systems do improve recall, various explanations for their facilitation can be offered. One suggestion is that use of a mnemonic increases subjects' motivation (e.g. McNulty, 1966), though most other writers do not favour this as an explanation. Delin (1968), for example considers that "imagery mnemonics have a direct effect on performance; an effect which cannot be accounted for in terms of such factors as motivation". Bower (1972) regards the motivation explanation as "useless, because it explains nothing, or everything", and Paivio has similarly pointed out the lack of real explanatory value and danger of circular arguments on several occasions (e.g. 1969, 1971, 1972).

Finally, it is particularly difficult for a "motivation" explanation to account for mnemonic facilitation in incidental learning, which has been increasingly studied with the growing realisation that student subjects may not follow mnemonic instructions if they expect a recall test later (c.f. Chapter I).

A second possibility is that the images used in many mnemonics facilitate recall by increasing stimulus distinctiveness. Paivio (e.g. 1965) and Bower (1970b) again oppose this explanation and further experimental support is provided by Groninger (1974), Morris and Stevens (1974) and Lowry (1974), who all describe imagery's value as a relational organiser. Lesgold and Goldman (1973) do suggest that differentiation is an important factor, and that the widespread recommendation that mnemonic images should be bizarre could in fact be "an approximation to encoding uniqueness", but most researchers prefer to emphasise a mnemonic's ability to relate items together as its most important contribution to aiding recall.

The items may be related visually or verbally - in a picture presented by the experimenter, a mediating image generated by the subject, or in a sentence produced by either.

One study of the importance of relatedness in verbal strings which is often quoted in the literature is that by Rohwer, Lynch, Levin and Suzuki (1966), who showed that pairs of words linked by a verb were easier to recall than pairs linked by a preposition, which were in turn easier than pairs linked by a conjunction. A less well-known study by Buggie (1974), however, points out that Rohwer et al used only two conjunctions, as opposed to sixteen prepositions and twenty-four verbs, and suggests that the results could simply be due to interference effects. (Buggie used the same verb for all pairs, and found no better results than for unconnected pairs, but did not take subjects' spontaneous mediation into account, which tends to weaken his argument). Rohwer's own explanation (e.g. 1970) is in terms of the images evoked by the various strings of words - verb connectives gave rise to an active mental image, while the conjunctions and prepositions produced more static images.

Certainly, support for the importance of relatedness in mnemonic images is widespread, and several writers have commented on the ability of an image to combine items into a unified whole, or "chunk" (e.g. Bower, 1970(b); Begg, 1973; Begg and Robertson, 1973; Kerst, 1976). It could be misleading, however, to assume that "relatedness" is a simple variable; Robbins, Bray, Irvin and Wise (1974) showed that it interacts with the concreteness of the materials - with abstract words, instructions to form interactive images were no more effective than instructions to form separate images.

Although there is always the problem of checking subjects' compliance with such instructions, this result could be seen as offering indirect support for Rohwer's explanation in terms of images, mentioned above. Bransford and McCarrell (1974), on the other hand, suggest a different emphasis, and are less convinced about the functional significance of imagery in recall (c.f. Chapter I). They suggest that relatedness is the most important factor, and that imagery appears effective simply because it is just one way of relating items together.

Similarly, experiments which have studied actual pictures rather than mental images as memory aids have also tended to support the value of relatedness (e.g. Holyoak, Hogeterp and Yuille, 1972; Odom and Nesbitt, 1974; Arnold and Brooks, 1976). Indeed, Wollen, Weber and Lowry (1972) suggest that interaction is really responsible for the facilitation apparently produced by bizarreness, using pictures as mnemonic aids, and their results have been replicated by Senter and Hoffman (1976).

It would thus appear that a mnemonic system, whether it uses words, images or external pictures, owes much of its success to "mediation" - the relationships which it establishes with the items to be recalled, rather than to any increase in item distinctiveness or subject motivation. If this is the case then the second objection to mnemonics - that they encourage rote learning - can be at least partly answered. Rote learning - James' "mechanical methods" - simply prolongs the stimulus, usually by repetition, without trying to transform it in any way. Research on "levels" of processing, which will be discussed more fully in Chapter Four, would suggest that rote learning thus involves only a shallow level of processing, and will tend to be relatively inefficient (e.g. Anderson, 1972; Swenson and Kulhavy, 1974; Pressley, 1977). Mnemonics, on the other hand, relate new information to something that is already known, though admittedly in an "artificial" way. In fact, when presented with a supposedly rote learning task, many student subjects adopt a whole variety of image and natural language mediators and try to avoid rote learning as far as possible (e.g. Paivio, 1971; Blick and Boltwood, 1972; Ott, Butler, Blake and Ball, 1973).

Seen in these terms, a mnemonic image or sentence allows the subject to form "mediating links" either between the items themselves (e.g. in a "first letter" mnemonic) or between the items and some system that he already knows (e.g. the method of "loci"). Mnemonics thus involve deeper processing in the form of redintegration (Montague, 1972; c.f. Prytulak, 1971), in which the new material is incorporated into larger but more meaningful units. Common sense might indeed suggest that such a process involves an increased burden on the memory, but this is clearly not the case. If the material is already meaningful for the learner it can be understood and assimilated directly, according to all the best principles of a "cognitive" approach to learning. But if the material appears arbitrary or meaningless, rote repetition is not the only possible answer. A mnemonic may well enable the learner to create some sort of artificial "meaning" and to learn the new information in a way that is not only highly efficient, but much closer to James' "judicious" methods than to the shallow, "mechanical" methods of rote learning.

(c) The source of The Mediators: Experimenter-Supplied or Subject Generated

It has been suggested that mnemonics facilitate recall because they supply mediating links, which may be images or in verbal form, between the items to be recalled or between the items and some already familiar system. Two issues in the research will be discussed in turn, the first involving the best source of these mediators, and the second involving the best form they should take. The first issue concerns the question of whether subjects should generate their own mediators when they use a mnemonic, or whether the experimenter should provide them. In the most general terms, both S-R and cognitive theories of learning prefer the idea of an "active" learner (c.f. Hilgard and Bower, 1975) and several researchers studying mediation have also emphasised that subjects should not simply receive information in a passive way (e.g. Anderson, 1970; Taylor, Josberger and Whiteley, 1973; Clarkson, Haggith, Tierney and Kobasigawa, 1973). The idea of learners making up their own images in mnemonics can be traced back to the anonymous author of Ad Herennium (Yates, 1966), who, unlike many other writers, avoided giving lists of "loci", while recent theories involving a "depth" approach to mediation also suggest that the generation of mediators will involve deeper processing, and thus facilitate recall (e.g. Craik and Lockhart, 1972; Griffith, 1976).

Several experimental results support this view, using subjects of various ages and both image and verbal mediators (e.g. Anderson and Hidde, 1970; Palermo, 1970; Bower, 1972; Bull and Wittrock, 1973). Anderson (1970) suggests that subjects will usually make the smallest effort necessary to perform a task, so that getting them to generate their own mediators should be more beneficial than simply asking them to attend to those produced by the experimenter, while Clarkson et al (1973) consider the effort of generation as much more important than its product, regarding images almost as a side-effect (c.f. I5). Montague (1972) and Rohwer (1973) go one step further, and suggest that a mediator provided by the experimenter may actually interfere with one generated spontaneously by student subjects, resulting in poorer rather than better performance. Similarly, Bobrow and Bower (1969) showed that presenting subjects with an anomalous sentence (rather than a meaningful one) to relate paired-associates resulted in poor recall, and they emphasise the importance of comprehension when subjects are allowed to generate their own links.

The subject's ability to generate his own mediators is thus an important factor, and is clearly related to age (c.f. I7).

There is, however, evidence that even quite young children are able to produce their own sentences or images to be used as aids to recall (e.g. Bull and Wittrock, 1973; Kerst and Levin, 1973; Levin, Davidson, Wolff and Citron, 1973) and some writers (e.g. Paivio, 1970, Wolff and Levin, 1972) describe a dramatic development in this ability roughly between the ages of five and eight.

Actual production of mediators, which was originally used as a check on subjects' compliance with instructions (e.g. Paivio and Foth, 1970) has also been shown to facilitate recall. In fact, production of mediators has involved three forms, active manipulation, illustration and verbalization, which appear to be related to the three forms of representation proposed by Bruner (e.g. 1964). Wolff and Levin (1972) suggest that motor activity is necessary for children of kindergarten age to form mediating images, but is no longer needed by eight year olds, and thus tend to support Piaget's view of imagery. (Piaget and Inhelder, 1971). On the other hand, a recent article by Anthony (1977) denies that motor activity is necessary for mediation to occur and claims that learning is just as good if the child simply observes.

As for modern language learning, Asher, Kusudo and de la Torre (1974), propose lessons based on "total physical response" in which adult subjects obey commands like "jump up in the air". The authors claim that such activity improves learning and motivation, though eleven of their twenty-seven students did leave during the first term.....

Similarly, the production of mediating images in the form of pictures has been shown to improve children's recall (Bull and Wittrock, 1973) especially if training was given (Lesgold, McCormick and Golinkoff, 1975). Manipulation of cardboard cut-outs also appears to produce similar improvements. (Lesgold, Levin, Shimron and Guttman, 1975).

Finally, producing a verbal description of a mediator may also improve recall - Taylor, Josberger and Whitely (1973) claim that the ability to verbalize is vital in assuring the success of a mediator, though they do admit that the actual production of sentences could also increase the influence of the experimenter on the subject's performance. Indirect support also comes from a series of experiments on discrimination learning by Levin and his associates (e.g. Wilder and Levin, 1973) which showed improvements related to age, due to verbalization. Similarly, a study by Radford (1966) showed that children could improve their performance on Raven's Progressive Matrices (a "non-verbal" intelligence test) by explaining their answers to the experimenter.

Taken together, the studies outlined above suggest that, when a mnemonic is being used, it should be generally better for the subject to produce his own mediating images or verbal links. Moreover, the research on children would support the idea that learning will be even better if they also overtly "produce" the mediators in the form of actions, pictures or verbal descriptions.

Unfortunately, the picture is not quite so simple, since a number of other studies have shown that mediating images or sentences produced by the experimenter can be just as good, or even better than those generated by the subject himself. For example, the "loci" method still works even if the experimenter describes both the loci themselves and the linking images (e.g. Briggs, Hawkins and Crovitz, 1970; Crovitz, 1971). Similarly, pictures have often been successfully used as mnemonic "images" provided by the experimenter (e.g. Reese, 1970; Davidson and Adams, 1970; Wambold and Hayden, 1975; Pressley, 1976) though there are obvious dangers in equating pictures and mental images.

When subject - generated and experimenter - provided mediators have been compared, the results have sometimes been the same (e.g. Buonassissi, Blick and Kibler, 1972; Kerst and Levin, 1973), while at other times experimenter - provided mediators have actually proved better than the subject's own (e.g. Gran-ger, 1973; Pines and Blick, 1974).

This confusion of results suggests that other factors are involved, and a particularly important one seems to be the length of time between initial learning and the recall test. Griffith (1976), who takes a "depth" approach, points out that a supplied mediator requires only shallow processing, which may well be adequate for an immediate recall test. On the other hand, a subject generating his own linking image or sentence is involved in deeper processing whose effects will persist in a delayed recall test. Unfortunately, studies comparing generated and supplied mediators in immediate and delayed recall have come up with highly contradictory results. Table 6 shows the results of just six such experiments.

	Immediate Recall	Delayed Recall
E - provided mediators better	Ott et al (1973)	Kibler and Blick (1972) Pines and Blick (1974)
S - generated mediators better		Garten and Blick (1974) (two days and seven days)
No significant difference	Buonassissi et al (1972) Kibler and Blick (1972) Garten and Blick (1974) Pines and Blick (1974) Kerst and Levin (1973)	Buonassissi et al (1972) Garten and Blick (1974) (eight weeks) Ott et al (1973) Kerst and Levin (1973)

Table 6 Results of six studies comparing experimenter - provided and subject generated mediators on both immediate and delayed recall.

This area would thus seem to be a clear case for that familiar conclusion "More research is necessary", since several other factors appear to be involved in deciding whether subjects should generate their own mediating links or not. The length of delay does seem to be one factor (Garten and Blick, 1974), but an equally important one could be the interaction between the age or ability of the subjects and the perceived difficulty of the material.

As Wolf and Levin (1972) point out, the ability to generate mediators develops later than the ability to use them, so that the provision of mediators could be especially useful with younger subjects, for whom simple instruction to form linking images or sentences for themselves may not be enough (c.f. Montague, 1972; Rohwer, 1973). Conversely, the attempt to force a mediator on to older subjects who spontaneously generate their own may lead to interference and actually reduce recall. Finally, the distinction between supplied and generated mediators may also interact with the type of mnemonic that is being used; Buonassissi, Blick and Kibler, (1972) found that, on a delayed recall test, supplied mediators were better if a "first letter" mnemonic was being used, but not if a "story" mnemonic was being used.

Given that the comparative value of supplied and generated mediators thus depends on a number of factors, and that no unified picture has yet emerged, it was felt that this was one of the issues that could be examined in the specific conditions of the experiment to be reported in the next chapter. The other research issue, to be discussed next, concerns not the source of the mediators, but the most useful form that they should take.

(d) Bizarreness as a Factor in Mnemonic Images

Most researchers in the area would probably agree that a mnemonic's ability to form mediating links is of central importance, and would accept that images are a traditional and effective way of forming such links. In fact, recent research has tended generally to support the ideas of the traditional mnemonic systems, except on one point. If imagery is to be used, what sort of images make the best aids for recall?

The ancient mnemonists were in no doubt about this, for an important recommendation made by them and passed on over the centuries, was that the best images for aiding recall are vivid, ridiculous and bizarre rather than realistic or common-place, and this advice has survived until the present day (e.g. Hunter, 1956, 1957; Gombrich, 1972; Buzan, 1974; Smith, 1975). Curriculum specialists also sometimes assume that a bizarre idea will help children to learn, and the two topics of foreign language vocabulary learning and reading comprehension are no exception.

Thus Cole (1976) advocates "attaching incongruous colours to objects in order to help retention", or the use of "an exaggerated comic effect to make a vivid impact and aid retention" of foreign words. Similarly, Putnam, (1975) describes how she bases comprehension work on "a humorous paragraph, using one root to the point of absurdity". Yet in spite of all this advice, only two experiments have actually supported the value of bizarreness, while all the others in this area have tended to find bizarreness either of no value in recall, or even a handicap.

Delin (1968) got subjects to write down bizarre images to link three pairs of words, and then had these image descriptions rated for bizarreness by two judges. On a surprise recall test fifteen weeks later, the more bizarre images resulted in significantly better recall than the less bizarre ones. Andreoff and Yarmey (1976) also found that bizarreness facilitated recall, particularly retention after a delay of twenty-four hours, but all the other evidence disagrees with these two studies. Bizarre images appear to be harder to form than common-place ones (Nappe and Wollen, 1973) and have been found no better, or even significantly worse, as aids to recall. (Wood, 1967; Reese, 1970; Wollen, Weber and Lowry, 1972; Collyer, Jonides and Bevan, 1972; Kulhavy and Heinen, 1974; Senter and Hoffman, 1976; Hauck, Walsh and Kroll, 1976).

Writing in 1971, Paivio concluded that the case for bizarreness had not been proved, while by 1975, Postman considered its unimportance as one of the few established conclusions of recent imagery research. On the other hand, at least two writers (Anderson, 1975; Neisser, 1976) suggest that bizarreness should facilitate recall, in spite of the lack of experimental support.

Several points can, in fact, be made before rejecting bizarreness entirely. First of all, some experimenters have used different sorts of instructions to get subjects to form normal or bizarre images, and, as Wollen, Weber and Lowry (1972) point out, this may sometimes have been very difficult for subjects to do. Moreover, other research has tended to show that student subjects often do not follow mediation instructions which differ from their own preferred strategies (e.g. Paivio and Yuille, 1969; Jones, 1972, Wittrock and Goldberg, 1975). All these writers thus suggest that manipulation of materials makes a better independent variable than instructional set in experimental studies of mediation. When experimenters studying bizarreness have used instructional set and have also questioned their subjects, up to a third or even more have admitted that they did not comply with instructions, particularly after the first few trials (e.g. Perensky and Senter, 1970; Collyer, Jonides and Bevan, 1972; Senter and Hoffman, 1976).

This effect is probably heightened if subjects are told, or even just suspect, that there will be a recall test, and there is evidence that they will spontaneously use a whole variety of verbal and non-verbal strategies instead of or as well as those proposed by the experimenter (e.g. Reinert, 1976; Kosslyn and Alper, 1977). This may not even be deliberate - Kosslyn and Alper's subjects "claimed to find themselves compelled to encode material to the best of their ability in the face of foreknowledge of an impending test". Incidental learning might avoid these problems (Nappe and Wollen, 1973; Hauck, Walsh and Kroll, 1976), especially if the use of different sets of instructions could also be avoided.

On the other hand, manipulation of stimulus variables has also caused problems. For a start, "bizarreness" is a poorly defined term (Bower, 1970) and various experimenters and subjects could well disagree on how bizarre an item is. For example, Reese (1970), who found no significant difference between recall of his normal and bizarre sentences, considered that "The cat has the umbrella" and "The chicken has the flag" were commonplace, while "The cat is carrying the umbrella" and "The chicken is carrying the flag" were bizarre. Asking subjects to rate their own images for bizarreness could avoid the problem, as well as providing a suitable incidental learning task.

An important point was made by Wollen, Weber and Lowry (1972), who suggested that some previous experiments could have confused bizarreness with interaction or relatedness. They used sets of pictures to separate these two variables, though there are dangers in equating pictures provided by E with images generated by the subject - a criticism later made by Wollen himself (Nappe and Wollen, 1973). If subjects are asked to generate their own images, sentences can be used instead of pictures, but interaction must be controlled, and the bizarre and ordinary strings should really be made up of the same words throughout.

Another aspect of this idea of relatedness or interaction is that any bizarre picture or image may well imply almost by definition that some form of interaction has taken place in the past, or is actually present. Taking Wollen, Weber and Lowry's example Piano-cigar, the "noninteracting bizarre" condition shows a piano tilted over on one leg and a cigar burning at both ends. Though they are admittedly not interacting with each other, some action has had to be performed on them to make them bizarre in the first place. The piano is bizarre only because of its unusual "interaction" with the floor, and the cigar because of its previous "interaction" with a match.

Perhaps a bizarre image always implies some unlikely relationship between an object and something else which may or may not be present? "Black dog" may conjure up a straightforward image, but "green dog" probably suggests a dog which has somehow been coloured green. This whole problem stems from the fact that "bizarreness" is not a well defined variable (see above), but it could also account for the fact that bizarre images take longer to form than commonplace ones and for some of the apparent confusion between bizarreness and relatedness. It may therefore not even be possible to form a bizarre image without also imagining some form of past or present interaction.

Finally, almost all the studies denying the value of bizarreness have been limited to short-term retention, while the only two positive findings have examined delayed recall. (Delin, 1968; Andreoff and Yarmey, 1976). Paivio (1971) had already suggested that any effects of bizarreness might be confined to delayed recall, and this question needs further examination.

When the above points were taken into account, a study by Merry and Graham (1978)⁽¹⁾ suggested strongly that, contrary to almost all the previous evidence, bizarreness was an important factor in both expected and unexpected immediate free recall, and in delayed free recall both with and without an immediate recall test.

(1) The original version of this study was submitted by the first author as part of an MSc in Education, University of Aston, August 1975.

However, the difference between this study and the rest could be explained by the fact that Merry and Graham's subjects were eleven-year-olds, while all the other experiments had used students. Consequently, a follow-up study was carried out using the same materials and procedure, but with 128 College and University students as the subjects. The results of both studies are summarised in Table 7.

Immediate Recall Test			
	Normal Sentences	Bizarre Sentences	Abstract Sentences
I Children, expecting test	31	45	11
II Children, not expecting test	23	41	10
IV Students, not expecting test	40	63	17
Delayed Recall Test (Unexpected)			
	Normal Sentences	Bizarre Sentences	Abstract Sentences
I Children, re-tested	22	47	8
II Children, re-tested	17	36	4
III Children, not previously tested	4	16	1
IV Students, re-tested	17	43	7

Table 7 Mean percentage recall, by sentence type and recall condition, children and student subjects.

In all cases, sentences rated as producing bizarre images were recalled better than sentences, composed of the same words re-arranged, but rated as producing commonplace images (p. 401). This was true regardless of the age of the subjects, and under a variety of recall conditions. (Expected recall was not tested with the students for the reasons discussed above).

Two possible explanations for the facilitation due to bizarreness have been offered. One rests on the idea that bizarre sentences or images are less likely to have occurred in the subject's past experience - in behaviouristic terms, lack of interference due to low frequency of association could facilitate recall. Lesgold and Goldman (1973) found that image uniqueness was a critical factor in mediation, and suggest that the bizarreness advocated by the mnemonists could have been "an approximation to encoding uniqueness". Similarly, Paivio (e.g. 1971, 1972) suggests that the ancient recommendation of bizarreness is in keeping with the ideas of interference theory.

Motivation or "subject interest" have also been offered as an explanation for the facilitation provided by mnemonics, as discussed earlier, without really explaining why this should be the case. Such explanations are thus only partial (Bower, 1972), but do form a basis for suggesting why bizarre images might be more effective than commonplace ones. The bizarre strings were generated in a way similar to that used by Miller and his associates in work on anomalous sentences (e.g. Miller and Isard, 1963).

Such sentences were formed from several normal ones to make strings like "The sticky young rhythm ate wonders", which violate semantic rules but not syntactic ones. These are generally harder to process because they cannot be related to the subject's experience through the stored meanings of the words. In contrast with the bizarre sentences, however, they contain a whole series of anomalies; the bizarre sentences, each formed from only two normal ones, generally contained only one discrepancy, which the subjects were able to resolve by use of a bizarre image. In other words, the bizarre sentences perhaps involved a certain amount of "cognitive dissonance" by presenting the subject with a "mis-match" to be resolved, but without making the task virtually impossible. In these terms, the sheer amount of effort needed to integrate such disparate elements could give rise to deeper processing and facilitate eventual recall. The construction of a suitable bizarre image could thus be related to concepts like Koestler's "bisociation" (1964), or the idea of "semantic distance", with elements normally well separated in a "network" becoming related to each other (c.f. Bower, 1972; Collins and Quillian, 1972).

Paivio (1974) argues against discussing images in terms of metaphors based on existing technology, and prefers to compare imagery with actual perception rather than wax tablets, tape recordings or computers.

In this case, the link with perception could be the work of Berlyne on "collative" stimulus variables, particularly novelty, incongruity and incongruous juxtaposition (e.g. Berlyne, 1966). The formation of images for bizarre sentences, unlike either normal or anomalous ones, might involve subjects at something close to the "optimum level of arousal", and thus result in better recall. An explanation of the facilitation due to bizarreness in mental images can thus be offered within the terms of a cognitive approach to perception itself; but even if this explanation is unsatisfactory, the results above still strongly suggest that the ancient Greek teachers were intuitively right in advocating the use of bizarre images to help their pupils over 2,000 years ago.

(d) Applications of Mnemonics

It has so far been suggested that the value of mnemonics lies in their ability to relate new information to a previously learned system, and that the most memorable interactions between the two will tend to be bizarre rather than commonplace.

Mnemonics are particularly useful when the material is arbitrary - that is when no relationship already exists between the new information and something that the learner already knows. This may occur for two sorts of reasons, one involving the learner himself, and the other involving the material. On the one hand, the learner may lack either the necessary knowledge or an appropriate strategy for relating the knowledge that he does have to the new information, and on the other hand, the material itself may be inherently arbitrary. Each of these possibilities needs to be briefly discussed.

- (i) If the learner does not have the appropriate knowledge or strategy, there are perhaps three possible solutions. Firstly, we can simply decide not to teach him until he does have them; in their different ways, the controversy over learning "Readiness" and Rohwer's (1971) suggestion that much of education should be postponed till adolescence offer this sort of approach. Second, we can take the rather more positive step of trying to give the necessary knowledge first; psychologists have been demonstrating the importance of preparatory "set" for many years, and the work of Bransford and his associates (e.g. Bransford, Johnson, 1972; Bransford and McCarrell, 1974) represents an interesting recent development of this tradition.

In more directly educational terms, Ausubel's "Advance Organisers" (e.g. Ausubel, 1968) offer the same sort of help. Similarly, if the learner lacks a suitable strategy, we can try to teach him one, and there has been some success here, at least in the short term, as Chapter One suggested. The third possibility is to try to alter the form of the material itself so that it can be related to the learner's existing knowledge. Bruner's (1960) notion of "the spiral curriculum" and his now famous hypothesis that "any subject can be taught effectively in some intellectually honest form to any child at any stage of development" are probably the best known examples of this approach.

- (ii) However, it is possible that the material we wish to present is sometimes inherently arbitrary - that there is no way for even the most skilful and well-informed learner to work it out for himself, no matter how it is presented. Why are there thirty days in September, but thirty-one in October? Why do we have to write "Yours faithfully" and not "Yours sincerely" after a letter beginning "Dear Sir"?

Using the principles outlined above to make material more meaningful may not always be enough. For example, another often - quoted study by Bruner (e.g. 1959) involved children being given a map of the Great Lakes, working out for themselves where the major cities should be, developing ideas about resources, communications, and multiple causes, and involving understanding rather than passive reception. One could hardly argue with this, but why are there five lakes, shaped and arranged as they are, and why is this city called Chicago and that lake called Huron? No understanding of general rules can convey the arbitrary labels of specific proper names, yet we obviously require children to know them if they are to communicate at all. The degree to which teachers still expect such knowledge of arbitrary material is open to discussion, though Bower (1973) again appears in little doubt:

Despite protests to the contrary by teachers, I am impressed that rote learning constitutes a large percentage of what gets taught and learned in schools, particularly after basic reading, writing and math are learned. We teachers are fond of denying that our students must rote memorise huge amounts of material in order to be certified as competent in our subject. Rather we state educational goals in terms of developing the student's greater love for and appreciation of the subject matter, or developing his ability to make insightful criticisms, theoretical comparisons, arguments and the like. By and large, that is a lot of high - sounding bunk - at the very least, these are only extra requirements beyond the rote memorization of basic facts that we demand of our students.

In a rather different way, John Holt (1965) points out the dangers of teachers' presenting arbitrary information to children as though it made sense, not recognising, or being unwilling to admit, that it does not. It is pointless, for example, to pretend that English spelling is entirely rule-governed, and that the accepted spelling of any word can be worked out from basic principles. Holt describes a little girl bursting into tears when told how to spell "once", and goes on:

She cried because the word was so crazy, because it smashed to pieces the understanding that she had been carefully building up in her mind about the way words are spelled. Even then, she could probably have lived with this crazy word if only the teacher had troubled to point out that it was crazy. What really makes school hard for thinkers is not just that teachers say so much that doesn't make sense, but that they say it in exactly the way they say things that are sensible, so that the child comes to feel - as he is intended to - that when he doesn't understand it is his fault.

Even Smith (1975), who consistently emphasizes the centrality of comprehension in learning, has to admit that much of it depends in the beginning on arbitrary information, though he insists that "there are few occasions at school where learning is, or should be, completely rote".

But rote learning is not the only alternative to what Smith calls "meaningful learning", for if we admit that at least some of the things we expect children to learn appear to be highly arbitrary, a mnemonic may still sometimes enable the learner to "manufacture meaningfulness", to enrich, transform and manipulate the new information in an active way by relating it to knowledge that he already possesses.

It could be argued that foreign language vocabulary may often involve exactly this sort of arbitrary information, especially to a beginner, for whom the most relevant existing knowledge is almost certainly his native language. For most eleven-year-olds, there can be little reason why "fenêtre" means "window", yet we do expect them to learn this arbitrary piece of information. Admittedly, it might seem much less arbitrary to a fluent adult who could relate it to a native language word (fenestral), to a word in another language (Fenster) or to its derivation (fenestra).

Similarly, knowledge of a rule about circumflex accents and the missing letter "S" could be very useful, and the teacher might well decide to teach these related pieces of information in the hope of making the French word more meaningful. Unfortunately, each of them, including the rule, is at root just as arbitrary as the fact that "fenêtre" means "window", and would have to be learned by rote (or, of course, by a mnemonic!) Clearly, if the link with existing knowledge is obvious (e.g. French words like "intelligent" or "football"), the new information can be acquired directly, but if such links are not apparent to the learner, a case can be made for creating artificial ones in the form of mnemonics. The comparative success of such methods and the best ways in which they might be carried out must be the subject of empirical investigation, and the next chapter will present an experimental study of various ways of learning vocabulary, using children who had just started to learn French, and words that they would actually be expected to learn as part of their course.

The three parts of this chapter have been concerned with three things:

First, a case was made for retaining the "traditional" task of vocabulary learning as part of a Modern Languages curriculum. Using evidence from recent research into learner's strategies, it was suggested that attempts to avoid native language are futile, if not actually misleading, and that some aspects of the currently popular audio-visual courses can be criticised for their lack of flexibility on this point, especially since successful learners often appear to learn new vocabulary covertly in a "traditional" way, however it is presented.

Second, even if these arguments remain unacceptable, the exploratory questionnaire survey strongly suggested that, in practice, many teachers still do expect their pupils to perform this task anyway, and that the days of the vocabulary test are far from over! On the other hand, teachers appear to give little advice as to how new words might best be learned, and leave this largely up to the children themselves.

Third, research into mnemonics suggests that there could be potentially useful techniques, so far examined only in "pure" research, that could be applied to children's vocabulary learning, perhaps particularly in the early stages, when much of the new material must appear unavoidably arbitrary.

These three points are in no way advocating a return to traditional "grammar - translation" methods, but are an attempt to apply the advantages of "meaningful learning" to an activity which lost favour with many language specialists because the shallow processing of rote learning had appeared to be the only possible method. If anything, this chapter argues for a flexible approach, where the teacher does not adhere strictly to any one system and is also concerned both with what his pupils actually do in order to learn, and with any method that might help them learn better.

CHAPTER III

SUMMARY

The third chapter consists of three parts. The first part describes how mnemonics could be adapted to foreign language vocabulary learning, and discusses the very small number of American studies which have been carried out, all but one using students and all using carefully selected materials. In spite of consistent recommendations made by all the researchers, the method has only once been tried on children and never on "real" school materials.

The second part describes a pilot study carried out simply to see if this mnemonic method is feasible with eleven-year-olds who have little experience of a foreign language. The results suggest that the method is in fact very successful.

The third part consists of a larger experiment, involving two control groups and three experimental groups. The reported spontaneous strategies of children in a control group left to their own devices suggest that some strategies are much more useful than others and that successful and unsuccessful learners approach the task of vocabulary learning very differently. Comparison with the experimental groups indicates that the mnemonic method resulted in significantly better scores on both immediate and delayed tests, especially if an "augmented prompt" was also used.

III 1 THE "KEYWORD" METHOD(a) Introduction

Chapter I discussed research in the general area of imagery and learning strategies, and Chapter II suggested that part of this research, into mnemonics, could be applied to a controversial area of the modern language curriculum - vocabulary learning. The provision of learning strategies in general, and the use of mnemonics in particular, have frequently been shown to improve recall in laboratory tasks like paired-associate learning, though many psychologists have themselves pointed out that such tasks may be very far removed from what goes on in schools (e.g. Palermo, 1970; Kintsch, 1974; Phillips, 1976; Pressley, 1977). The first research attempting to relate such findings to foreign language vocabulary learning was published by Ott, Butler, Blake and Ball in 1973. College students learned twenty-four one-syllable German words, using a variety of methods, including two types of "mnemonic prompts".

Group 1 were given what Rohwer would call an "augmented prompt", (see below), though Ott et al call it a "visual elaborative prompt", in the form of an interactive picture, relating the English meaning to another English word which sounded like the German word. For example, the German word Ei means "egg", and this was depicted by a picture of an eye peering out of an egg, as in figure 3.



Figure 3 An interactive picture presented to Group 1. From Ott, Butler, Blake and Ball (1973).

Group 2 were instructed to generate their own images to link the German and English words. They had to select an English word which sounded like the German one, and then generate an interactive image to link the two English words together.

Group 3 were told to use any method to learn each German word and its meaning.

Group 4 were instructed to repeat each pair of words over and over again to themselves.

All subjects filled in a self-report form describing how they had learned each word, and all were given the same test shortly after learning, and again two weeks later. The results were clear; the two groups given elaborative strategies (groups 1 and 2) recalled about twice as many words as subjects in groups 3 and 4. (Group 1 performed better than group 2, but only on the immediate test). Moreover, when the self-reports were analysed, it was found that over 75% of the words learned by subjects in all groups had involved some form of elaborative strategy, either spontaneous or on obedience to instructions, and if a subject in any group had used some form of elaboration, he was much more likely to remember the word, this effect being particularly significant in the delayed recall test.

Finally Ott et al point out that in the classroom, the strategy used by the learner is usually left up to him, or at best he is encouraged to repeat the words to himself - as in the two least successful conditions. They conclude that their research could have important implications for the "real world of foreign language instruction" which might result in considerable gain in performance for a small amount of instructional time.

This research was taken up by Atkinson and Raugh at Stanford University, and three studies were published in 1975. (Atkinson; Atkinson and Raugh; Raugh and Atkinson). The basic method developed by Ott et al remained the same, but it was tested under a wider range of conditions, and its component parts analysed more closely. The three reports are very similar in many ways, and their main conclusions are presented below, along with some further questions they raise.

(b) How Widely Can the Method be Applied?

- (1) The Stanford experiments showed clearly that the "key-word" method, as Atkinson calls it, could be applied to vocabulary learning of other languages apart from German. As one might expect, it was successful with Spanish vocabulary (Raugh and Atkinson, 1975), where the experimental group scored an average of 88% correct on the tests, as compared with a control group score of 28%.

A more stringent test of the method was carried out using Russian words, which have many phonemes not present in English, thus making the selection of suitable "keywords" more difficult. Even here, the method gave highly significant results, with the experimental group scoring 72% on average over the various tests, and the control group 46%.

One would thus expect that the method could apply to French, by far the most common second language learned in this country. However, a much more serious problem involving the application of the keyword method lies in the fact that all the research so far has obviously used very carefully selected words, chosen because they fit the method well. What is not known is how many of the words actually expected to be learned in a second language course could be successfully learned in this way. If the proportion is small, the method may be psychologically interesting, but educationally useless, in spite of the implications claimed by its supporters. This issue is one that it was hoped to examine in the main experiment to be described later.

- (ii) Subjects A second consideration involves the application of the keyword method using other subjects, particularly children.

In common with the majority of published research in Psychology, Ott et al and the Stanford studies all used college students as their subjects. These subjects have been consistently reported elsewhere as using various elaborative strategies spontaneously (I 7), and similar findings have been made in this area of research, both with control group subjects left to their own devices (Ott et al, 1973), and even with subjects explicitly asked not to use any form of elaboration (Raugh and Atkinson, 1975). Consequently, students may well have found it simple to use the strategy provided by the experimenter, perhaps regarding it as a more sophisticated version of what they would do anyway. In contrast, there is the possibility, mentioned in the previous chapter, that provided mediators might actually interfere with ones spontaneously generated by the subject. It has also been suggested that the likelihood of such spontaneous generation increases with age (e.g. Hagen, 1972; Rohwer, 1973; Reese and Porges, 1976; Pressley, 1977), but that even quite young children can be successfully taught to use elaborative strategies (I 2).

**PAGE
MISSING
IN
ORIGINAL**

We now expect almost all children to start learning a second language at the age of eleven, or even sooner. Can the keyword method be used by such children, to learn words that they would really have to learn in school?

In fact, one experiment, published only after the present experiments had been completed, has begun the study of the keyword mnemonic with children. Pressley (1977) reported that children given external pictures and keyword instructions recalled Spanish vocabulary better than children in control groups. However, several points remain to be examined. In common with the other studies, Pressley used specially selected words - all nouns of less than three syllables, chosen from the whole language for their expected suitability. Only an immediate test was used, and no attempt was made to see if the keyword method facilitated the translation of the English word into the foreign word. Though two age groups were studied, the children's spontaneous strategies were not examined, and no measure of ability was involved. Finally, only one keyword condition was used, with pictures provided, and Pressley himself says that the ability of children to generate their own images is still in question.

Finally, apart from consideration of age, it is also possible that the student subjects were already well-motivated and successful learners of another language, though not the one tested in the experiments. Ott et al, for example, used students who were already attending "an intensive foreign language training school for missionaries", though they did exclude anybody who had studied any foreign language in the previous two years. At any rate, Ott and his fellow researchers have since agreed that the subjects in their original study were probably already skilled at memorizing, if not at language learning, and may not have been typical "beginners". (Ott, Blake and Butler, 1976).

In conclusion, the applicability of the keyword method remains to be examined using "real" vocabulary rather than specially selected words, and developed with children rather than highly-motivated college students who use various elaborative strategies spontaneously anyway. This is especially important in view of the claims consistently made by all the researchers that the method could well have important implications for the classroom, particularly for those just starting a second language and for those who find the acquisition of new vocabulary difficult.

(c) Analysis of The Method

Two distinct stages are involved in the strategy - an "acoustic link" in which the subject must relate the foreign word or part of it to a similar - sounding English word, and an "image link" in which he must relate this keyword to the original English word by means of an interacting visual image. The experimenter may provide the acoustic link, or both links, or may leave the subject to provide them for himself; for the sake of clarity, each link will be discussed separately.

(i) The Acoustic Link

The original study by Ott et al (1973) uses only single-syllable German words, each of which sounded like an English word, and there was thus no problem in the choice of the acoustic link. The first experiments at Stanford tried to extend this idea to words with more than one syllable, but with the keyword still linked to the whole of the foreign word. Some examples from Atkinson (1975) will make this clear:-

Atkinson himself seems generally in favour of teachers making most of the decisions in a learning task, rather than leaving them to their students (e.g. Atkinson, 1972), and suggests that it is usually better for the experimenter to provide the keyword, unless the subject definitely prefers to think up his own. However, as subjects get more experienced in the method, it becomes possible to use the "free choice procedure" in which a keyword is made available only if the subject requests it. (Atkinson, 1975). Given the choice, subjects asked for a keyword 89% of the time anyway, and their results were almost identical to those of subjects automatically provided with a keyword by the experimenter, on trials over a three day period.

(ii) The Image Link

In contrast, Atkinson suggests that it is better to let the subject generate his own image, rather than to have a mediator presented for him in the form of a picture or verbal link. It has been emphasised that the age of the subjects and the specific conditions of the study are vital in this respect, and there has certainly been no attempt to compare subject - generated and experimenter - provided mediators in children's use of the keyword method.

A useful theoretical approach to the problem is provided by Rohwer's "elaboration hypothesis" (Rohwer, 1973), which postulates a hierarchy of prompts and relates the facilitation which they provide to the spontaneous strategies used by subjects. Rohwer suggests five levels of prompt:-

Antagonistic prompts (e.g. instruct subjects to attend to irrelevant attributes, as in an incidental learning study).

Minimally explicit prompts (a simple instruction like "learn this").

Explicit prompts (a strategy is offered - e.g. S is instructed to form an image to relate items).

Augmented prompts (the experimenter actually presents the prompt, in the form of a picture for example, as in the original study by Ott et al).

Maximally explicit prompts (the subject actually re-enacts the prompt). This level appears to be similar to the method of getting subjects to produce a drawing or verbal description of the image - a device used to check that subjects really are following mediation instructions (e.g. Paivio and Foth, 1970; Yuille, 1973; Rasco, Tennyson and Boutwell, 1975), but which also appears to facilitate recall (II 3, C).

In general, the hypothesis predicts that the older the subject, the less explicit is the prompt needed to produce optimum learning; indeed, with college age subjects, an explicit or augmented prompt may even interfere with spontaneous mediation and cause a reduction in learning. Rohwer commented that, at the time of writing (1973), "there seem to be no experiments

that include a comparison of pictorially augmented prompts with simple explicit prompts". The study by Ott et al (1973), described above, did in fact look at this issue, but no experiments have yet been carried out on the relative value of explicit and augmented prompts in children's use of the keyword mnemonic, and this was a question to be considered in the main experiment.

However, the first task was to see if the keyword method was even feasible for use with eleven-year-olds with little experience of a second language, and it was felt that a pilot study would be useful. If the method did prove feasible, a larger study could examine some of the issues involved.

III 2 EXPERIMENT I - PILOT STUDY⁽¹⁾

(a) Summary

The pilot study involved the comparison of two similar groups of children on four vocabulary tests, two immediate and two delayed, one group having been taught twelve French words

⁽¹⁾ The data collection for the pilot study was carried out in collaboration with Miss J. French while she was a student at St. Peter's College, Saltley, Birmingham.

by the keyword method, and the other group literally "left to their own devices" to learn the same words over the same period of time. The null hypothesis, that there would be no significant differences between the two groups' scores, was rejected on all four tests ($p < .05$). Some obvious limitations of the experiment are discussed.

(b) Subjects

The subjects were forty-two eleven-year-olds, divided into two groups (their normal mixed ability classes for French). They were in their first year at Secondary School, and had been learning French for about six months. There were twenty-two boys and twenty girls, and both classes were normally taught French by the same teacher, who regarded them as similar in ability.

(c) Materials

A list of twelve French words with suitable keywords and English equivalents was used, and a check was made with the teacher to ensure that the children had not already learned any of the words chosen. The list is shown in Table 8, in order of presentation.

French Word	English Equivalent	Keyword
lunettes	spectacles	net
tapisserie	tapestry	tap
imprévu	unexpected	pray
matelot	sailor	mat
loupe	magnifying-glass	loop
canapé	settee	can
lapin	rabbit	lap
côteau	hillside	cot
baguette	rod	bag
batelier	ferryman	bat
pample mousse	grapefruit	moose
pataud	clumsy person	pat

Table 8 Pilot study vocabulary list.

Following the advice of Raugh and Atkinson (1975), the keywords were all monosyllables, and were similar to the first syllable of the French word wherever possible.

(d) Method

One class, chosen at random, served as the experimental group (n = 20), and the other class as the control group (n = 22). Both groups were told that they would be taking part in

an experiment to see how well they could learn some new French words, and that they would therefore be tested at the end of the experiment. Both groups were given the same three practice items first, and were allowed the same length of time to learn each word.

Control group Each French word, and its English equivalent, appeared together on a large "flash card" (a technique already frequently employed in the school with pictures). E first exposed the French word, pronouncing it twice, and then its English equivalent, which was also pronounced twice. Subjects were then instructed to learn the pair in any way they wished. At the end of ten seconds, E again pronounced both words before moving on to the next item.

Experimental group received exactly the same treatment, except that after exposing and pronouncing each pair, E went on to expose and pronounce its keyword, and to describe an interactive image to link the two together. Ten seconds were allowed for each description.

e.g. "lunettes.....spectacles.....net"

"Imagine a football net with a pair of spectacles caught up in it, left there by the goal keeper".

Whenever possible, the connecting images were made bizarre rather than ordinary (II 3, d), and a full list is given in Appendix B.

Finally, E again pronounced each French word and English equivalent, as with the control group. After all twelve words had been given, answer sheets were distributed, and the same two tests were then given to all subjects.

Test I French into English E exposed the twelve French words, previously written in random order, on the blackboard, and pronounced each word in turn. Subjects were asked to write down what they thought the word meant in English, and were reassured that spelling would not matter. There was no time limit.

Test II English into French E then read out the twelve English equivalents, in a different random order. Subjects were instructed to copy from the board the French word which they thought meant the same. It should be noted that this test was therefore one of recognition rather than recall, and that results cannot strictly be compared with test one. The second test was carried out because Atkinson and his associates have published only studies testing recall in the manner used in test one, and appear less certain about the value of keywords in helping subjects learn translations from English into the foreign language. It was felt, however, that a recall test, in which subjects would have to write out the French words after a single exposure, would be expecting too much of eleven-year-olds, and so a recognition test was used. All subjects were

unexpectedly given exactly the same two tests a week later. The null hypothesis was that there would be no significant differences between the scores of the keyword and control groups, on any of the four tests. Finally, after all the tests were over, E asked control group subjects informally about what strategies they had used. All the suggestions came from the class originally, and the children were encouraged to write down any comments they had about how they had learned the words. Analysis of this data was not to be attempted, but it was hoped that it might suggest some spontaneous strategies which could be examined in the main experiment.

(e) Results

Table 9 shows the mean score out of twelve for each group on each test, and Table 10 the number of "perfect" scores in each condition.

	Immediate English - French	Recall French - English	Delayed English - French	Recall French - English
Control group	7.1	7.0	6.0	6.3
Keyword group	8.7	9.3	9.1	8.3

Table 9 mean scores by group condition, pilot study.

	Immediate English - French	Recall French - English	Delayed English - French	Recall French - English
Control group	3	6	3	3
Keyword group	8	9	7	5

Table 10 Number of subjects with maximum scores, control and keyword groups, pilot study.

Table 10 suggests that there was a marked ceiling effect, so a non-parametric test of significance was used. Mann Whitney U tests showed that subjects in the keyword group had recalled significantly more than subjects in the control group on all four tests ($P < .05$ in all cases), and the null hypothesis was therefore rejected.

Moreover, the ceiling effect did not appear to be equally strong in both groups. Because of the small numbers in some cells, categories were collapsed to meet the requirements of a Chi squared test (Siegel, 1956), which showed that there were significantly more maximum scores overall in the keyword group than in the control group ($\chi^2 = 7.0$, $p .01$).

(f) Discussion

The results of the pilot study thus suggest that the keyword method can be successfully used, even without provided pictures, and not by college students, but also by eleven-year-olds just beginning French, if both the keyword and a description of the linking image are provided. The superiority of the experimental group was not so great as that sometimes reported in the Stanford experiments on students, but this could be at least partly due to the marked ceiling effect - overall, there were almost twice as many maximum scores in the experimental group as in the control group (33% and 17% respectively). The pilot study also suggests that the keyword method could be especially useful in delayed recall - obviously an important point in classroom learning. (Overall, the experimental group showed a superiority of 28% over the control group in the immediate tests, and 42% in the delayed tests).

The results suggest that the keyword method can be used to improve scores on a test where subjects have to recognise the foreign language equivalent of an English word, and not just remember the meaning of the foreign language word, as in all the previously published studies. However, such a recognition test cannot be directly compared with a recall test, and it is also possible that the recall test itself had a powerful rehearsed effect. The importance of covert rehearsal has been stressed by several writers (e.g. Bugelski, 1962; Montague, 1972), and Yuille, (1973) suggests that an immediate recall test could serve as valuable overt rehearsal which would have a pronounced effect on later tests.

The pilot study was, however, obviously very limited, and raises several issues that could be examined further:-

- (i) The vocabulary used in the pilot study was specially selected to fit the keyword method. How far could the method be used profitably with "ordinary" vocabulary that children starting French would really have to learn?
- (ii) Results might be improved if the ceiling effect, especially marked in the experimental condition, could be avoided.

- (iii) The pilot study did not establish a base-line; it was possible that subjects already knew, or could have guessed, the meaning of some words.
- (iv) The pilot study did not give any detailed information about the strategies children spontaneously use in such tasks, when left to their devices, though this was discussed informally with the control group.
- (v) It is possible that any form of instruction would improve learning beyond that of the control group. (c.f. the notorious "Hawthorne effect").
- (vi) Conversely, it is possible that even richer mediation prompts would facilitate recall even more than the use of images alone. (c.f. Rohwer, 1973).
- (vii) With a larger group, it would be possible to look for possible interactions between individual differences and the effectiveness of various strategies.
- (viii) Subjects might possibly prefer some methods of learning to others, and this factor should also be taken into account, though motivation alone is not a satisfactory explanation for the general superiority of groups given imagery instructions. (II 3, b).

III 3 EXPERIMENT II - MAIN STUDY

(a) Summary

The main experiment was carried out using the whole first year group (n = 195) of a local mixed Comprehensive School. Scores were obtained on the Schonell "B" Reading Test, and subjects were randomly assigned to five treatment groups. A control group learned seventeen French words, drawn from their actual course material, in any way they wished, and the learning strategies they reported using were compared with their success on four tests similar to those used in the pilot study. Three experimental groups were instructed to learn the words either by rote repetition, the keyword method, or the keyword method plus an "augmented prompt" in the form of a line drawing supplied by E. A final group was given the tests only, to establish a base line. Analysis of the control group's results suggested distinct differences between the reported strategies of successful and unsuccessful learners left to their own devices. Comparison with the experimental groups indicated that

use of the keyword method had resulted in significantly better performance on all four tests, with even better scores produced in the "augmented prompt" condition.

(b) Introduction

The main experiment was an attempt to examine some of the issues raised by the pilot study. Based on the points mentioned above, the pilot study was modified as follows:-

- (i) The vocabulary in the pilot study like that used in the American research, was specially chosen for its suitability for the keyword method. However, the method would not appear to be suitable for learning all new words, and it would be interesting to find out roughly what proportion of new words actually expected to be learned by children starting French could be learned in this way. In other words does the reported success of the keyword method depend on specially chosen vocabulary, or can it be used in a more "real" situation, with some of the vocabulary which is actually part of the first year French curriculum? In an attempt to answer this, all the words used in the main study were taken from the vocabulary which the children would have to learn anyway in the next few weeks. (See "Materials" below).

- (ii) Because of the ceiling effect in the pilot study, the number of words was increased from twelve to seventeen. Similarly, the delayed recall period was extended from one week to two.
- (iii) It was possible that subjects might have been able to guess, or already knew, some of the words used in the pilot study. Consequently, a small group of subjects were given both tests without any sort of previous learning, to serve as a base-line against which to judge the results of the experimental groups.
- (iv) Apart from informal discussion, little was known about what the control group in the pilot study had done to learn the words, and it was therefore decided to ask the control group in the main study what strategies they had used. It might also then be possible to compare the reported strategies of successful and unsuccessful learners, and to see if any of the subjects spontaneously used visual or verbal strategies without being instructed. There is a large amount of research which indicates that student subjects spontaneously use a variety of strategies, even if asked not to do so (see point 5 below), and Ott et al (1973)

found that subjects left to their own devices when asked to learn vocabulary frequently used some form of elaboration. The previous chapters have suggested that spontaneous use of strategies appears to develop with age (e.g. Rohwer, 1973; Kulhavy and Swenson, 1975; Paris and Lindauer, 1976; Paris and Upton, 1976). It would thus be interesting to know what spontaneous strategies are reported by children in the first year at Secondary School, on a "real" school task, and how their reported strategies relate to their success.

- (v) Rote learning instructions have been used in several mediation strategy studies, and rote learning generally appears inferior to most other strategies, at least when experimenters have ensured that instructions were followed (e.g. King and Russell, 1966; Paivio, 1971; Rohwer, 1973; Levin, Davidson, Wolff and Citron, 1973). However, student subjects often appear not to follow mediation instructions which seem unsuccessful or which differ from their own spontaneous strategies (e.g. Paivio and Yuille, 1969; Jones, 1972; Wittrock and Goldberg, 1975). If children generally tend to lack such appropriate spontaneous strategies, they might be expected to benefit from any mediation instructions, even rote repetition,

and be more likely than older subjects to follow such instructions (c.f. Kulhavy and Swenson, 1975). Consequently, it was decided to run a second experimental group given standard rote learning instructions. An additional reason for running a rote learning group was that the informal discussion with the control group in the pilot study had suggested that some form of rote repetition was a popular spontaneous strategy. It would thus be interesting to see if actual instructions to use rote repetition would have any significant effect, resulting in scores superior to the control group left to their own devices

- (vi) Chapter two suggested that pictures could be used to form mediating links in a mnemonic, and pictures have sometimes been taken to represent "images" supplied by the experimenter, as opposed to mental images generated by the subject (II 3, b and c). Although there are obvious dangers in equating the two (e.g. Nappe and Wollen, 1973), the dual coding theory would also suggest that mental images should be generated relatively easily from picture stimuli. (e.g. Paivio, 1975a). In Rohwer's terms (1973) a

presented picture could constitute an "augmented prompt" and should thus usually improve children's learning more than a single "explicit prompt".

There is some evidence that pictures as aids lead to better performance in a variety of tasks than do words (e.g. Rowe and Paivio, 1971; Levin and Kaplan, 1972; Rowe, 1972; McClure, 1974; Rohwer and Matz, 1975), but several other factors seem to be involved, just as they were in the more general comparison of experimenter - provided and subject - generated mediators. The facilitation provided by pictures also appears to depend on factors like the age of the subjects (e.g. Pressley, 1977), and the ability to use a presented picture may actually increase with age, so that young children sometimes perform better with verbal mediators than with pictures (e.g. Milgram, 1967; Davidson and Adams, 1970). Other factors include the type of test given (e.g. Rohwer, 1973; Paivio, 1975a) and the rate of presentation, where pictures are better only at a slow rate (e.g. Paivio and Csapo, 1973; Nelson, Reed and Walling, 1976). The nature of the picture is also understandably important. For best results, the items must be inte-

grated rather than separated in the picture (e.g. Simon, 1972; Arnold and Brooks, 1976; and c.f. Bransford and Johnson, 1972), and a series of very similar, schematic pictures may be no better, or even worse than their verbal equivalents. (Nelson and Walling, 1976). A related finding by Murraray and House (1976) was that a simple line drawing or a verbal link produced roughly the same facilitation, while a colour slide gave significantly better results, though this appears to be contrary to the finding of Dwyer (1967), where a photograph was worse than a line drawing. This area appears to be another complicated one! The best course is therefore probably to present both a picture and a verbal description at the same time which should also allow for individual differences in strategy preferences (c.f. Reinert, 1976), and this policy has been supported in several studies (e.g. Jones, 1974; McMahon, 1975; Rohwer and Harris, 1975). Apart from the keyword group, it was therefore decided to run another experimental group who would not only hear a description of the image they were to form, but would also see a line drawing depicting the appropriate interaction, at the same time.

- (vii) It was hoped that with the larger number of subjects involved in the main experiment it would be possible to look more closely at individual differences. In particular, would there be any sex differences in the usefulness of the various strategies, and might there be some interaction between ability and strategy? Some studies (e.g. Lebrato and Ellis, 1974) suggest that pictures might be of particular benefit to slow learners, for example, though this might appear to conflict with findings mentioned earlier that the ability to use pictures as mediators develops later than the ability to use words (e.g. Rohwer, 1970). Schonell "B" reading age scores were available for all the children, and it was hoped to include these in the analysis of the data. As for sex differences, an established body of research suggests that, very generally, females are superior to males on verbal tasks, but inferior on visual-spatial tasks (e.g. Coltheart, Hull and Slater, 1975). Blick and Boltwood (1972) found no sex differences in reported strategy preferences, though other researchers from Galton onwards have suggested that female subjects have more vivid imagery (e.g. White, Ashton and Brown, 1977) and

are more likely to use it (e.g. Richardson, 1977). Some experimenters have used only female subjects, as Richardson advises, (e.g. Murray and Roberts, 1968; Anderson, 1974) in an attempt to control for sex differences, while others have tried to explain them. Krashen (1975), for example, suggests that visual-spatial skills are more clearly lateralised in the right hemisphere for men, than women.

- (viii) Finally, it might be useful to know which, if any, of the various methods the children enjoyed most. At the end of the experiment, all subjects were therefore asked to say whether or not they liked that method of learning, indicating their response on a simple five point scale, although it was not considered that a motivation explanation alone could account for any differences between groups. (c.f. discussion on image bizarreness in Chapter Two).

(c) Subjects

The subjects were the whole first year group of a local mixed Comprehensive School. The school is a denominational one, and draws its pupils from a very large catchment area, so that there is a wide spread of both socio-cultural background and ability. It was thus felt that the sample, which consisted of 100 boys and 95 girls, was reasonably representative of first

year pupils in Secondary Schools. The children are taught French in mixed-ability classes, and had been at the school for about eight weeks when the experiment was carried out. At the time, their mean age was eleven years eight months.

All the children had been given the Schonell "B" reading test the previous May, and their scores are shown in Figure 4. (Since both tasks to be studied obviously involved language, and the second in particular involved a good deal of reading, it was felt that a measure of each child's reading ability might prove interesting). Their mean reading age was 11.0, with a standard deviation of 2.1 decimal years, at a time when their mean chronological age was 11.1 decimal years, suggesting again that the sample was reasonably representative of children in this age group.

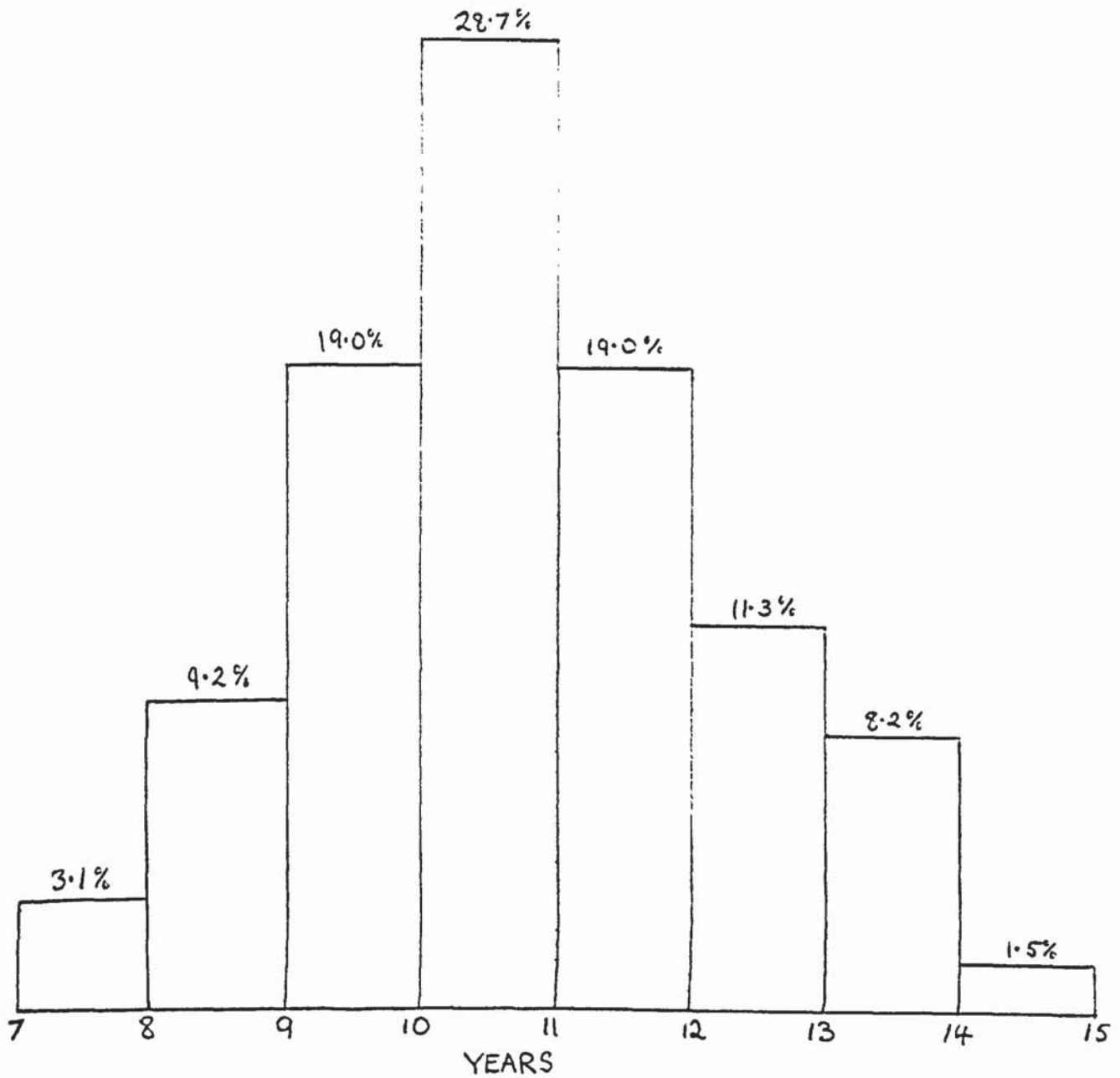


Figure 4 Distribution of the sample's reading ages
(Schonell "B")

(d) Materials

In accordance with the conclusions reached in the previous discussion, it was felt important that the materials should be as "real" as possible, rather than a list of words specially selected entirely to suit the keyword method. All first year classes follow the Longman's audio-visual course (Moore and Antrobus, 1973) very closely, and all take the same examination at the end of the year. The Longman's course is easily the most popular one, being used in about 70% of Secondary Schools.

Having found a point in the course which none of the classes had yet reached, E went on to analyse the text to find the next fifty new words that would be introduced. From this pool of fifty, eighteen words were selected which seemed appropriate to the keyword method. It is important to point out that the method has never been considered suitable for all words, the only suitable ones being those which have at least a partial similarity to some English word. With some new words the method is not applicable, while with others it is not even necessary

(See discussion below). Table 11 presents the fifty words and their suggested classifications.

ami (K)	entrer (E)	paquet (E)
après-midi	épicerie (K)	parler
arriver (E)	être	porter (K)
attraper (K)	fatigué (K)	poser
ballon (E?)	fleur	préparer (E)
beaucoup (K)	football (E)	quitter (E?)
café (E)	l'heure	regarder (E)
chasser	intelligent (E)	sage (K)
chercher	jouer	soir
classe (E)	là	sonnerie (K)
commencer (E?)	lancer (K)	tard (K)
content (E)	manger (K)	tasse (K)
croissant	matin (K)	tomber (K)
de	midi	travailler (K)
déjà	minuit	trop
demie	moins	trouver (K)
	nuît	vilain (K)

Table 11 The pool of fifty vocabulary items.

In the table, (K) indicates that the word was used in the main keyword experiment, and (E) that the word appeared similar to an English word with the same meaning, which would be known to the children. These two conditions are both important. "Travailler", for example, is obviously similar to the English word "travel", but is in fact a "false friend" because it does not mean the same as the English word it suggests. As has already been discussed, this French word might not appear arbitrary to a fluent adult who knew the English word "travail" and who would therefore not need a mnemonic aid. Similar points could be made about several other words (e.g. "sage". "matin"), but it would be wrong to assume too much about the existing vocabulary of eleven-year-olds and their ability to relate it to new words. For example, "fatigué" might seem obviously related to the English word, yet when a group of nineteen children, who had not been taught any of the words, were presented with a list of seventeen words in French, and asked to guess which of them meant "tired", only two guessed correctly, and one of them admitted that she already knew the word anyway! This was in fact part of the main experiment, and the point will be taken up again in later discussion.

The next task was to think up suitable connecting images for the selected vocabulary items, to relate the keyword (e.g. "mat" in "matin" or "tar" in "tard") to the meaning of the word in English. However, as Atkinson (1975) has pointed out, the

selection of keywords and linking images involves a good deal of subjective judgement, and there is yet no firm theory to predict the success of either. Atkinson therefore suggests that

a committee of individuals familiar with the language should select the keywords rather than having one person make the decisions. Experience indicates that individual experimenters can come up with some pretty bizarre keywords that work for them but for no one else.

In fact, all four English teachers and a local authority Foreign Languages adviser proved very willing to examine the proposed list, and made a number of suggestions, resulting in several changes. One of the two proposed practice items was dropped completely, and the other was altered, though they proved rather more benign with the main list: one word was dropped completely, reducing the total to seventeen, and three or four others were improved by the suggestion of a different keyword or linking image.

The final list appears in Table 12 in order of presentation to all groups.

French word	Translation	keyword
tard	late	tar
attraper	to catch	a trap
fatigué	tired	fat
lancer	to throw	lance
sage	good	sergeant
manger	to eat	manager
beaucoup	many	bow
porter	to carry	porter
tasse	cup	ass
épicerie	grocer's	a piece
tomber	to fall	Tom
travailler	to work	travel
trouver	to find	trove
sonnerie	bell	sun
ami	friend	me
matin	morning	mat
vilain	naughty	villain

Table 12 Main study vocabulary list.

On the basis of an admittedly very small sample of fifty words, it was thus possible to think up acceptable keywords and images for about a third. However, the fifty words also included those like "football" and "intelligent" and several others which are the same or very similar in both languages. About a dozen words seemed to fit this category (see previous table). If this pool of words is a reasonable sample, the keyword method could therefore probably be tried with quite a large proportion - at least a third - of the words which children would actually have to learn. A recent report (Wilby, 1978) claims that children are expected to master a vocabulary of about 1,5000 words by the end of their fifth year. The keyword method might thus reasonably be used with several hundred words altogether, although its usefulness may decline as pupils increase their knowledge of the language, so that some new words appear less arbitrary, and "artificial" links less necessary. There is certainly no claim being made that the method is valuable with all words or for all subjects. Atkinson and Raugh (1975) conclude that the keyword method

may prove useful only in the early stages of learning a language and more so for some classes of words than others. The method may not be appropriate for all learners, but there is the possibility that some, especially those who have difficulty with foreign languages, will receive particular benefits.

A full list of the linking images, and copies of the pictures used, appear in Appendix C.

(e) Procedure

All the children were tested in their normal French classes, randomly assigned to the various experimental treatments. All classes were told that they were going to take part in an experiment in which they were going to try to learn seventeen new French words. It would be useful to learn the words because they would meet them all in the next few weeks anyway. It was explained that different first year classes were going to learn the words in different ways, so that if anybody felt they hadn't done very well, it could just be that that way of learning wasn't very good - it didn't mean that they were stupid. They were warned that they would be tested on the words afterwards (though no warning was given of the delayed recall test to be given two weeks later). Interestingly, when asked if they had any questions, somebody in each of the seven classes wanted to know if spelling would matter, and all appeared highly relieved when told that it would not.

After this general introduction, the various treatment groups were given instructions as follows:-

Control Group I - Test Only (n = 19) It was explained to this group that E had to find out if people might know some of the words already, or be able to guess what they meant. These subjects were given both tests without any form of teaching at all, to establish a base-line against which the results of the other groups could be judged.

Control Group II - No Instructions (Two French classes, n = 51)

This treatment was very similar to the control group treatment used in the pilot study. However, so that the visual part of the presentation would be identical for all subjects, an overhead projector was used. Subjects were told that E would show them a new French word and its meaning, and would pronounce both words twice, pointing to them on the screen. They would then have to learn the words in any way they wished, and would have fifteen seconds to do this. Finally, E would pronounce and point to each pair of words again, before going on to the next item. A total of thirty seconds was allowed for each pair, and subjects were given two practice items - both words which they already knew. All words were hand printed in block capitals, giving a screen height of about 6cms. The two practice items were:-

armoire.....cupboard

cartable.....satchel

Rote Group (n = 24) After the general introduction, E discussed rote learning - "repeating something over and over again" - very briefly, then went on to the practice items. Here, the group received identical treatment to that of Control Group II except that they were instructed to repeat each pair of words quietly

to themselves during the fifteen second interval. It could be argued that one advantage for the experimental groups described below was that they actually heard E pronounce each pair of words at least once more often than did the control group, while listening to E's description of the image to be formed. Consequently, the rote group, after hearing each pair pronounced twice, heard an extra two repetitions of the phrase that they themselves were to repeat. The final repetition at the end of the fifteen seconds thus meant that the rote group had actually heard E repeat each pair five times.

Image Group (Two classes, $n = 51$) After the general introduction, E discussed images - "pictures in your mind" - very briefly, then went on to the practice items. After pronouncing and pointing to each word twice, as with all the other groups, E pointed to the keyword, printed in small green block capitals in brackets beneath the French word. (Smaller size, brackets and the different colour were used to distinguish the keyword from the English word). On the practice items, E stressed that the keyword and French word did not mean the same, but that one should simply remind you of the other. E then went on to describe an image to link the keyword and the English word. A final repetition of each pair followed, and the same overall time of thirty seconds

was allowed for each item. In keeping with the previous discussion, and with the pilot study, the pairs of words were related in a bizarre or unusual way wherever possible. The keywords and images suggested for the two practice items were as follows:-

"armoire.....cupboard.....arm".

"Imagine a cupboard on the wall. Suddenly the door bursts open and a big hairy arm comes out".

"cartable.....satchel.....car".

"Imagine a car driving past you. Your satchel gets caught on its bumper, and it drives off down the road, dragging your satchel behind it".

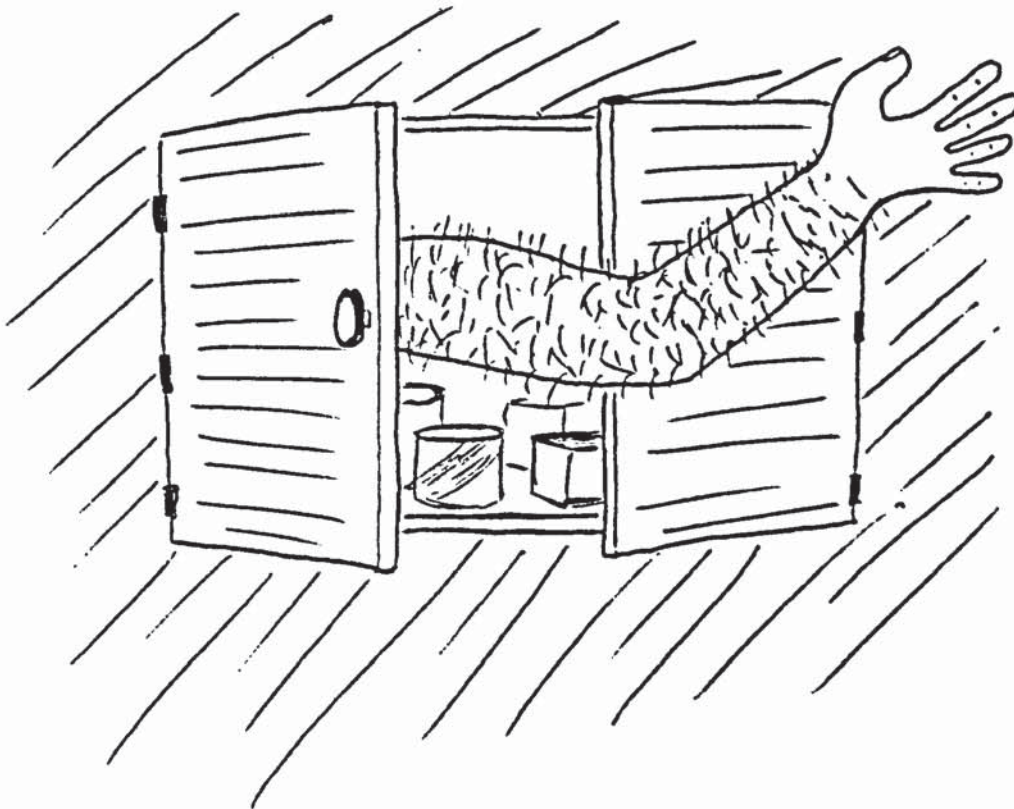
(For a full list of keywords and images, see Appendix C).

"Picture" Group (Two classes, $n = 49$) This group received identical treatment to that of the Image Group except that as E described the image, a line drawing was exposed below the three words. The descriptions of the images were the same for both groups, though E also pointed to the relevant aspects of each drawing as the image was described. (It would perhaps also be noted that the pictures were far from professional, and were certainly no better than most teachers could produce in a couple of minutes. This is not a trivial point - if the method works, it would be advantageous if the effect did not depend on complex, professionally produced pictures (c.f. The introduction to this experiment, section vi). Each pair of words was then repeated, exactly as with the other groups. (Copies of all the pictures

used also appear in Appendix C).

"armoire.....cupboard.....arm".

"Imagine a cupboard on the wall. Suddenly the door bursts open and a big hairy arm comes out".



"cartable.....satchel.....car".

"Imagine a car driving past you. Your satchel gets caught on its bumper, and it drives off down the road, dragging your satchel behind it".

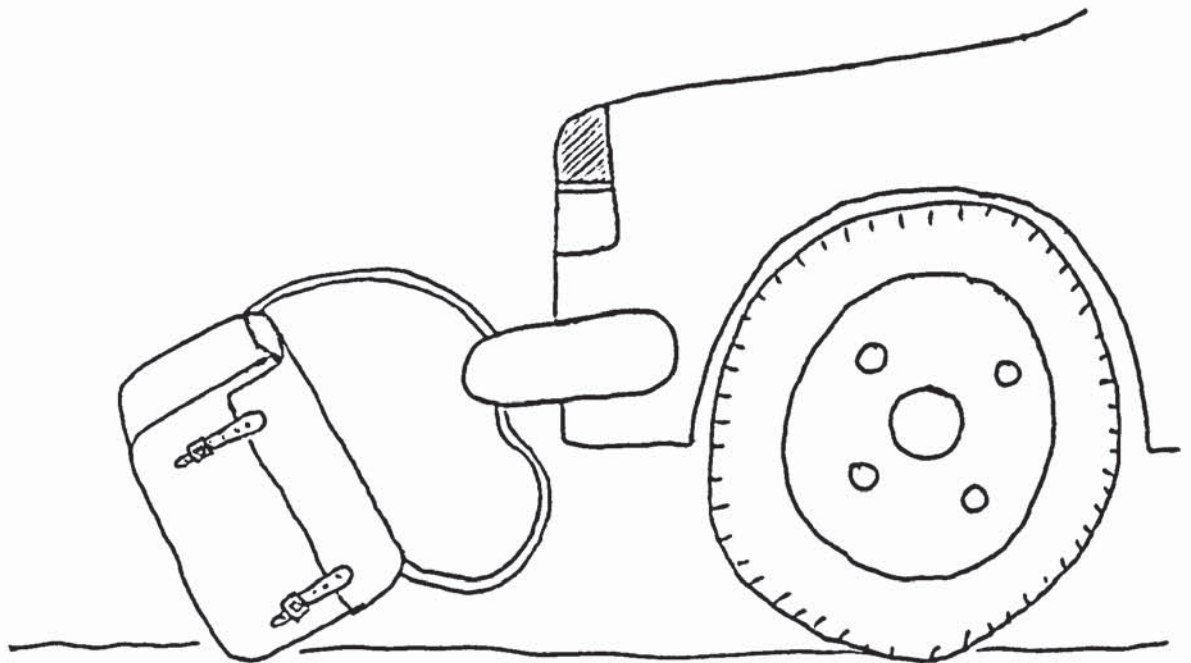


Figure 5 Practice items for "picture" group, main vocabulary study.

At the end of the seventeen learning trials, paper was given out, and all the subjects were tested.

Tests

All subjects were given exactly the same two tests, both involving all seventeen words.

Test I French into English

The procedure here was very similar to that used in the pilot study. E wrote each of the French words on the blackboard in turn, and pronounced it twice. Subjects had to write down the English equivalent, with no time limit set. Test one always consisted of the words in the following random order - tomber, vilain, ami, travailler, fatigue beaucoup, epicerie, matin, lancer, attraper, sonnerie, manger, sage, trouver, porter, tasse, tard.

Test II English into French

As before, a recognition task was used, possibly making this test generally easier than test I for most subjects. All the French words already on the board were re-labelled with a letter of the alphabet, and E read out the English words in a different order, or the corresponding letter of the alphabet, or both as they wished. Test two always consisted of the words in the following random order: cup, many, to work, to catch, morning, to eat, to fall, good, grocer's, to find, naughty, friend, bell, to throw, tired, late, to carry.

Subjects were not given the correct answers after the tests.

Rating

After the tests, all subjects were asked to rate the method they had been given to learn the words, using a simple five point scale. E discussed this briefly with each group, and asked them to be honest - nobody would be upset if they said they didn't like the method used, and everybody would have a different opinion anyway. It was also stressed that the tests themselves were not to be considered as part of the method. The five points on the scale are given below:-

Liked it a lot.

Quite liked it.

OK - not bothered either way.

Didn't like it much.

Hated it.

Subjects were asked to copy down the most appropriate phrase, to avoid any confusion, and their responses were scored from 1 - 5 by E later, with "Liked it a lot" scoring 5 and "Hated it" 1.

Exactly two weeks after the first learning and testing session, all available subjects were re-tested on exactly the same two tests as before. None of them had been warned that this would happen. These two tests are therefore:-

Test III French into English (delayed)

Test IV English into French (delayed)

Finally, an attempt was made to find out what strategies had been used for each item by subjects in Control Group II (no instructions) and if any of the subjects in the Rote Learning Group had used methods other than rote for any of the items.

Expectations based on the literature, and discussion with the pilot study control group suggested five possible categories. All subjects in Control Group II were given another sheet of paper, and E went slowly through each pair in the order presented in tests I and III. Subjects were asked to indicate which of the following five strategies they had used for each individual word. The two practice items were used as examples in each case.

"English" "The French word reminded you of an English word, and you used this to help you to remember the French word". (This strategy is related to the "acoustic link" part of the keyword method).

"Image" "You made up a picture in your mind to help you to remember the word". (This strategy is related, at least in part, to the "visual link" in the keyword method). This strategy was called "picture" for the benefit of the children, but will be referred to as the "Image" strategy here to avoid confusion with the experimental group given an external picture as well as keyword instructions.

- "Repeat" "You repeated the two words several times to yourself to help you remember". (This strategy had appeared very popular in the pilot study control group).
- "Other" "You used some other way of learning the words". E encouraged subjects to write down any details - including the possibility that they already knew the word.
- "Don't Know" "You don't really know how you tried to learn the words, you just learned them".

It was stressed that subjects might put down more than one category for any word, and that some might have used the same method throughout, while others used several different methods. Care was also taken to avoid the suggestion that the "don't know" category implied low ability, or that any one answer was somehow "better" than any other. The children were also encouraged to write down any other methods they might have used, or to say if they had already known any of the words, and these responses, amounting to about 5% of the total, were collected into an "Unclassified" category to be analysed later. The possibility that some of the children might find this rating task difficult had been considered, and it was also realised that those who remembered the most words might also be more likely to remember what strategies they had used. In fact, neither of these seems to have happened, as the discussion of the results will show.

(f) Results

The analysis of the results of the main experiment can be divided into two sections:-

- (i) The breakdown of the reported strategies used by the Control Group and their relative success, and
- (ii) the comparison of the control and various experimental treatment groups.

This perhaps slightly unorthodox presentation of results was chosen because causal relationships could not be inferred from the control group data alone - any apparent link between reported use of a strategy and success in the test could be due to a third factor, for example. However, the experimental groups' data can be seen as a logical extension of the idea that strategies are related to success, and in this case a causal relationship is much more likely, though other factors can never be entirely ruled out.

(i) The Control Group

Table 13 shows the correlations between the four test scores, for the control group alone. (As in all other tables, Test I FE IMM is the French into English test given immediately after learning, Test II EF IMM is the English into French test given immediately after learning, while Test III FE DEL and Test IV EF DEL are the same two tests repeated after a delay of two weeks).

	Test I	Test II	Test III	Test IV
Test I FE IMM	-	+0.93	+0.84	+0.81
Test II EF IMM		-	+0.82	+0.85
Test III FE DEL			-	+0.91
Test IV EF DEL				-

Table 13 Control group (n = 51). Correlations between the four vocabulary tests.

All the correlations are significant at well beyond the .001 level (two-tail test); it was thus felt that they were largely measuring the same thing, and that scores could be combined if necessary.

The next step was to analyse the use of strategies as reported by the control group, and compare each strategy with its relative success on the tests. Table 14 thus shows how frequently each strategy was reported and, when it was used, the likelihood of successful recall.

Strategy Reported	% Frequency Reported	% Success of Strategy on Recall tests
Rote	36	30
English	28	56
Don't know	22	25
Image	4	59
Rote and English	2	59
Im. and English	1	88
Im. and Rote	(0.3)	(33)
Unclassified	5	72

Table 14 Reported use of strategies and their success in recall tests, control group, (to nearest%).

(Descriptions of each strategy were given in the previous section). Thus, for example, rote repetition, whilst used spontaneously most of all, gave only a 30% success rate; the "Image and English" combination, involving aspects related to both parts of the keyword method, was used very rarely, but resulted in 88% success when it was used.

However, the various combinations of strategies were all reported very rarely, and little reliance can be placed on their apparent success rates. The "Image and English" combination was in fact only reported on sixteen occasions, fourteen of which involved successful recall. Analysis of the apparently highly successful "unclassified" category revealed that, when an explanation was given, it was almost always that the subject already knew the word (e.g. "learned it at other school" or, in the case of "sonnerie", "knew a French song"). Only in a very few cases was a genuinely unclassifiable response given. For example, one girl explained how she learned "tomber" by saying "I rymed it with tumble which means fall". This was counted as use of the "English" strategy, but when she simply said "I rymed it" as her explanation of how she learned "épicerie", this had to be counted as "unclassified".

Discounting the very infrequent reports of combinations of strategies, and the "unclassified" strategy on the grounds that it almost always indicated previous knowledge, the actual frequencies for the remaining four strategies were re-cast as in Table 15.

Reported Strategy	Number of Wrong Answers	Number of Correct Answers
Rote	440	186
English	208	268
Don't Know	286	96
Image	30	44

Table 15 Frequency of correct and wrong test answers, according to reported strategy, control group.

A chi-square test gave a result of $\chi^2 = 128.2$, with 3 d.f., significant at far beyond the .001 level, suggesting that there are significantly different proportions of right and wrong answers among the various strategies.

The next stage involved the isolation of two sub-groups to represent those who had done particularly well on the four tests, and those who had done particularly badly. Cut-off points are obviously arbitrary, and in the end the two groups chosen were subjects scoring over the 75% ile and below the 25% ile. Analysis by t-test showed that the two groups differed significantly on their overall scores and reading ages, and separate chi-square tests were used to test the frequency data showing how often each sub-group reported each strategy.

	Total Score	Reading Age	Reported Frequency of Strategy				
			Eng.	IM.	Rote	Unclass.	D.K.
High Scorers (over 75% ile)	52.8	12.6	8.3	0.9	3.8	1.8.	3.0
Low Scorers (below 25% ile)	11.4	9.8	2.3	0.3	10.5	0.8	3.5
Significance of difference	$p < .01$	$p < .01$	$p < .01$	n. s	$p < .01$	$p < .05$	n. s

Table 16 Control group, high and low scorers. Comparison of mean total scores, reading ages and reported frequency of each strategy.

The results are shown in Table 16. The maximum total score on all four tests would be 68, and the maximum score for any strategy would be 17, indicating that it was used to learn all the items, either alone or in combination with another strategy. The frequency total for the High Scoring group is 17.8, indicating that, on average, they reported using two strategies combined only 4% of the time. Similarly, the total for the low scorers is 17.4, indicating use of two strategies combined only 2% of the time. Combinations of strategies were thus reported very rarely by either group, as one would expect from the previous analysis. In other words, the superiority of the high scoring groups would appear to be related to their spontaneous use of single strategies rather than combinations of them.

The low-scoring group was exactly equally divided between boys and girls, but the high-scoring group contained three times as many girls as boys, suggesting that there could be sex differences in performance on the task. This was examined in more detail by comparing the mean reading ages, scores and reported use of strategies for boys and girls in the control group separately. The results are shown in Table 17.

	Total Score	Reading Age	Reported Frequency of Strategy				
			Eng.	IM.	Rote	Unclass.	D.K.
Boys (n = 23)	25.0	10.9	4.0	1.0	6.9	0.4	5.1
Girls (n = 28)	34.5	10.8	6.1	0.9	6.6	1.1	2.9
Significance of difference	$p < .01$	n.s.	n.s.	n.s.	n.s.	n.s.	$p < .01$

Table 17 Control group, boys and girls. Comparison of mean total scores, reading ages and reported frequency of each strategy.

(As before, t-tests were used for the scores and reading ages, and separate chi-square tests for the frequency data). These results suggest that, although the difference in total scores was significant, the differences between reported use of the various strategies were not, except for the "don't know" category, which was used significantly more by the boys. In spite of the lack of significance, however, the results might suggest

that the girls did use the "English" strategy slightly more, and perhaps also knew more of the words to begin with, according to the trend of results in the "unclassified" category. Lack of significance makes this obviously no more than a tentative suggestion, however.

Discussion

Taken as a whole, the results of the control group suggest several things about the various spontaneous strategies used, which in turn could be related to the results of the experimental groups actually instructed to use some of these strategies.

Rote Repetition

Was the most popular strategy spontaneously reported by the control group, though its success as a learning device was apparently poor, the only worse category being reports of "don't know" (i.e. "I don't know how I learned the words, I just learned them"). Moreover, children who did badly on the tests reported using rote repetition much more often than children who did well, and it was by far the most popular method used by the low scorers.

(196)

"English" strategy was where the French word reminded the subject of an English word, and was thus related to the "acoustic link" of the keyword method. In the group as a whole, it had almost double the success rate of rote repetition, though it was used only about a quarter of the time. However, this proportion is doubled in the high-scoring group, who used the strategy for about half the words, and for whom it was easily the most popular method. In contrast, the low scorers much preferred rote learning, and used it over four times as often as the "English" strategy. For both these strategies, the differences between reported use by High and Low scorers are highly significant.

"Image" Strategy was the deliberate use of images to aid recall and its spontaneous use was very rare - only 4% of the time overall. Like the "English" strategy, it gave roughly double the recall of rote repetition, and although it was reported more frequently by the high scorers than the low scorers, the difference was not significant.

Combinations of these three strategies were also very rare, and account for less than 4% of all the responses. It is tempting to conclude that the combination of "Image" and "English" strategies, which involves both parts of the keyword method, was by far the most effective (88%), but there are obvious dangers in forming strong conclusions on the basis of so little data;

(197)

the apparently poor result of the "image and rote" strategy should also be interpreted with care, since the reported success rate is well below that found for "image" alone. The only safe conclusion is that combinations were rare, and do not appear to differentiate between high and low scorers.

"Don't Know" as one might expect, had the lowest success rate of all, though it was reported 22% of the time by the group as a whole. More interesting is the similarity between high and low scorers in this category - no significant difference emerged. This is in fact a very important point, since it had been thought that any differences in reported strategies of good and poor learners might involve the superior ability of the good learners to remember or label the strategies they had used. However, this does not seem to be the case, since both groups used the "don't know" category about the same number of times, and the difference between them was not significant. (Table 16). In other words, the low scorers appear to be just as good at remembering and categorising their strategies as the high scorers were.

"Unclassified" responses, as previously indicated, tended to be almost always cases where the children already knew the word, and appeared significantly more often in the high scorers than the low scorers. However, they appear to have had little

bearing on the total success rates - if the differences in prior knowledge are allowed for, the mean score of the top group is still 11.6 out of 17, in contrast with 2.1 for the bottom group. While the high scorers thus knew on average about one word more than the low scorers to begin with, this difference does not account for much of their superiority. In fact, prior knowledge would account for 14% of their scores, in comparison with 19% of the scores of the bottom group.

In conclusion, the superior results of the top group do not appear to be due to their use of combinations of strategies, their ability to recall or categorise what strategies they used, or their previous knowledge. On the other hand, there are highly significant differences between the two sub-groups' reported use of rote repetition and the "English" strategy. Taken overall, the results of the control group thus suggest that some strategies are more effective than others, and that children who did well tended spontaneously to use different strategies from children who did badly. However, all the results presented so far have been in the nature of correlations, and causality cannot be inferred - use of certain strategies and high scores could both be due to some other factor, for example. Moreover, these results alone obviously do not indicate whether instruction in the use of strategies might be beneficial. The spontaneous use of imagery was reported very rarely, for example, so could actual instructions involving the use of imagery improve scores? Conversely, rote repetition was frequently reported, particularly

by less successful learners, so what would be the effect of actual instructions to use it? The use of the various experimental groups can be seen as an extension of the correlational results so far presented, and these results are given below.

(ii) The experimental Groups

Tables 18 to 20 show the correlations between the four tests for each of the three experimental groups. As with the control group, all the correlations are highly significant ($p < .001$, except for the correlation between tests II and III for the Rote group, where $p < .003$.)

	Test I	Test II	Test III	Test IV
Test I FEIMM	-	+0.78	+0.84	+0.82
Test II EFIMM		-	+0.60	+0.71
Test III FEDEL			-	+0.85
Test IV EFDEL				-

Table 18 Rote group (n =23) Correlations between the four vocabulary tests.

(200)

		Test I	Test II	Test III	Test IV
Test I	FEIMM	-	+0.90	+0.64	+0.71
Test II	EFIMM		-	+0.67	+0.73
Test III	FEDEL			-	+0.93
Test IV	EFDEL				-

Table 19 Image group (n = 46) Correlations between the four vocabulary tests.

		Test I	Test II	Test III	Test IV
Test I	EFIMM	-	+0.83	+0.79	+0.68
Test II	FEIMM		-	+0.69	+0.69
Test III	EFDEL			-	+0.88
Test IV	FEDEL				-

Table 20 Picture group (n =46) Correlations between the four vocabulary tests.

It was thus felt that scores on the four tests could again be combined if necessary, for reasons of simplicity, while still avoiding any direct comparison between the recall (French into English) and recognition (English into French) tests. The data from the main study are obviously more complex than the equivalent data from the pilot study, and overall results are therefore presented in the form of histograms, for ease of comparison.

**PAGE
MISSING
IN
ORIGINAL**

Figure 6 shows the immediate and delayed recall scores for all groups, and Figure 7 the immediate and delayed recognition scores. As expected, the "test only" group, who received no training at all, did very badly, with a mean score of 0.5 out of 17 on the French-into-English test, and 2.3 out of 17 on the English-into-French test, where they could guess at the answers. Their results establish a baseline against which to judge the scores of the other groups. The rote group appears to be rather better than the control group on all tests. (25% better on Test I, 14% better on Test II, 20% on Test III and 21% on Test IV). The image group is more clearly superior to the control group on all four tests (64%, 48%, 90% and 68% respectively), while the superiority of the group also given a picture is even more marked. (79%, 59% 111% and 84% better than the control group on the four tests respectively). The superiority of the image and the picture groups appears greater the more difficult test - the unexpected delayed recall test (FEDEL) was almost certainly the most difficult, and all groups recorded their lowest score here, but it is on this test that the superiority of the two keyword groups is most apparent. In fact, the pictures may have been even more helpful than the results suggest, since there was a ceiling effect here.

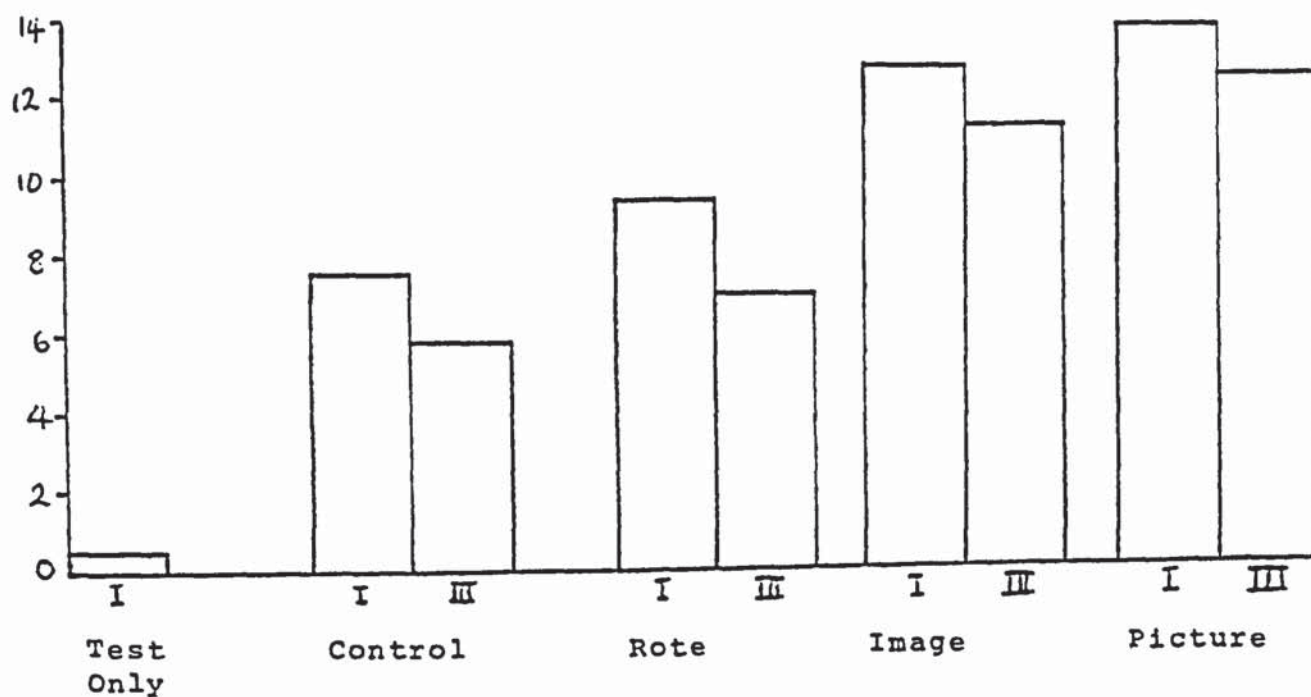


Figure 6 Recall scores of control and experimental groups on Test I (FEIMM) and Test III (FEDEL).

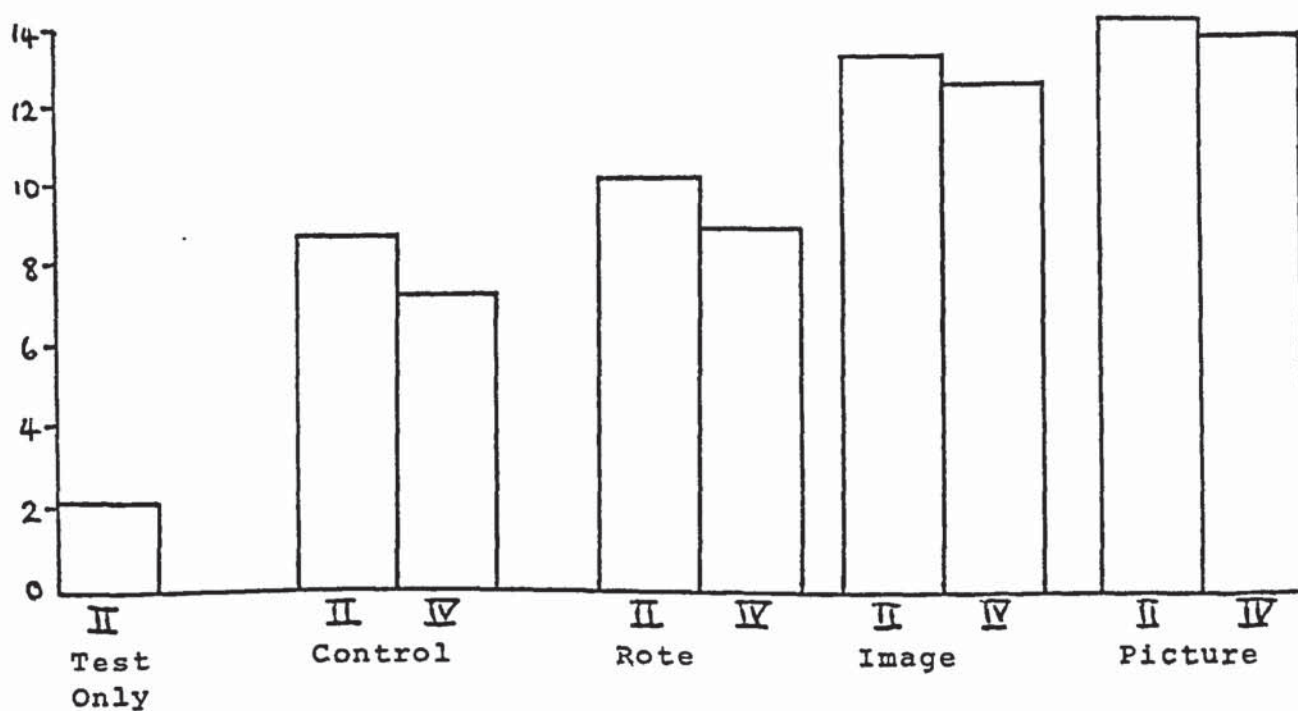


Figure 7 Recognition scores of control and experimental groups on Test II (EFIMM) and Test IV (EFDEL).

Analysis of variance was used to compare the mean scores of each treatment group on the reading age scores, the scores on the four tests, and the ratings of each method. Tables 21 to 26 present the results.

Condition	Mean Reading Age	Standard Deviation
Control	10.4	2.8
Rote	10.3	0.9
Image	10.8	1.5
Picture	10.5	2.4
Test Only	10.5	1.3

Source	s.s.	d.f.	m.s.	F.	Sig.
Between groups	6.2	4	1.5	0.36	n.s
Within groups	814.8	190	4.3		
Total	821.0	194			

Table 21 Reading age by condition, with analysis of variance. This result suggests that the various control and experimental groups did not differ significantly in reading age, so that any differences in their scores cannot be due to this factor.

(205)

Condition	Test I (FEIMM) Mean Scores	Standard Deviation
Control	7.7	4.2
Rote	9.4	3.1
Image	12.7	3.7
Picture	13.7	3.6
Test Only	0.5	0.8

Source	s.s.	d.f.	m.s.	F.	Sig.
Between Groups	2547.3	4	636.8	50.1	.01
Within Groups	2416.5	190	12.7		
Total	4963.8				

Table 23 Scores on Test II (EF IMM by condition, with analyssis of variance.

Condition	Test III (FEDEL) Mean Scores	Standard Deviation
Control	5.9	4.3
Rote	7.0	4.5
Image	11.2	4.7
Picture	12.4	4.6

Source	S.S.	d.f.	m.s.	F	Sig
Between groups	1323.2	3	441.1	21.5	.01
With groups	3326.1	162	20.5		
Total	4649.3	165			

Table 24 Scores on Test III (FE DEL) by condition, analysis of variance.

(207)

Condition	Test IV (EFDEL) Mean Scores	Standard Deviation
Control	7.6	4.5
Rote	9.2	3.9
Image	12.8	4.3
Picture	14.0	3.8

Source	S.S.	d.f.	m.s.	F	Sig
Between groups	1221.0	3	407.0	23.4	.01
Within groups	2813.0	162	17.4		
Total	4034.0	165			

Table 25 Scores on Test IV (EFDEL) by condition, with analysis of variance.

Table 26 shows how each group rated the method it used for enjoyment, on a 1-5 scale with 5 indicating "liked it a lot" and 1 "hated it". Again, as analysis of variance showed, there were highly significant differences between the groups.

Condition	Mean Rating	Standard Deviation
Control	3.2	1.1
Rote	2.7	0.8
Image	3.6	0.9
Picture	4.2	0.9

Source	S.S.	d.f.	m.s.	F	Sig
Between groups	48.2	3	16.1	17.3	.01
Within groups	159.6	172	0.9		
Total	207.8	175			

Table 26 Rating by condition, with analysis of variance.

Taken together, these results suggest that, although the various treatment groups did not differ in reading age, there were some highly significant differences between their scores on each of the four tests, and on their ratings of the method they used. In each case the level of significance was far beyond .01

Not all the differences were significant, however. Separate t-tests showed that the difference between the rote and control groups was barely significant on immediate recall ($t = 1.96$, with 76 f, $p < .05$, 1-tail test only) and was not significant on the delayed test. ($t = 1.03$, with 72df, $p > .05$). Similarly, the picture group's score was not significantly higher than the image group's on either immediate or delayed recall. ($t = 1.4$, with 98df, $p > .05$, and $t = 1.2$ with 90df, $p > .05$, respectively).

The final stage of the analysis involved a regression program to assess what proportion of the total variance was due to the different ratings, treatments, reading ages and sex differences, and to possible interactions between them. When ratings were included in the analysis, they had only a very small effect, explaining less than 0.2% of the total variance above and beyond that caused by the other factors ($F = 0.1$, $p > .05$), and suggesting that the children's ratings of enjoyment did not directly affect their results. However, the other variables did appear significant, and table 27 shows the results of the analysis of variance.

Source of Variation	Proportion of Variance	d.f.	m.s.	F	Sig
Covariates	0.10536	2			
Sex	0.01454	1	0.01454	4.6	.05
R.A.	0.08808	1	0.8808	28.1	.01
Treatments	0.25655				
Interactions	0.06353	6			
Sex X Treatment	0.00843	3	0.00281	0.9	ns
R.A.X Treatment	0.05717	3	0.01906	6.1	.01
Residual	0.57456	183	0.00314		
Total	1.00000	194			

Table 27 Analysis of variance - interactions between covariates (sex and R.A.) and treatment conditions.

These results suggest that, taking all four tests together, sex and reading age both had significant effects on the total variance - reading age about 9% and sex about 1%. The various treatments had a much larger effect, and explain about 26% of

the total variance. As for interactions, the effect of treatments does not differ by sex, but there was a significant interaction between treatment and reading age. For this reason, the relationship was examined in more detail by splitting the reading age scores into three groups of roughly equal size and working out separate means for each treatment. Table 28 shows the result.

	"Low" R.A. (Under 10)	"Middle" R.A. (10 - 11)	"High" R.A. (Over 11)
Control	15.3	30.1	39.4
Rote	32.5	34.8	48.8
Image	39.8	47.6	57.8
Picture	46.5	48.5	64.4

Table 28 Mean total scores (out of 68) by reading age and treatment.

Taking each level of reading age separately, the trends are fairly clear. Firstly, as the analysis of variance suggested, reading age is clearly related to success on the tests, and the different strategies had different effects on the three sub-groups. In comparison with the middle and high groups, the low

group benefited a great deal from rote instructions, which more than doubled their score. In contrast, the middle group benefited much less from rote instructions, but improved considerably when given the keyword mnemonic. Interestingly, the addition of a picture hardly helped at all, though it improved the scores of both the high and low R.A. groups to the extent that half the high group obtained maximum scores on all four tests, and the low group mean rose to 46.5, a score clearly better than that of the high group subjects left to their own devices.

(g) General Discussion

The results of the control group suggested that the spontaneous use of different strategies was related to success in the tests, taking both the group as a whole, and in the comparison of the reported strategies of successful and poor learners. The results of the experimental groups extend these findings and strongly suggest that children with little experience in French, can greatly improve their vocabulary learning if they are provided with a mnemonic strategy, even though they would hardly

ever use such a method spontaneously. In this respect, the results generally support the findings of Ott and his associates and the Stanford research, and extend the application of the keyword method to children's learning of "real" vocabulary and their ability to recognise the foreign equivalents of English words, as well as to recall what the foreign words mean, on both immediate and delayed tests.

Rote repetition was the most common spontaneous strategy reported in the control group, and although the experimental subjects actually instructed to use rote appeared to score slightly higher than the control group on the recall tests, the difference was barely significant on immediate recall ($t = 1.96$, with 76 df, $p < .05$, 1-tail test only) and was not significant on the delayed test ($t = 1.03$, $df = 72$, $p > .05$). One problem frequently encountered in the research on student subjects is their failure to follow mediation instructions, though it does appear more likely that children will follow such instructions (e.g. Bull and Wittrock, 1973; Wittrock and Goldberg, 1975). Because of this, the subjects in the Rote group were asked if they had used any other methods to learn the words, and six of them indicated that they had, though only for an average of four words each. No spontaneous use of imagery was reported, but twenty-four words

altogether had been learned by noticing a similarity with an English word - the "English" strategy in fact. Recall of these words was roughly double the recall of words which the same subjects reported learning by rote, adding more support to the general pattern that has emerged.

In terms of a "depth" approach to memory, which will be discussed in Chapter Four, rote learning probably involves only shallow processing: repetition simply prolongs the stimulus, without enriching or transforming it in any way (c.f. Anderson, 1972; Swenson and Kulhavy, 1974; Pressley, 1977). In contrast, the keyword mnemonic involves both visual and verbal mediation - the recognition that the foreign word is similar in some way to an English word, and the generation of a mediating image. When left to their own devices, it appears that the children in the low scoring group tended to use the "shallow" method of rote repetition, while successful learners often tried to form a mediating link for themselves, and thus processed the information more deeply. However, there are dangers in equating the control group's reported spontaneous strategies with the image and acoustic links provided by the experimenter.

Taking the "image link" separately, the value of the actual image, as opposed to a purely verbal link, is a matter of dispute: Anderson and Bower (1973) found that imagery instructions gave the same result as sentence instructions, while Atkinson (1975) claims that images result in significantly superior recall in the Keyword mnemonic, quoting figures of 73% as opposed

to 64%. Although the role of imagery has been emphasised right from the introductory chapter, there is always the problem of being sure that it was actually used in any experiment, and it is always possible that the children in the keyword and picture groups did not use imagery at all. In spite of this possible objection, however, instructions to form images as part of the keyword strategy certainly improved the scores of the children in the present study. There is also the possibility that many of the children in the control group did form images while they were learning the words, but did not report them because imagery is so much a part of their normal thinking anyway. On the other hand, the keyword mnemonic should involve the deliberate and controlled use of imagery in a way that may be very different from casual, spontaneous images formed for no deliberate purpose. In other words, there may be problems in trying to equate the children's spontaneous reports of an "Image" strategy with the use of images in the keyword mnemonic.

A similar possibility arises with reports of the "English" strategy, which seems similar to the "acoustic link" of the keyword technique. It is possible that control group subjects reporting the "English" strategy noticed a similarity with an English word, but did not actually use it as a memory aid (c.f. previous discussion on "mediation deficiency"). In spite of

these possible objections, however, the results of the control group (tables 14 - 16) do suggest some spontaneous and conscious use of effective mediators similar to the "acoustic" and image" links of the keyword method, especially by the more successful subjects, and the results of the Image group certainly imply that children can use the method when actually provided with both the necessary links by the experimenter.

The provision of an external picture also involves the controversy discussed earlier about experimenter-supplied versus subject-generated mediators. In spite of the attempts to avoid the ceiling effect that had been noted in the pilot study, 26% of the subjects in the Picture group scored seventeen out of seventeen in all four tests (i.e. thirteen subjects, in contrast to two in the Image group and one in the Control group). The picture condition also appeared very beneficial to the low reading age group, where there was most room for improvement. An external picture and the keyword instructions more than tripled their scores in comparison with the low reading age subjects in the Control group, and brought their mean score to a level distinctly higher than that of Control group subjects in the high reading age group (Table 28).

Previous discussion has suggested that the value of a picture would depend on factors like the age of the subjects and the nature of the material and test. Atkinson (1975) describes

unpublished studies suggesting that students perform better when asked to generate their own images in the keyword system, rather than use "cartoonlike drawings". Ott et al (1973) found that provided pictures like that in figure 3 improved immediate but not delayed recall in their students. Separate t-tests on immediate and delayed recall scores in the present study showed that the superiority of the Picture group over the Image group was not significant in either case ($t = 1.4$ with 98df, $p > .05$, immediate recall; $t = 1.2$ with 90df, $p > .05$, delayed recall), though the lack of significance could be explained by the ceiling effect. Further studies could increase the number of words given, but the questionnaire data reported in Chapter Two suggested that in practice, teachers rarely expected their pupils to learn more than twenty words at a time. Given this target, the extra work for the teacher involved in producing pictures might not produce a significant improvement, but their provision could perhaps allow an increase in the number of words to be learned, this being one area of possible further research.

The practical applications of the keyword method raise several other questions, however. One possible danger is that, since the whole point of redintegrative processes like mnemonics is that they involve the generation of additional elements beyond the presented material, these additions could result in

interference. A nice example is provided by Bower (1973), who describes attending a meeting of mnemonic experts. As he left, one of them shook his hand, looked him squarely in the eye and said, "It's been nice meeting you, Mr. Flowers"! In the present study, taking the practice item "armoire.....arm....cup-board" as an example, subjects might remember the keyword (arm) but find that this interfered with recall of what the French word actually meant. Examination of the errors made by the children suggested that this was in fact quite rare - only occasionally did subjects in the Image or Picture groups put that "lancer" meant "lance", for example. One exception was the "false friend" - "travailler", where a number of children did answer "travel" (the keyword) instead of "work". However, examination of the other groups showed that this particular mistake was equally common there, suggesting that it was not necessarily interference from the keyword that had caused the error. A less direct form of interference was the occasional "semantic intrusion", where interference from the image link, rather than from the actual keyword seemed to have occurred. For example, given "tomber.....Tom thumb.....fall" and the description (with or without a picture) of a tiny Tom thumb walking across your desk and tripping over your thumb, a few children in the keyword group put "little" or "small" as the meaning of "tomber". However, both kinds of interference were quite rare, and the

research generally suggests that interference becomes less likely as time goes on. Prytulak (1971), for example, found that language mediators "appeared to be prevalent during early learning, but to disappear with continued practice". Similarly, Paivio (1971) states that "associative aids are used during the initial stages of associative learning, when the task is difficult. As training progresses, mediators drop out". In the keyword mnemonic, Atkinson (1975) concludes that:-

Early in the learning process, the memory structure for a given item involves only these two independent links; however, with continual practice, a third link is formed directly associating the foreign word with its English translation. It is this direct link that sustains performance once an item is highly practiced.

Another possible problem area involves the extended use of the keyword method over a long period of time, and there are perhaps two dangers, one specific and one general. The specific one involves the occasional use of the same keyword for different vocabulary items. For example, the keyword "mat" was used to recall "matelot" on the pilot study and "matin" in the main study, though with different subjects. The research literature on mnemonics is controversial on this point, however. Postman's "interference paradox" (Postman, 1962) had suggested that interference would occur if responses competed with each other, and Crovitz (1971), for example, did find that recall was poorer if the same "loci" were re-used to learn several lists. On the other hand, Ross and Lawrence (1968) found no such effect, and

Bugelski (1970) and Bower and Reitman (1972) both suggest that additional items can be incorporated into an existing image without interference. However, the ease with which extra items can be incorporated will depend partly on the nature of the image; Lesgold and Goldman (1973), studying the method of loci, distinguish between "expansive locative words" like "street" which contain many features (and could, by implication act as the keyword for several words at once), and "static locative words" like "television" which tend to give a stereotyped image and would probably cause confusion if re-used. Finally, it might also be wrong to consider each item as if it was learned in isolation: Hunter (1956) pointed out the importance of context cues in avoiding inter-list confusions, though these could still arise in delayed recall, which is of less concern to professional mnemonists than to teachers. As for the classroom, if Paivio, Prytulak and Atkinson are all correct, then the keywords will eventually disappear in any case, so that a gap of a few weeks between learning, say, "matelot" and "matin" should avoid any specific interference effects.

A more general possible danger is that the use of the keyword system over a long period might cause a decline in student interest which would make the method less effective. However, Bower (1973) points out that professional mnemonists maintain a

life-long enthusiasm (though it could be argued that they are hardly "normal" people!), while Atkinson (1975) reports that his students were still enjoying the keyword method at the end of a ten week course. The results of the present study suggested that the pupils' enjoyment was not an important causal factor in its own right, although they did rate the keyword method as more enjoyable than the rote or control conditions especially if a picture was also provided (Table 26). It is also possible that the picture and keyword simply provided extra cues for recall, rather than encouraging improved strategies, though Pressley's (1977) study, using control groups given either the picture or keyword without full instructions, suggests that this explanation is unlikely. However, research on strategies often involves making inferences about what subjects do and it is very difficult to rule out all other possibilities. For example, it is always possible that the sheer novelty of the keyword and picture conditions could have been partly responsible for their facilitation. Such factors can not be entirely ruled out, though further studies could examine the effects of long-term use of mnemonic methods, and an interesting development would be to see if children could be taught to generate effective mediators for themselves in a variety of situations.

Finally, it must be emphasised that the keyword method is certainly not being advocated as suitable for all new vocabulary items, or even as equally helpful for all children. The second chapter suggested theoretical reasons, based on recent research on learners' strategies, for the retention of vocabulary learning, and went on to show that many teachers actually set such work in practice anyway, while giving little advice as to how it could be done. The keyword mnemonic is offered simply as another technique, and apparently a very effective one in some cases, for the teacher to add to his existing repertoire; it certainly should not replace the presentation of new words in context, or the use of "flash-cards" or tapes, for example. As for the wider application of mnemonic techniques in other areas of the curriculum, this will depend on how much arbitrary and specific information children are expected to acquire. If all school material can be presented in a "meaningful" way, related to things the child already understands, and according to the best principles of cognitive learning as expressed by writers like Bruner and Smith, then mnemonics are clearly unnecessary. On the other hand, it may be that we still expect children to master a good deal of unavoidably arbitrary information in practice (c.f. Bower, 1973), in which case mnemonics offer a method of extending the principles of meaningful learning to the acquisition of materials which could otherwise be learned only by rote. Coming to that familiar conclusion "more research is needed", it is quite possible that mnemonic systems

like the keyword method have only very limited educational applications, but that two rather different research approaches need to be used to find out, for even if experimental studies can suggest such potentially useful techniques, they still need to be related to the actual demands of classroom tasks, which in turn must be examined by the established methods of applied research.

7

-ooOoo-

CHAPTER IV

SUMMARY

The fourth chapter provides an introduction for the second experiment, and can be seen as a parallel to the second chapter. It looks first at some current issues in English teaching as a background to a more specific discussion of the status of comprehension work. A survey of the literature suggests that most English curriculum specialists see the traditional comprehension exercise as a pointless activity largely irrelevant to current views about the nature and aims of English teaching.

However, the second part of the questionnaire survey suggested that the vast majority of children are still frequently given such exercises. As with French vocabulary learning, it would seem that an activity despised by those who write in the journals is still common, though little advice is given as to how it might be done.

Psychological research might offer some useful techniques for improving comprehension, and the final part of the chapter briefly surveys some recent research, particularly using a "levels-of-processing" approach, which seems to offer a framework for looking at some of the complexities of comprehension. However, the materials and tests used in such research often appear very different from those actually used in schools so that if claims of educational relevance are to be justified, the suggested techniques need to be re-assessed using "real" tests and materials.

IV 1 INTRODUCTION

In spite of the apparent success of the keyword mnemonic, it might be argued that the learning of individual words or other items is comparatively rare in schools and that mnemonics, which tend to apply only to specific pieces of information, are thus limited in value. However, experimental research in psychology has been moving on from the study of such individual items - Kintsch (1974), describes how articles in the Journal of Verbal Learning and Verbal Behavior have progressed from studies of (so called) "nonsense syllables" in the 1950's, to lists of words in the 1960's and to sentences in the 1970's. This development can be related to the influence of Behaviourism, with its emphasis on individual words as S - R units, giving way to the influence of Generative Linguistics, with its emphasis on the syntactical structure of sentences. (Paris, Mahoney and Buckhalt, 1974). The recognition of the importance of semantics has meant that psychologists are in fact now beginning to study subjects' comprehension and recall of passages of connected discourse, and such research would seem very relevant to much of the work in schools - the researchers themselves have frequently claimed that their studies have educational implications, especially if they have used children as their subjects.

The second area to be examined here is therefore children's reading comprehension of connected passages. Although this is

obviously important in all academic subjects, English is still the only one that consistently involves the comprehension of non-specific information as a valid activity in its own right rather than as a means of gaining information about some body of knowledge. Consequently, the first part of the chapter will limit itself to the role and status of comprehension work in English only, and will not be concerned with the added problems of different sorts of comprehension which might be involved in reading a history textbook as opposed to a chemistry textbook, for example. The nature and relevance of any one part of a subject must however be seen against the background of current ideas and general issues in that subject so that, just as it was felt important to suggest some of the current issues in modern languages before looking in detail at vocabulary learning as one aspect, it is now proposed to discuss some of the issues which seem to concern English teachers, as a background to the second specific topic of reading comprehension.

IV 2 SOME CURRENT ISSUES IN ENGLISH TEACHING

In parallel with the section of the second chapter on current issues in French teaching, this section can be divided for convenience into Hirst's same three areas - objectives, content and methods, bearing in mind that, although this

division provides a useful framework, the distinctions between the three are often rather arbitrary. Another useful framework is D'Arcy's (1973) suggestion of three basic models of English teaching - the "skills" model, the "cultural heritage" model and the "personal growth" model, each of which can be seen as implying different objectives, content and methods.

(a) Objectives

In terms of objectives, the "skills" model emphasizes correctness and refinement in the use of language, according to some externally imposed standard. Thus, Douglas (1972) describes the first two aims of such a course, published 50 years ago, as -

The ability to spell correctly without hesitation all the ordinary words of one's writing vocabulary
 The ability to speak, in conversation, in complete sentences, not in broken phrases.

The same sorts of objectives can be seen, if rather less bluntly stated, underlying the introductory discussion provided by Pendlebury (1958) thirty years later.

Perhaps the main difficulty in teaching English is to combine liveliness with method. The teacher wants to stimulate the pupils' minds by providing them with exercises that are piquant, varied, and sometimes amusing; yet at the same time he does not want to find at the end of the year that there have been serious omissions in the programme of learning and practice.

In these terms, the attempt to "stimulate the pupils' minds" is really little more than a device to get them to concentrat

on the more important aims of "the programme of learning and practice" emphasised by the "skills" model's approach to language as a body of knowledge to be mastered.

In a way the "cultural heritage" model is an extension of this view, since its basic aim is to introduce the child to the "Great Writers" who embody the best use of language, though again using criteria that may be very different from those of the children themselves, imposed by those who feel they have the necessary authority. Such a view is implied, for example, in the H.M.I. report of a survey of English lessons carried out in 1971. Having observed 140 lessons, they comment that "Three contemporary authors of quality were represented" without bothering to name them or to justify their decision that there were three, no more or less. Jones (1974) describes the central aim of this approach as the attempt "to emphasize the leading names and ideas in English writing so as to fix these in the minds of pupils", offering Shakespeare and Lamb in the third year and running the risk of encouraging what Jones scathingly describes as a "top of the form mentality." More sympathetically, the Newson Report (1963) considered that

all pupils, including those of very limited attainments, need the civilising experience of contact with great literature, and can respond to its universality sympathetically presented, literature can stretch the minds and imaginations of the pupils, and help to illumine for them, in wider human terms, their own problems of living.

Such reasons for studying the "Great Writers" thus merge into the personal growth model, and Stratta (1972) sums up the difference -

Twenty years ago a teacher of English might have confidently asserted that his main concern was to introduce his pupils to literature. Today he might less confidently assert that his main concern is with the process of helping his pupils to develop their abilities in using the language for a variety of needs and purposes.

The Newson report had itself emphasized that "the overriding aim of English teaching must be the personal development and social competence of the pupil and it is this basic division between the imposed standards of correctness on the one hand, and the personal development of the individual on the other, which lies at the heart of the whole "traditional V progressive" debate in English Teaching.

This division is a central one, and it appears equally in the content and methods of the English curriculum. It is embodied even in Mr Stuart Froome's "note of dissent" in the Bullock Report (1975), whose long-awaited recommendations can be seen as a culmination of widespread thinking about English in schools. It would seem obvious that both aspects are important - the child must feel he has something worthwhile to express, and should benefit from its expression, but he must also be able to communicate it according to the

accepted rules that allow others to understand him. If the Bullock Report was expected to settle the debate once and for all, it was doomed to failure. Writing at the time of its publication, in the T.E.S., "Aristides" concluded, rather wistfully -

The committee now hope that their report will knock simplistic arguments about prog V trad., basic skills V creativity etc on the head for good. Some hope.

What has emerged, as so often, is something of a compromise, and the report itself reflected this. Spelling and grammar, emphasized by the "skills" model, should not be taught as separate entities, but should ideally spring from the child's own work, and appear as relevant to him rather than as artificial "exercises" to be performed in isolation from everything else. The report concludes that, "provided teachers work on intentions first, then on techniques to match them, a good deal of spontaneity can survive the transition from artlessness to art."

If a compromise of objectives really is desirable and obtainable, then it should ideally be reflected in the various examinations, which can be seen as expressing formalised sets of objectives in the curriculum. Yet several writers expressed dissatisfaction with the role of examinations in French (II 1), and dissatisfaction appears

if anything to be even stronger among English specialists. For many, the values upheld by the examining boards appear not only irrelevant, but even in actual opposition to the aims of any sort of "personal growth" view of English. Dissatisfaction is not new - the Newson Report (1963) expressed "considerable disquiet at the possible impact of examinations on English teaching", while the 1965 Schools Council Paper was equally concerned that "innappropriate external examinations can too easily frustrate all that is best in the teaching of English." Dixon (1972) points out the difficulties for pupils who must accept exam. cramming in one English lesson, and yet are expected to become "purposeful and independent judges" in the next. Nor does there seem to be much hope for the future. In 1978, Knight, discussing the values implied in examiners' reports, could still conclude -

Divergencies between good English teaching in the schools and the demands of the examiners, not to speak of those between a living English prose and the conventions candidates must observe will continue to result in the expressions of amazement, irritation and fatigue characteristic of these reports.

There is thus an obvious discrepancy between the aims of many teachers and the values implied by the English language examinations; the "O" and "A" level literature examinations also appear unsuccessful even in the terms of a "cultural heritage" model. Jervis (1974) describes a

survey by Yarlot and Harpin of the opinions of 800 "O" level and 200 "A" level literature candidates in Nottingham. Only one boy in every eleven and one girl in seven said they ever wanted to read any more poetry, and in spite of the children's strong preference for modern literature, the choices of almost all the boards remain "antiquated." Jarvis points out that 25 years is more than a lifetime to these pupils, so that writers like D.H. Lawrence are not really "modern" for them at all. The advent of the CSE, particularly mode three, with its emphasis on course work and avoidance of "set books" seemed very promising at one time, but it is still largely seen as "second best" by most pupils, teachers and employers, and not really a serious alternative for able children. As in the case of French, the examining boards therefore do not seem to reflect the changing emphasis in the objectives felt by many English teachers to be important.

(b) Content

It is not surprising that the same basic issues re-appear in an examination of the content of the English curriculum, especially if the distinction between objectives, content and methods is in practice less sharp than Hirst suggests. Jones (1974) summarises a basic change in the nature and content of English books in the title of his article -

"From courses to resources." Similarly, Merrick (1974)

describes how -

the last decade has produced some striking developments in English course books. Up until 1963, it seemed to me, the authors of English course books were people who knew how and were concerned with teaching how ... They also knew what we needed to know in order to be literate, articulated and cultured.

The changes in objectives already discussed were thus obviously reflected in changes in the content of course books which became viewed more as flexible resources to be adapted to the needs and abilities of the pupils rather than vehicles to convey a fixed body of knowledge.

The issues involved have probably been nowhere more clearly demonstrated than in the controversial area of "creative writing." On the one hand is the view, typified by various Black Paper writers, that creative writing has been largely responsible for a decline in standards. The Bullock Committee in fact found no evidence for such a decline, but Froome's note of dissent insisted on "the dilution of English teaching and the reaction against spelling and grammar", putting much of the blame firmly on the rise of creative writing and the attitudes that it implies. Froome goes on to describe -

Schools where in the desire to foster creativity, it is held that children will develop the power to use

language simply by being encouraged to speak and write, and that any critical intervention will stem the flow it is believed that the child's spontaneous effort is sacrosanct and to ask him to improve it is to stifle his creativity.

A good deal has been written on the subject, much of it in the form of heated argument, but creative writing does have its own values, not necessarily incompatible with the traditional emphasis on correctness. Its basic concern is to help children to perceive more accurately and feel more acutely through the expression of perceptions and emotions; language is seen as an important means of extending self-awareness, and it can be made even more explicit by being acted out in the form of improvised drama or the role-play activities which are closely related to creative writing. Moreover, if the child is expected to act out his feelings or commit them to paper in an honest way, an important factor becomes the "sense of audience" described by Thomas (1972) and in the Schools Council Writing across the curriculum project (1973). Sometimes, the child may not have an audience in mind, and may be writing simply "for himself"; at other times he may be writing for "the teacher as examiner." But true creative writing does demand that he should be able to trust the teacher without putting on some special "persona" for writing, in a way that is obviously not necessary for doing spelling tests or comprehension exercises.

The work produced by children is now probably regarded more sympathetically - as valid in its own right, and some of it even appears in print! Several of the popular anthologies (eg Things Being Various, Impact) include children's work among the accepted literary names. The use of reading materials is also strongly related to the emphasis on "personal growth"; one of the earliest Schools Council working papers - English: a programme for research and development in English Teaching (1965), stressed that children should be encouraged to respond creatively to the expressions of other people, as well as to express themselves creatively -

It is important to their development as people that they should, by means of books and drama, extend their imaginative understanding of human actions and motives beyond their immediate circle, and increase the range and quality of their aesthetic responses.

The idea of extending self-awareness is thus at the centre of the "progressive" model of English teaching, but there has also been a growing recognition of the importance of language in other school subjects. The slogan "every teacher a teacher of English" is hardly a new one - Whitcombe (1973) says it can be traced back to the twenties, if not before - but the Bullock report again urged all teachers to examine the language demands of their subjects. Messenger (1974), for example, feels that many recently-developed curriculum schemes in a variety of subject areas demand too much linguistic

competence; Stratta (1972) considers that the "transactional writing" expected in science subjects should be delayed until the child has mastered "expressive writing" - the ability to describe his own experiences. Conversely, if English departments appear to be abandoning their traditional role of maintaining standards of language use, it is important that they should explain what they are trying to do, to other teachers: There is a danger that "changing values" might be seen as "avoiding responsibility."

Another area of change in the content of English curriculum has been a much greater emphasis on oral work - both speaking and listening, though there is a danger that the latter may be sometimes neglected. Listening is a deceptively difficult skill, even for apparently mature and fluent students - Abercrombie's book The Anatomy of Judgement (1960) suggested that such students learned very little from each other in group discussion, at least initially. More stress on oral work does provide increased opportunity for pupil interaction, in which the teacher may be faced with the extremely difficult task of keeping quiet! (cf the controversy over the role of the teacher as a neutral and non-contributing chairman in the Schools Council's Humanities Project). As with written work, there is less emphasis on "correct" use of language as judged by standards which might well be based on a culture

different from the child's own; again one can see a relationship between this sort of approach and work in other areas, in this case the work of socio-linguists like Bernstein and Labov, as well as a great increase of research into language as it is actually used by teachers and children in the classroom. The Bullock report again summarises current thinking here -

"Correct" language varies in different situations and communities. The aim should be to enlarge his repertoire so he can meet new demands and situations and use standard forms when they are needed teachers should start by accepting the language the child brings to school.

The approach to oral work is thus very much in line with the issues already mentioned in reading and writing. Indeed, this is again reflected in Froome's comment that "one of the causes of the decline in English standards today is the recent drift away from the written to the spoken word."

Finally in the area of "content", there has been a recognition of the importance of other media apart from language, though it is arguable how far they should be the responsibility of English teachers. Certainly, they could hardly subscribe to the view of McLuhan that "print is dead" and can be replaced by the visual media; the complementary nature of verbal and non-verbal communication has already been briefly discussed, but the attitudes of English teachers towards other media do sometimes seem ambiguous. On the one hand there

is the feeling that the teacher is in a sense competing with the powerful mass media, especially television (eg 1965 Schools Council Paper), whose implied values may appear to differ from those that the teacher is offering. To this extent, teachers may encourage a critical examination, rather than a passive acceptance of television, looking for example at the language of commercials in the hope of making children less vulnerable to them. On the other hand is the feeling that television and other media can be used by English teachers to study situations or ideas that might otherwise be unavailable to their pupils either for practical reasons or because they involve difficult abstract concepts. Television, like film, has a distinct advantage in being a medium "which has attained the level of an art form but which is also an accepted pattern of leisure, accepted (unlike reading) by the masses." (Simmons, 1967). A programme might therefore be used to illustrate part of a theme also being covered in more conventional ways, or might even be studied in its own right as a form of communication - a proposal strongly made by the Newson Report. For example, Longmans published a book of four "Z Cars" scripts for use with 13-15 year olds, pointing out that the children could consider ideas like responsibility and justice through the concrete examples, as well as becoming more aware of the actual use of film techniques like rapid cutting from one shot to another in "action" sequences. Such

developments can be seen in relation to the idea of "iconic education" mentioned in the first chapter, and the complementary values of verbal and visual forms of communication mentioned in the second chapter.

(c) Methods

Such use of other media clearly brings the discussion into the realms of "methods", and is in fact a central part of a controversial issue in this area - the "thematic" approach. The advent of comprehensive education has already been shown to have serious implications for the modern languages curriculum, and one of its major effects on methods of English teaching has probably been its contribution to the growth of the thematic approach. (Jones, 1974).

Froome, in his "note of dissent" to the Bullock Report, points out that a first year mixed ability class could have reading ages ranging from 7-14¹ and "an accompanying wide divergence in maturity of reading interest and taste", offering

¹The reading age data for the children in the experimental studies does seem to support this. According to the Schonell "B" scores, the reading ages of the 200 11-year-olds ranged from 7½ years and below to 13½ years and above.

this as evidence in favour of streaming. Faced with such a class, the English teacher will need materials that are highly flexible, to say the least, and will have to use a whole variety of approaches. The traditional English curriculum was organised on a "timetable" basis - a comprehension exercise every Monday, an essay on Thursdays and the next chapter from the "reader" on Friday afternoons. In contrast, an H.M.I. report (1971) based on observation of 140 English lessons with first and second year pupils, found that 40% of them were "organic" - lessons involving several different tasks and experiences. The organisation of such lessons often stems from a central theme which is examined in a number of different ways, though this idea is not a particularly novel one. Marland (1968), for example, advocated the "flexible planning needed without the chapter by chapter sequence of a course book as a standby", and concluding that "the teacher of English is an anthologist by nature."

Whole books may be organised on a thematic approach (eg the Impact series), but there has also been a major swing away from the use of textbooks alone. Abbs (1970), for example, produced a series of "Broadsheets" on themes ranging from death to riddles, with a teachers' book of "Extensions" suggesting further developments. He points out that -

the trouble with textbooks, designed to be used

frequently and over a long period of time, is that once the books become known, they become familiar, and as soon as they become familiar they are in danger of becoming dead. This doesn't matter if the books are concerned with learning drills, but it matters greatly if they are concerned with genuinely creative work.

There is thus perhaps a link between the method of presentation and the underlying aims of the "personal growth" model discussed previously. Similar ideas lie behind a number of other commercial products using tapes and slides as well as the printed word, and such materials are seen as flexible learning resources to be adapted for use by the teacher, rather than a set exercise to be completed as an end in itself - a distinction between methods and objectives of which Hirst himself might have approved.

There are dangers in the thematic approach however. Although it does appear more suitable for the mixed ability classes of comprehensive schools, hopefully allowing each child to work at his own level and rate, and involving a whole range of activities and media other than the printed word, the Bullock report saw a danger in that "literature is selected solely because it matches the theme, and its complexities are ignored. There is a tendency to use extracts, and pupils have minimal experience of whole books." Fears about superficial treatment of literature in the thematic approach were perhaps best summed up by the (hopefully apocryphal) story of

the distraught English teacher optimistically offering King Lear as part of a theme on "The Family." In spite of this, the thematic approach embodies many of the current ideas in English teaching which are involved in the general swing away from the traditional emphasis on a body of knowledge to be acquired towards the more flexible approach implied in a "personal growth" model.

(d) Conclusions

Such a brief review as this has obviously had to ignore many areas of interest, ranging from remedial English to the teaching of English as a foreign language, but enough has been said to illustrate the major issues in current thinking about the objectives, content and methods of English teaching.

As with many other subjects, the roots of the controversies lie in the basic "progressive-traditional" debate, represented in English by extreme forms of the "personal growth" and "skills" models respectively. The social importance of language has always been recognised, but notions about its social functions appear to have changed. In the most general terms, there seems to have been a swing away from language viewed in terms of products, with skills that can be directly taught and validated by reference to external standards, to a view of

language in terms of processes and a concern with its role in the cognitive and emotional development of the individual. This change appears in aspects of the objectives, methods and content of the curriculum - dissatisfaction with examinations, for example, can be seen as a result of their unavoidable emphasis on products for assessment and their corresponding inability to satisfy the objectives of the "personal growth" model. In turn, this change in emphasis is perhaps related not only to changes in education generally, but also to the developments in psychology described in the first chapter - a return of interest in mental processes themselves, and not just their external products in the form of observable behaviour.

In spite of similar trends in other areas, however, English as a subject retains certain unique features. The "New English", if there is such a thing, cannot exist in the same way as the "New Maths." or "New Science" embodied in the Nuffield Models. The Schools Council Working Paper of 1965 emphasised this point -

The Nuffield model of organised development assumes the existence of a broad consensus on aims and methods capable of being expressed in materials such as teachers' guides, equipment and audio-visual aids. In English, there is certainly a growing consensus of aims, but it is intrinsic to the subject that methods and materials must always be highly personal.

This personal quality is central. Chapter two suggested that the French teacher, as the possessor of a body of knowledge which initially his pupils simply do not have, tends to be seen as an un-questioned authority on his subject. In contrast, children come to school already able to meet most of the linguistic requirements of the world outside the classroom, with a language which has been shaped by their cultural backgrounds, and which may be very different from that of the traditional grammar books. With such a starting point, there is the danger that English teachers may become what Squire (1972) calls "vague and ill-defined gurus with no content of their own, but a kind of mystic process as their domain", in an attempt to avoid the other extreme of an insistence on imposed standards of correctness; as Marland (1968) commented, "the generalities of English teaching are easy to articulate but difficult to make operational", and the English teacher is now deprived of the firm structure of the "traditional" English course book. Similarly, there is a danger that any form of "traditional" work will be seen as unacceptable, while activities like group work and improvised drama will be seen as "good things", automatically valuable in their own right, rather than as means to an end. If compromise is possible, it will probably lie in a flexible approach, where the teacher feels free to use a variety of materials, adapting them to the individual needs, and perhaps

there is some truth in Jones' (1974) comment that, "if nothing else, the comprehensive schools have lifted the shackles from the English teacher and enabled him to be the master again in his own classroom."

(e) Comprehension Work in English

In a very general sense, "comprehension" can be seen as a central part of the research described in the first chapter, and cognitive psychologists have for some time emphasized the concept of an active learner, testing hypotheses and constantly trying to relate new information to existing schemas. Anglin, for example, in his preface to the collection of Bruner's papers in the book Beyond the Information Given (1974), sees this sort of theme running through Bruner's research right from his early studies of perception to his more recent work in education, and describes Bruner's "persistent concern with the process of knowing, with what man does with information and, in particular, how he goes beyond it to achieve insight, understanding, and competence." In this wide sense, the study of comprehension has become an accepted research area, and psychologists have developed a number of hypotheses, models and experimental techniques to study it, some of which will be discussed in the final section of this chapter. Moreover, much of this work has been done with children as subjects, and the authors frequently claim important implications for the classroom.

But if one turns to the actual status of comprehension in the English curriculum, a very different picture emerges. Seen against the background of the changes in the objectives, content and methods of English teaching that have been described so far, comprehension work appears as a highly traditional activity, based on the "skills" model, in which an isolated passage is presented and a standard exercise, with right or wrong answers, is carried out. Farnes (1973), for example, regards the use of a series of questions at the end of a passage as a useless and artificial exercise which gives the child no real purpose for reading. Smith (1971) considers that the task encourages children to learn only the superficial knack of copying out the appropriate sentence from the passage, and D'Arcy (1973) points out that the questions often involve only the understanding of isolated details or even individual words, ignoring aspects like the overall tone of the passage or the possibility of any emotional response. As usual, the Bullock Report summed up much current thinking in the area, emphasizing that traditional comprehension exercises involve only narrow, superficial skills applied to a restricted context with little real purpose. Even a very brief glance at recent issues of journals like English in Education or Use of English suggests that comprehension is almost a taboo subject and has thankfully disappeared entirely from the activities of any self-respecting English teacher.

If this is the case, then psychologists would appear to be studying and making recommendations about an outdated task which is regarded as irrelevant by English specialists, and it is really this disparity that forms the basis for the second area of study here.

Part of the problem lies in the distinction between comprehension as a general and vital cognitive process, and comprehension as an artificial exercise involving only superficial processing and no real understanding at all. The traditional view of the comprehension exercise was that it was largely a "passive", or at most a "receptive" skill (as opposed to the "productive" skills of speaking and writing) and such a view was implied even in the influential Schools Council Working Paper of 1965. However, there has been a growing realization among English specialists, probably related to developments in psychology and psycholinguistics, that the comprehension of spoken or written material in fact involves a great deal of active processing from the learner. In reading a passage, for example, a good reader does not simply interpret each symbol on the page in turn; he makes predictions based on samples of the text, involving both the wider context of what he is reading, and existing knowledge not derived from the text at all. Such a view underlies Goodman's (1967) description of reading as a "psycholinguistic guessing game" and is implied

in Bruner's famous phrase about going "beyond the information given."

This sort of approach will be discussed more fully in the final part of the chapter, on the psychology of comprehension, but its relevance for English teaching is slowly becoming recognised. The Bullock report was suggesting the same sort of idea with its proposal that

the aim, with comprehension work, is not simply to provide children with hours of practice in picking up the "right" answer, but to encourage the pupil to become "an active interrogator of the text."

In more detail, in one of the very few articles in English in Education actually concerned with comprehension, Wells (1976) concludes that -

At all stages of development, comprehending is an active process of matching expectations derived from various non-linguistic sources with the linguistic cues provided by the utterance itself in order to construct the message judged to be intended by the sender.

Viewed in this light, comprehension can be seen, not as an isolated and artificial exercise, but as an important general skill, or group of skills, with the added implication that such skills can be both taught and tested. As for teaching, D'Arcy (1973), describes how this "skills model" is widely accepted in the United States, though in Britain the idea of building specific comprehension skills using

materials like the American S.R.A. Reading Laboratory is probably more acceptable to remedial departments than English departments in secondary schools. (The 1971 H.M.I. survey in fact commented on exactly this point). As for testing, Flower (1966) describes how the use of comprehension passages began as a form of test. But then -

teachers start to narrow their general work in comprehension to meet the specific requirements of the comprehension test in an examination. Text-books of "comprehension" are then written based upon the syllabus of this or that examination, and a new section of the English Syllabus is born . . . What started out as a test has become an end in itself.

When the emphasis comes to lie on the test itself, and on the assessment of the pupil's products, in terms of right or wrong answers, and children are encouraged to become what Holt (1965) calls "producers" rather than "thinkers", and there is little concern with developing the general processes of comprehension, then the narrower aims of the examinations diverge from those of many English teachers, as has been already suggested.

In this general sense, the process of comprehension involves many different techniques, and works on several levels, so that it becomes no longer the responsibility of the English teacher alone.

Raitt (1976) proposes that -

by the time that pupils are studying recognizable subject areas, there are specific language techniques that they can acquire to improve the efficiency of their learning, and all subject teachers should help them to develop these. In reading, pupils should acquire a variety of comprehension skills: literal, inferential, and evaluative, and should do so across the various subjects of the curriculum.

Similarly, Messenger (1974) and Davies (1977) agree that recognition of the importance of comprehension in all school subjects implies that it is too important an activity to be left to a weekly ritual exercise in English lessons. In conclusion, the apparent opposition of English specialists to "comprehension" as part of their curriculum can thus perhaps be reconciled with the idea of comprehension as a general cognitive process, vital in the development of the individual and in a whole variety of school tasks, extending far beyond the artificial and limited demands of the traditional comprehension exercise.

Britton (1974) asked himself the question "What is English?", decided that it was impossible to answer, and suggested instead that a more useful question would be "What goes on in English lessons?" An H.M.I. survey had attempted to find out more precisely this in 1971, by observing 140

lessons in 12 schools; however, they had to admit that "the purpose of our visits was made known in advance and therefore it can be assumed that lessons reflected what the teachers thought we ought to see." There may be a great difference between "What goes on in English lessons" and "What goes on in English lessons observed by Her Majesty's Inspectorate", so that, with regard to comprehension work, their finding that about a third of the lessons they observed involved some form of reading or listening comprehension may not present a true picture. Mason (1978) points out that we still do not really know enough about the actual language demands made in school, and suggests that if we tried to find out -

One might discover that, in spite of a "broad" syllabus, they are still doing comprehension exercise after comprehension exercise in their English lesson, and, leaving there, doing more comprehension in history.

The first part of the survey questionnaire suggested that vocabulary learning and testing were still common in French, in spite of the recommendations of French curriculum specialists. The second part of the questionnaire is in many ways a parallel exploration of the actual status of comprehension work in English lessons. Those who write in the English journals have been consistently condemning the

traditional collections of isolated comprehension exercises, but are teachers still using them in practice? Are teachers perhaps instead trying to relate comprehension work to other materials that the children would be using anyway - to a "reader" for example? What help, if any, are they giving pupils to improve their comprehension skills? Similarly, the psychological research on comprehension described in the final section of this chapter often claims to have educational implications, but how far are its materials and methods of presentation and testing really similar to those actually used by teachers? The second part of the questionnaire attempted to suggest answers to at least some of these questions.

IV 3 . THE SURVEY - PART TWO

The development of the pilot version, the sampling and administration procedures, and an attempt to establish the reliability of the questionnaire have already been discussed, the only difference being that the whole sample of 155 children from 39 schools filled in the second part, on comprehension work in English lessons. As before, the questions appear below, along with any additional explanations given at the time, the proportion of children giving each

alternative answer, and a short comment on the responses. The whole questionnaire as actually presented appears in Appendix A.

- 12) In English lessons, did you ever use a book with comprehension passages and questions in it?

(A short explanation was added here - "comprehension passages" were defined for the children as "a short story or part of a story, with questions about it." Although strictly speaking the passage need not be "a story", it was felt that this would convey the idea, and E did try to make sure that the children understood exactly what was meant. In fact, they seemed only too familiar with the exercise, as their responses below show.

often (say every few weeks or more often)	81%
Sometimes (say once or twice a term)	9%
Rarely (say once or twice ever)	8%
Never	3%

As with several of the answers in the first section on French, the result here seems very clear, even taking into account possible sample bias and inaccurate responses. The vast majority of the children in the sample reported that books of comprehension exercises were frequently used. Again, as

with French, there seems to be some difference between the recommendations of those whose views are published in the journals, and what is probably the common practice in many classrooms, and again it seems that an aspect of the curriculum that many would label "traditional" is still widespread.

- 13) Did you ever have an English lesson using a "reader" -
a book which you read over a long period of time?

(answers as question 14)

Often	55%
Sometimes	24%
Rarely	7%
Never	15%

The answers here suggest that "readers" are also used quite often, though apparently slightly less frequently than comprehension books. How they are used is of course a different question, and possibilities range from the "traditional Friday afternoon" lesson, unrelated to other English lessons, to the use of a reader as a central part of a thematic approach. (cf the first part of this chapter). The next question therefore looked at a particularly relevant possible use of the "reader" as a basis for comprehension work which could be seen as rather less artificial and isolated than the use of books of comprehension exercises.

14) If so (ie if a "reader" was used), did the teacher ever set comprehension questions on this book or part of it?

often	19%
Sometimes	9%
Rarely	21%
Never	51%

A slight problem here, which had not been noted in the pilot version, was that 18 of the 22 children who had never used a "reader" also reported that they never had questions on it, instead of leaving this question out. Slightly different results are obtained if their responses are excluded, though the general trend is very much the same. Taking children who did report that they used a reader, the following proportions of them were given comprehension questions on it -

Often	22%
Sometimes	10%
Rarely	24%
Never	44%

However, it still seems that questions set by the teacher on a "reader" are less common than ready-made comprehension exercises. In particular, almost half (40 out of 83) of those who "often" used a reader never had questions on it. It would thus appear that many teachers still regard comprehension as a set activity best done by using specially selected passages in isolation, collected together in a course book.

15) If you ever did any sort of comprehension questions, were you usually allowed to look at the story or passage while answering? (E emphasised that the question referred to any sort of comprehension work).

Yes 83%

No 17%

These figures in fact refer to the whole sample, since only 4 children had reported that they never used books of comprehension exercises, and all 4 of them reported doing comprehension questions set by the teacher on the class "reader." (in other words, every single child in the sample did some form of comprehension work).

The trend is obvious again, and the replies raise an important issue to be examined in the experiment in the next chapter. The vast majority of the published experimental research on comprehension has not allowed subjects to look back at the passage when answering questions on it. The claims of such experiments to have "educational implications" will therefore be seriously affected if it can be shown that presence or absence of the text is an important variable. This issue will be discussed in more detail, particularly with regard to research on the effects of adjunct questions on passage comprehension. Some later items in the questionnaire were also included to look at this point in more detail.

16) Did you ever do "multiple choice" comprehension questions on a story - where you have to pick the right answer from several answers? (It was emphasised that this question referred to any kind of comprehension work, and care was taken that the children understood the idea of "multiple choice" by reiterating that, after the question, three or four possible answers are given, and the pupil has to pick out the one he thinks is right. S.R.A. was given as an example).

Often	17%
Sometimes	27%
Rarely	30%
Never	27%

Like the previous question, this item was included in an attempt to see how "real" the experimental research conditions are. In the research, comprehension questions are almost always in multiple choice form, and while this has obvious advantages for scoring, it would appear that its use is rather less common in the classroom. It is also possible that the format of the questions affects pupils' strategies and scores, and this too will be discussed in the next section. (of Levin, Ghatala, Guttman, Subkoviak and Bender, 1978).

(258)

17) Did you find comprehension work generally

Very easy 14%

Quite easy 74%

Quite difficult 11%

Very difficult 1%

This question was parallel to question 10 on French Vocabulary, and again most subjects opted for the "Quite easy" response. It might be interesting to compare responses to these two questions, and table 29 shows the number of subjects choosing each answer on questions 10 and 17.

		Comprehension Work (Question 17)			
		Very easy	Quite easy	Quite difficult	Very difficult
Vocabulary learning (Question 10)	Very easy	1	4	2	0
	Quite easy	13	64	2	0
	Quite difficult	6	20	9	0
	Very difficult	0	0	0	1

Table 29 Number of responses in each category, questions 10 and 17.

Categories were collapsed as before because of the very low frequencies in some cells, giving a value for Chi Square of 9.2 ($df = 1$, $p > .01$), and suggesting that subjects who found one task generally easy were likely to find the other task also generally easy. However, sample or response bias could well have affected these two questions in particular, and it may not really be possible to conclude that most 11-13 year olds find French vocabulary and English comprehension work quite easy.

18) Did the teacher ever give you advice about how to do comprehension exercises - eg what to do while reading the passage? (E gave examples of "Trying to picture the scene in your mind as you read it" and "Looking at all the questions before you start to read.")

Often	7%
Sometimes	41%
Rarely	29%
Never	24%

Compared with question 11, which had asked about advice being given on how to remember vocabulary, the results here perhaps suggest that a little more help is given. Rather more interesting however, is the comparison between this question and the previous one: is there any apparent

relationship between reported ease of doing comprehension exercises and help given by the teacher? Table 30 shows the breakdown of answers to these two questions.

		Advice on comprehension (question 18)			
		Often	Sometimes	Rarely	Never
Comprehension work (Question 17)	Very easy	3	2	7	9
	Quite easy	5	52	34	20
	Quite difficult	2	7	3	5
	Very difficult	0	1	1	0

Table 30 Number of responses in each category, questions 17 and 18.

Using the same collapsed categories, a non-significant value for Chi square was obtained. ($\chi^2 = .05$, $df = 1$, $p > .05$). As with the French vocabulary questions, there would appear to be no relationship here between the reported ease of doing comprehension work and the reported amount of advice given by the teacher.

- 19) Did the teacher ever test you on how well you remembered a story or part of it, without looking at the book?

Often	13%
Sometimes	25%
Rarely	41%
Never	21%

As previously indicated with question 15, the experimental research (which frequently uses recall as a test of passage comprehension) would appear to be at variance with normal classroom practice, according to the answers here.

- 20) If you had to remember a story, did you generally find this -

Very easy	8%
Quite easy	62%
Quite difficult	30%
Very difficult	0%

In comparison with question 17, these results suggest that, as one might expect, the children generally found answering questions on a story more difficult if they were not allowed to look at the story. Before this result is dismissed as trivial however, an experiment by Tuinman and Farr (1972) should be noted, which suggested that having the

passage there or not made no difference to children's comprehension scores, even on long passages. This point will be discussed in more detail later, but it is worth noting that experimental research again appears to disagree with the children's answers, not to mention common sense!

21) Did the teacher ever give you advice about how to remember a story - what to do while reading it? (E gave similar examples to question 18).

Often	3%
Sometimes	10%
Rarely	15%
Never	73%

The answers to this final question suggest that even less help is felt necessary for remembering a story than for answering comprehension questions on it with the text there, though the scores in the "never" category could obviously have been boosted by the fact that teachers appear to give recall tests less often than comprehension work (Cf questions 12 and 19). At any rate, tests of recall of a passage or story appear to be much less important than the use of questions to test comprehension of a passage to which the child can refer.

Finally, the children's spontaneous comments on their English lessons need to be briefly analysed. As with French, most of them could be quite easily divided into the two basic categories of "for" and "against." 31 children commented favourably about their English lessons, and 49 unfavourably. Again, many of the comments were really about the teacher - 16 of the favourable and 20 of the unfavourable clearly fell into this category. Other criticisms included the opinion that English was "boring" - as with French the single most common complaint about the actual lessons (18 comments). On the other hand, 13 children actually said they found English "interesting." Another apparent difference between comments on the two subjects was that several children appeared to find English easy: the comment "English is good. He explains and it is not difficult. He tells you what to do" was typical of several. Some perhaps even found English undemanding -

"All we ever do is read books, only very rarely do we do grammar."

"We don't really learn a lot from English." (This same child added "We learn too much at once in French.") Only six children described English as "difficult."

Analysis of the children's comments is obviously rather subjective, and ideally a panel of judges should be used;

however, some very tentative conclusions can be drawn.

Firstly, a recent report in the Educational Research News (N.F.E.R., 1978), describing the Northern Ireland Schools Curriculum Project, concluded that "at early secondary school level, pupils barely distinguish between a school subject and the teacher who teaches it", and this conclusion tends to be supported by the free comments of the children in this sample. Roughly half the comments they made altogether were about their teachers, and the identification between subjects and teachers appears very strong.

Secondly, there do also appear to be several differences between the children's perceptions of the two subjects. An attempt has been made to summarise the children's comments about the two subjects in table 31, though the figures given are obviously only a very rough indication.

	French	English
All Favourable comments	15	31
Liked teacher	12	16
Subject "interesting"	2	13
All Unfavourable comments	79	49
Disliked teacher	38	20
Subject "boring"	17	18
Subject "difficult"	20	6

Table 31 Number of spontaneous comments on French and English lessons.

Taking simply the favourable and unfavourable comments for the two subjects, a Chi Square test gave a value of $\chi^2 = 10.4$ ($df = 1$, $p < .01$), indicating that the children in the sample made significantly more favourable comments about English than about French. However, a similar test comparing comments about the teachers of the two subjects gave a non-significant result. ($\chi^2 = 3.1$, $df = 1$, $p > .05$). It had originally been hoped to compare the results here with those of Duckworth and Entwistle, (1974) whose 12 year-old subjects rated English, French and other subjects

for Interest, Difficulty, Freedom and Social Benefit, but the children's comments in this case suggested that they did not use categories like these spontaneously, and often limited what they wrote to a simple statement of liking or disliking the subject or teacher.

In conclusion, the results of the questionnaire suggest that traditional books of comprehension exercises are still being widely and frequently used, and that teachers prefer them to setting their own comprehension work based on a "reader." However, use of multiple choice questions is less common, and the children said they were usually allowed to look at the passage while answering, so that on both these points, normal classroom practice appears to differ from experimental studies of children's comprehension, which frequently use a multiple-choice recall test. Finally, some advice does appear to be given, though not often, and over half the children said that the teacher rarely or never gave them advice on how to do comprehension questions. Like the learning of French vocabulary, the task thus appears to be widely used, in a rather traditional form, in spite of the recommendations of the curriculum specialists, and often with very little help on how to go

about it. Such help might be offered by experimental research into the psychology of comprehension, which claims to have isolated a number of important variables involved, and to have developed several techniques which should improve children's comprehension of what they read. A particularly relevant approach seems to be that of various levels of processing, and it is to this research that discussion now turns.

IV 4 PSYCHOLOGICAL STUDIES OF COMPREHENSION

(a) Comprehension and levels of processing

In the psychological literature, the best-known and most influential statement of a "levels-of-processing" approach has probably been that of Craik and Lockhart (1972), who proposed it as an alternative framework for memory research, as opposed to the numerous multistore theories of memory. They suggested that processing involves a hierarchy of levels -

This conception of a series or hierarchy of processing stages is often referred to as "depth of processing" where greater "depth" implies a greater degree of semantic or cognitive analysis. After the stimulus has been recognised, it may undergo further processing

by enrichment or elaboration. For example, after a word is recognised, it may trigger associations, images or stories on the basis of the subject's past experience with the word Analysis proceeds through a series of sensory stages to levels associated with matching or pattern recognition and finally to semantic-associative stages of stimulus enrichment.

There are certain problems or weaknesses in the approach - Eysenck (1978), for example, points out the dangers of circularity, and considers that "depth" is a vague concept which ignores factors like stimulus discriminability and response bias, while Postman (1975) suggests that if the levels are truly on a continuum, then the identification of stages is bound to be arbitrary. In spite of these criticisms, however, the "levels of processing" approach would appear to be very relevant to the study of comprehension, especially if one accepts Lockhart and Craik's (1978) comment that memory is just "one aspect of a general theory of cognitive processing" and the view that comprehension is itself bound up closely with other cognitive activities (eg Farnes, 1973; Smith, 1975, and cf Brewer, 1974) to the extent that recall is itself often used to test comprehension.

Certainly, the idea that comprehension is not unitary, but involves several discriminable levels of processing,

underlies several pieces of research into comprehension, particularly the comprehension of sentences. (eg Glucksberg, Trabasso and Wald, 1973; Mistler-Lachman, 1972, 1974), though there may be some differences in emphasizing the levels of information embodied in the stimulus itself (Friedman and Bourne, 1976) as opposed to levels of processing largely under the control of the subject. (Montague, 1972).

At the "shallowest" level, reading comprehension obviously requires considerable visual skills, and the perceptual aspects of reading have been the concern of a number of psychologists. We know, for example, that the eyes do not move smoothly over lines of print, but proceed in a series of fixations alternating with "saccadic jerks." The conventional view is that processing can take place only when the eyes are still (eg Wright, 1972), though it is possible that the "iconic image" could be used to prolong the stimulus perceived at each fixation (eg Haber, 1971), or even that good readers could actually perceive material as their eyes move rapidly over it during each saccadic jerk (Edfeldt, 1975). However, it seems that poor readers can actually scan lines of print in a "visual search" task just as quickly as good readers can (eg Katz and Wicklund, 1972), and that they do not differ in average fixation time (eg Wright, 1972). On the other hand, good readers require

fewer fixations per line and appear better able to make use of peripheral cues (Marcel, 1974), particularly at the right-hand edge of the visual field, thus providing valuable clues about the text which follows. However, concentration on the perceptual attributes of the text results in the least effective way of reading, involving what Smith (1975) and others call "tunnel vision", where the struggling reader concentrates on trying to decode the visual symbols of tiny sections of the text, one at a time.

Successful reading comprehension clearly requires far more than even very skilled visual decoding, and although such skills may be a vital first stage, it has already been suggested that, in the most general terms, comprehension involves the attempt to relate stimulus information to what is already known. The good reader uses the whole context of what he is reading, combined with existing knowledge not derived from the text at all, to sample, infer and predict, so that a wide variety of perceptual and non-perceptual linguistic and non-linguistic skills are involved. It is obviously not enough to be able to decode the individual words in a sentence, or even to understand the syntactic relations between them, and for comprehension to occur, the reader must be able to draw on semantic contexts beyond

those explicitly stated in the text (cf the often-quoted experiments by Bransford and Johnson, 1972a). Wells (1976) sums up -

Comprehension takes place when the receiver of a message is able to make a match between the cues provided by its linguistic form, and the expectations set up by his interpretation of its situational context in the light of his current purpose and of his past experience. The decoding of the linguistic input can never be sufficient in itself for, although this can lead to an abstract specification of the sort of meaning that the sender intends, the way in which this specification is to be applied on a particular occasion can only come from the information that the receiver brings from his currently activated knowledge of the physical and social world in which such messages occur.

Seen in these terms, the psychological study of comprehension must take account not only of the nature of the message and how it is presented, but also of the complexities of individual differences and the nature of knowledge existing outside the message. Faced with this daunting task, many researchers have accepted that comprehension cannot be viewed as a unitary concept, and have proposed different "levels" or "depths" of comprehension arranged in a hierarchy. While offering flexibility, such an approach runs the risks of making arbitrary divisions (cf Postman, above) with the result that different researchers have decided on different divisions and labels, though there is general agreement about the overall pattern. At one

extreme is the idea that reading comprehension is responding orally to printed symbols (Gattegno, 1970), while at the other is the notion of the reader changing his whole view of life as a result of what he reads (Clymer, 1972), and there is thus at least a partial similarity here between these two views and the "skills" and "personal growth" models of English teaching discussed earlier.

In between these extremes, different writers have offered different classifications, some tending towards one end and some the other, while still others try to cover the whole range. Thus Anderson's three levels of encoding - perceptual, phonological and semantic (Anderson 1972), and the three offered by Forster and Olbrei (1973) - lexical, syntactic and semantic, can both be seen as emphasizing the more measurable skills of comprehension. A rather different emphasis is placed by Strang (1972) or the writers of the Bullock Report (1975), whose classification systems are more in sympathy with the "personal growth" end of the spectrum. Strang uses Dale's often-quoted phrases "reading the lines", "reading between the lines", and "reading beyond the lines" to distinguish between three levels rather similar to the Bullock Report's literal, inferential and evaluative levels of comprehension.

The result is that chosen classifications obviously reflect the interests of the writer so that, for example, while Anderson's classification does not really distinguish between inferential and evaluative comprehension, which would both be seen as "semantic encoding", the Bullock Report in turn is not concerned with the distinction between perceptual and phonological encoding, both of which might be classified as "literal" comprehension, (or hardly regarded as comprehension at all). Finally the few attempts to cover the whole spectrum must either involve rather vague and all-embracing categories, or at least a larger number of them; Cleland (1965), for example, offered six, ranging from "perception" to "application."

One of the most frequently-quoted classification systems is, surprisingly, an unpublished one by Barrett, the usual published source being given as Clymer's (1972) paper. Barrett distinguished five levels - literal, reorganisation, inferential, evaluation and appreciation, and also suggested that teachers often have two misconceptions about comprehension. The first is to consider comprehension as a single unitary skill and the second, conversely, is to assume that it contains so many skills as to be unmanageable!

If the present discussion seems in danger of suggesting this latter view, it is only fair to point out that there is in fact general agreement about the most basic division of comprehension into two very broad categories which might be called "literal" and "inferential." (Pettit and Cockriel, 1974), the former being limited to information explicitly given in the text, and the latter also involving the reader's own cognitive and affective contributions which allow him to make inferences and value judgements about what he reads. (cf Peel's "in" and "out" concepts, eg Peel, 1974). Such a basic division is apparent in most of the classification schemes already mentioned, and is obviously consistent with the "depth" or "levels" approach to comprehension. In terms of information - processing in general, these two categories would seem to be related to Jensen's level I and level II abilities (eg Jensen, 1969), level I being the capacity to register, store, recognise or recall information, and level II involving elaboration, transformation and manipulation. It is perhaps significant that most of the controversy surrounding Jensen's views has involved their political and educational implications, rather than this basic premise, and even Craik and Lockhart (1972) cannot resist the temptation of talking about "Type I processing",

which they describe as "repetition of analyses which have already been carried out"¹ and "Type II processing, which involves deeper analysis of the stimulus."

Much of the research on comprehension can therefore be related to the "levels of processing" approach, with a basic division between two levels, one involving only the literal comprehension of explicitly-presented information, and the other involving inferences and judgements based the reader's existing knowledge and his active cognitive and affective contributions. Further sub-divisions frequently occur, but they tend to reflect the individual interests of the researcher, and there is consequently rather less general agreement about them. In spite of this, the "levels of processing" view of the nature of comprehension has two important classes of implications, the first concerned with ways of testing comprehension, and the second with the range of factors that will affect it.

¹ cf the suggestion made in chapter III that rote repetition of French vocabulary items involved only shallow processing, with no elaboration. In contrast, mnemonics would seem to involve level II or deeper processing.

(b) Testing Comprehension

It is apparent that, to say the least, the psychological study of comprehension is very complex. Faced with this daunting task, Anderson (1972) comes pragmatically to the rescue -

It is not necessary to know the precise form and the mental organisation of meanings in order to recognise the occurrence and non-occurrence of comprehension. The trick will be to devise techniques for constructing questions that can be answered if a person has semantically encoded a communication but not answered if it has been encoded only perceptually or phonologically.

Asking questions about a passage would certainly seem to be an obvious way of testing comprehension, and the levels-of-processing approach would suggest that several different types of question could be asked to test comprehension at different levels. Again, there is general agreement about a basic division into two types of question, to test literal and inferential comprehension, with further sub-divisions largely reflecting the interests of the writer. Anderson himself offers five - Verbatim , transformed verbatim, paraphrase, transformed paraphrase and concept questions. Given the unlikely sentence "The toper torrefied toque" one could confidently answer the verbatim question "What did the toper torrefy?" or the transformed verbatim

question "By whom was the toque torrified?" on the basis of only a very shallow level of comprehension. The paraphrase question "What did the drunkard roast?" or the transformed paraphrase question "By whom was the hat roasted?" would require an understanding of the meaning of at least the individual words, while the concept question "How was the clothing destroyed?" would involve an understanding of underlying categories. However, Anderson's classification is confined to the shallower levels of comprehension; a taxonomy like Barrett's (Clymer, 1972) would suggest at least three other levels - inferential (eg "What makes you think the man was drunk?"), evaluative (eg "Do you think he could be blamed for what he did?"), and appreciative (eg "What do you think of the writer's choice of vocabulary?").

In practice, however, it seems that children have enough trouble coping even with levels of questions like those described by Anderson. Bormuth, Carr, Manning, and Pearson (1970), who describe levels of questions similar to Anderson's, gave 9-10 year olds examples of each type, and found the following percentages of correct answers (using Anderson's labels for clarity):

(278)

Verbatim	77% correct
Transformed Verbatim	71% correct
Paraphrase	69% correct
Transformed paraphrase	67% correct

Bormuth et al comment that many of the children failed on the basic verbatim questions, let alone those involving syntactic or semantic changes, and their concern is expressed in their interest in "an instructional theory of comprehension." Their figures would also indirectly support the finding of Paris and Upton (1976) that performance on verbatim questions improves with age, but much less dramatically than performance in inference questions, and this finding is again clearly compatible with previous discussion on levels of processing in general.

There also seems to be recognition that, in a way, verbatim questions are not really a test of comprehension at all, and that deeper questions are generally much more useful to the reader. Anderson (1970) points out that students will process text only to the shallowest level necessary to answer questions; given obvious cues like having the correct answer underlined in the text, for example, results in very poor delayed recall, and Anderson in fact

goes on to describe research suggesting that students learned more from a Skinner program if the right answers are not provided, at least in a form that allowed cheating. Similarly, Putnam (1975) claims that only "deep" questions can improve children's comprehension, while Andre and Sola (1976) suggest that even paraphrase questions require some semantic encoding and thus produce better recall than verbatim questions do. An actual training programme was carried out by Belch (1974), who found that practice in answering "high order" questions each day for a month improved reading comprehension scores significantly more than similar practice with "low order" questions, which gave no more improvement than simply reading the passages with no questions at all. This area of research involves the use of questions as a training rather than a testing device, and one particular approach - the use of "adjunct questions" - will be discussed in more detail in the next chapter.

It has been emphasised that comprehension often involves the reader's existing knowledge as well as information he picks up from the text; one implication of this is that it may be possible to answer some questions purely on the basis of prior knowledge, without really having "comprehended" the passage at all, and Peel's

distinction between "in" and "out" questions (eg Peel, 1974) serves to emphasise this point. One attempt to assess how far existing knowledge has an effect in programmed instruction is to work out the "blackout ratio" of a passage (eg Anderson, 1970) - the percentage of words that could be blacked out without reducing the frequency of correct responses. As for comprehension questions, a more straightforward approach is simply to get subjects to answer the questions without having read the passage, to give an indication of the "passage dependency" of the questions. Tuinman (1974) gave only the questions in five widely-used tests to over 9,000 children and found that, with four multiple-choice answers to choose from, giving a chance score of 25%, the children obtained average scores ranging from 32% to 50% on the five tests. On the very popular S.R.A. reading laboratory tests, for example, children who had read the passage had a mean score of 62%, while those who had not read it scored 37% rather than the chance guessing level of 25%, and scores obtained through prior knowledge also increased with age. While it might be argued that the ability to make inferences based on a combination of prior knowledge and passage information is a vital part of comprehension, Carroll (1972) does not entirely agree and holds to the strict view that "an ideal test of comprehension should be valid in the sense that it reflects solely

comprehension and not any other behavioural process such as memory, inference, guessing or the like."

On this basis, Carroll also has reservations about the assessment of comprehension by the "cloze" technique, where the reader has to fill in gaps left in the text. Although it is derived from information theory, particularly the concept of redundancy, the cloze technique can also be related to a levels-of-processing approach. Ammon (1975), suggests that the reader generates language on three levels - phonetic, syntactic and semantic. Syntactic and semantic features in the text provide redundant information which enables the reader to fill in the gaps, particularly if a phonetic cue (usually the first letter of the missing word) is provided to limit the possible answers. The technique is obviously flexible - the experimenter can choose the words he wishes to delete to make the reader focus on desired aspects of the text, or can simply delete, say, every fifth word, giving a "20% cloze" test. Cues or multiple choice answers may be provided, and Ammon (1975) considers the technique a valuable training device.

However, Carroll's reservations (Carroll, 1972) are echoed by several other writers; Neville and Pugh, (1974) for example, consider that the cloze technique cannot be used

with younger children, while Carver (1973) concludes that "the cloze method is relatively insensitive and thereby invalid as a measure of what is understood during reading." (Carver himself prefers the "Chunked test" in which the reader has to recognise which of several "chunks" of text have been altered). It could also be argued that the cloze technique is necessarily limited to the literal levels of comprehension, and cannot be used to examine inferential, evaluative or appreciative comprehension, (see Barrett, above). Moreover, although the technique is quite widely used in experiments, it appears to be quite rare in schools, and even a brief look at some of the comprehension materials in current use suggests that the cloze method, if used at all, is restricted to the occasional exercise. (For example, the S.R.A. reading laboratory materials sometimes have a short exercise in which the child has to fill in gaps in sentences, using words selected from the text he has just read).

Unfortunately, a similar criticism can be levelled at the use of questions by experimenters, who frequently claim quite direct educational implications for their research, for in spite of these claims, the testing of

comprehension in experiments differs from comprehension testing in schools in at least two important ways.

First, the questionnaire data strongly suggested that the use of multiple choice questions, while common in experiments for reasons of control and objectivity, is comparatively rare in schools. (81% of the sample said they often had comprehension exercises, but only 17% often did multiple choice questions, in contrast with 57% who rarely or never did them). In psychological terms, the use of multiple choice items can perhaps be seen as related to the use of recognition or true/false tasks, to the extent that they all involve matching, with a minimum of overt production, and it is for precisely these reasons that several English curriculum specialists reject them. Among the very few articles in the English journals that actually discuss comprehension, the use of multiple choice items does appear of interest. Honeyford, (1973), and Messenger (1974) consider that their use encourages guessing and completely ignores the "creative" aspects of language; they interfere with the child's active response to the text and offer instead only the passive selection of one provided response. (There is thus perhaps an interesting similarity here with "experimenter-provided" versus "subject-generated" mediators discussed in chapter two and it is also possible

to see this sort of view as related to the "personal growth" model of English teaching). In contrast, Davidson (1974) points out that any form of comprehension test interferes with the reading process, and claims that good multiple choice items require considerable active processing without the additional, and possibly unfair demands on the reader to generate new text in the form of his answer, which is not part of "comprehension" anyway.

In spite of the disagreements, the fact remains that a form of question often used in laboratory studies is not the same as the one found most commonly in the classroom. The second important difference involves the use of recall as a test of comprehension. The questionnaire again suggested that this was comparatively rare in schools: 83% of the children said they were usually allowed to refer back to the passage when answering questions, yet this is very rarely allowed in experimental studies. It is true that recall is strongly related to comprehension (eg Carroll, 1972) though there is certainly not a one-to-one correspondence between them, since subjects may remember something without understanding it, or understand something without remembering it (eg Bransford and Johnson, 1972b). Moreover, while good comprehension usually implies successful recall, the deliberate

attempt to remember something, perhaps for a test expected later, may actually interfere with comprehension (Smith, 1975) - a suggestion that can clearly be related to the effects of concentrating on stimulus attributes and relatively shallow processing. (There may therefore be particular dangers in using an expected recall test as a measure of comprehension). Again, ease of comprehension may not be related in a straightforward way to ease of recall: the experiment on bizarreness described in chapter two (Merry and Graham, 1978) suggested that images formed for sentences like "The hen smoked a cigar" were easier to recall than ones for sentences like "The horse ate the hay", and yet, however one assesses comprehension, the latter sentence is surely easier to comprehend.

Without wishing to deny that recall is probably a very useful indicant of comprehension, it would thus appear that it needs to be used with care, though Wilson (1974) actually claims that asking literal comprehension questions without the passage being there is pointless, and urges teachers and researchers to concentrate on the "real" task of comprehension. There appear to have been very few studies comparing children's answers to questions with and without the passage present, and an experiment by Tuinman and Farr (1972) could actually be used to support the

widespread use of recall tests, since they found no significant differences on children's comprehension scores whether they were allowed to look back at the passage or not, using both long and short passages of both high and low passage dependency. However, their "long" passages ranged from 166 to 356 words (as opposed to "short" ones of 42-140 words) and they admit that differences could appear with longer passages. This possibility and the fact that they used only an immediate test suggested that this point could be examined in the experiment to be described in the next chapter.

In conclusion, it is clear, as Carroll (1972) points out, that there is no one satisfactory method of testing comprehension. Success or failure could occur at various levels, and no one method can rule out the effects of all other factors like guessing, inferences, prior knowledge or memory ability. The form of the test is almost bound to affect the subject's strategies, and even different forms of recall test, for example, make very different cognitive demands: Aaronson (1976) suggests that free recall requires little inferential comprehension in comparison with a true/false recognition task, for example.

Similarly, the exact form of a multiple choice test strongly affects subjects' answers - the plausibility of "distractor" items is one factor, and the use of verbatim extracts from the text, rather than paraphrases, as the correct choice item, significantly increases comprehension scores. (Levin, Ghatala, Guttman, Subkoviak and Bender, 1978).

If it is true that the nature of the test used to assess comprehension does have an effect on the strategies adopted by subjects and their apparent success, then the widespread use in experiments of the cloze technique, multiple choice questions, and recall tests implies that results produced in the laboratory may not hold up in the classroom, and this possibility underlies the experiment described in the next chapter.

(c) Major factors affecting comprehension

If the most "shallow" level of comprehension involves simply the accurate perception of words, then perceptual attributes like their length and shape should have some effect, and this does appear to be the case (Wolpert, 1972). At the "lexical" level, properties of individual words like their frequency and meaningfulness will also matter, and

the research described in the first chapter strongly suggests that the imagery values of the words in a passage greatly affect its difficulty. (eg Yuille and Paivio, 1969). However, there are problems in generalizing from the rated values of individual words to their use in context (cf Bransford and McCarrell, 1974), and a well-established body of research, often influenced by the work of Chomsky, has long emphasised the importance of the syntactical relationships between words. (eg Olson, 1977). On the other hand, several researchers claimed that Chomsky's logical rules about language competence might not be the same as the psychological "rules" of language performance (eg Aaronson, 1976), and it soon became clear that in practice, syntactic and semantic constraints interact and that both affect comprehension. For example, while it is true that passive sentences are generally harder to comprehend than active ones, reversibility is also an important factor which depends entirely on semantic considerations. (eg Slobin, 1966). Similarly, good readers are more sensitive than poor readers to both syntactic and semantic cues (Isakson and Miller, 1976). In fact, the priority of semantics has been strongly advanced in the past few years, and a whole new field of "Generative Semantics" has grown up in opposition to the

earlier emphasis on Syntax, while Paivio's contribution includes the idea that syntax is in any case based on a foundation of imagery. (eg Paivio, 1971). Finally at this level come most of the measures of "readability", which use factors like vocabulary and sentence length to assess the difficulty of a passage (eg Wall, 1969), though they do accept the importance of other aids like pictures and increased redundancy which go beyond the properties of individual sentences. (eg Siegel, Lambert and Burkett, 1974).

At its lowest levels, then, comprehension can be strongly affected by the properties of individual words and the relationships between them, though "readability" is alone an inadequate guide to the ease with which material can be comprehended. (eg Kretschmer, 1975). At a deeper level, the reader must be aware of the context provided by the whole passage, which is obviously more than a collection of individual sentences, and passage structure has also been shown to affect comprehension (eg Lee, 1965; Thorndyke, 1975; Bower, 1976).

Yet all these factors, from individual word properties to overall passage structure, are really only concerned with "literal" comprehension. At the deeper levels, comprehension

involves the reader's own contributions, and the relationship between these and the materials themselves has been a topic of recent interest to several psychologists. For example, the idea of "semantic integration" (eg Bransford and Franks, 1971; Paris, Mahoney and Buckhalt, 1974) proposes that subjects who read a passage store "wholistic situational descriptions" which will later lead them to "recognise" sentences not included in the passage, provided they are compatible with the ideas and inferences they have stored, although further research has suggested that such results are partly due to experimental conditions (eg Katz, Atkeson and Lee, 1974), and that some surface properties are also stored. (Postman, 1975). A rather similar idea is that of "surrogate structures" which the reader generates for himself, and which contain the basic ideas of a passage. Pompi and Lachman (1967) describe how -

the general theme or idea of a passage may be thought of as one huge chunk containing much information in the sense that it is the nucleus of a complex matrix of associations. This chunk may be part of what is meant by the "meaning of a passage."

The suggestion that the surrogate structure could be in the form of an image was naturally seized on by Paivio and his associates (eg Yuille and Paivio, 1969), and this would seem reasonable in view of what has already been said about

imagery's value as a relational organiser. If this is the case, then instructions to form images might help those who would not do so spontaneously to create surrogate structures, and thus improve their comprehension of a passage, particularly at a deeper level, and this hypothesis is in fact the basis for one of the experimental conditions described in the next chapter.

Conversely, a speaker or the writer of a passage must be sensitive to the attempts of his audience to relate new information to their various forms of existing knowledge. The "Given-New contract" describes how the form of the message is affected by what the speaker or writer thinks the receiver already knows. (eg Haviland and Clark, 1974; Clark, 1977). Similarly, the idea of "foregrounding" (eg Chafe, 1972; De Villiers, 1974) sets out in detail how the choice of vocabulary and sentence structure is affected by "transitory constraints" involving the limited amount of information that the audience can retain while attending to the message. Ideas like these are beginning to examine the complex interactions that take place between discourse structure and the cognitive contributions of the reader or listener, at the deeper levels of comprehension.

The "Given-New Contract" and "foregrounding" were originally related to spoken rather than written discourse, however, and it is clearly dangerous to equate the two. There are important differences between listening and reading comprehension, and comparative success in one mode rather than another appears to interact with variables like age and ability (eg Swalm, 1971) and socio-economic status (eg Peskin, 1973). Apart from the more obvious factors like the "ephemeral" nature of speech (eg Smith, 1977) and its use of extra-linguistic cues, there are other considerations like the possibility that the art of reading could interfere with the generation of images (eg Brooks, 1968; Sutherland, 1971), and the effects upon comprehension of rate of presentation (eg Riding and Shore, 1974), which is controlled by the experimenter more easily and naturally in listening comprehension tasks. On a more general level, Olson (1977) makes the important point that inherent differences between written and spoken language could be largely responsible for the continuing lack of common ground between the syntactic and semantic approaches.

Comprehension can be affected at all levels by the conditions under which the material is presented. The provision of pictures for example, either along with the text (eg Pressley, 1976) or as "advance organisers" (eg Bransford and Johnson, 1972; Arnold and Brooks, 1976) has

been very effective, (though again studied more with listening rather than reading comprehension), while the provision of various other strategies can be seen as affecting comprehension along the lines already discussed in the first chapter and in the descriptions of Rohwer's elaboration hypothesis, where various levels of prompt can improve, or may possibly interfere with, the spontaneous strategies of the subjects. In fact, the provision of two particular strategies, in the form of imagery instructions and adjunct questions, forms the basis for the experiment to be described in the next chapter.

Finally, since the whole emphasis of the levels-of-processing approach to comprehension involves the reader's imposing meaning on the text as well as drawing information from it, individual differences between readers will be just as important as the nature and presentation of the text, and will interact with both. Younger and less able subjects, for example, are less likely to have requisite pre-knowledge or suitable spontaneous strategies, and will therefore benefit from advance organisers and highly explicit prompts rather more than older and more able subjects will. Kretschmer (1975) has applied Piagetian theory to children's

errors in reading comprehension and, though he does not particularly favour the idea of different levels of understanding, other research suggests that such an approach can be successfully applied to younger children's failure to make inferences about sentences (Paris and Lindauer, 1976) or to select the key units from a passage (Brown and Smiley, 1977).

In the closely-related research area using retarded subjects and poor readers it is clear that, if comprehension is not unitary, then failure to comprehend can be due to a breakdown at a number of levels. Weiner and Cromer (1967) and Cromer (1970) propose four models of reading difficulty: the "defect" poor reader fails to comprehend at even the "shallowest" level because of sensory impairment; the "deficit" poor reader has vocabulary or decoding problems, while the "difference" poor reader can understand individual words but has poor reading habits and tends, for example, to "bark at print" - to read word-by-word without "chunking." The success of a provided strategy will thus depend on the level at which the breakdown occurred (eg Levin, 1973).

Other differences between subjects which have been shown to affect comprehension include socio-economic status

(SES), sex, and anxiety. High SES is generally correlated with superior reading comprehension (eg Cockriel, 1973; Peskin, 1973), and in the more general terms of "advantaged and disadvantaged populations" (Pressley 1977) this sort of research has been extended to comparisons of black and white children of different SES levels (eg Rohwer and Matz 1975), explaining part of the controversy over Jensen's claim that while level I abilities are not affected by SES, level II abilities definitely are. (eg Jensen, 1969). As for sex differences, the finding that girls are generally superior on reading comprehension is widely accepted - boys are up to six months behind girls on reading age tests, and form 60 - 90% of remedial reading classes. However, Asher and Markell (1974), who quote these figures, found that boys' reading comprehension was worse than girls' when each read passages on topics they had rated as uninteresting, but that the difference vanished when both sexes read "interesting" passages. Another finding going against the general trend is that boys may well score better than girls if tape recordings are used to pace their reading, encouraging them to process larger "chunks" (Neville, 1975), so that the levels-of-processing approach can perhaps even be related to sex differences! Similarly a high level of anxiety means that the reader is less

willing to sample cues from the text in order to infer and predict, and tends instead to concentrate on careful sensory decoding, the extreme form of which is "tunnel vision" (Smith, 1975). However, Edmunson and Nelson (1976) found that the provision of a good strategy (interactive imagery) brought the performance of their high-anxiety subjects virtually up to the same level as their low-anxiety subjects, while the provision of a poor strategy (rote repetition) had no such effect.

(d) Conclusion

In sum, it seems that an alarming number of variables, ranging from word length to the socio-economic status of the reader, can affect comprehension on any of a number of levels. Moreover, these variables are almost certain to interact, so that quite large-scale experiments may be required to study the interaction of even a few of them (eg Rohwer and Harris, 1975). More optimistically, a levels-of-processing approach can perhaps offer a framework for understanding at least some of them and for suggesting possible strategies for improving comprehension.

However, a final word of caution is necessary. Faced with the obvious complexities of studying comprehension, yet usually wishing to isolate a single variable or technique, several experimenters have been in danger of creating "artificial" conditions which make their frequent claims of classroom implications less valid. It has already been pointed out that the use of the cloze technique, multiple choice items and recall tests makes the testing of comprehension in the experimental research rather different from what the questionnaire suggested was normal classroom practice, and such a difference is also noticeable in the actual reading materials used. Control of factors like sentence length or word concreteness may result in rather unnatural-looking passages (Stromnes and Nyman, 1974), but the need for control may have even stronger effects. Newman and Saltz (1960), for example, having pointed out that very little previous research had studied connected discourse, went on to construct a "passage" that consisted entirely of a string of sentences like "Historic sights are on the wide river Gibson" and "Fishing villages are on the deep river Lawson", containing no overall structure whatsoever. Conversely, experimenters who wish to emphasize passage structure construct very different passages with strong sequential or narrative elements (eg Bower, 1976).

Similarly, if you wish to avoid the effects of prior knowledge on comprehension, you make up passages about imaginary islanders or breeds of monkey (eg Rohwer and Harris, 1975; Kulhavy and Swenson, 1975), although children's previous knowledge or misconceptions could greatly affect the inferences they draw from a "real" geography or biology text. If, on the other hand, you wish to emphasise the effects of prior knowledge, you make sure that the passage you construct is obscure and unintelligible without it. (eg Bransford and Johnson, 1972a).

Control is obviously vital in experimental research; Bransford and Johnson's passages, for example, are psychologically valid in the same way that research on visual illusions or "impossible objects" is - both demonstrate the importance of one factor in performance by studying it to the point of breakdown. But if the experimental studies of comprehension are also to be educationally valid, as their authors often claim, then the importance of these various factors needs to be re-assessed using "real" materials - a point that was also made about the vocabulary study. Similarly, if the actual form of a comprehension test affects the cognitive demands made on subjects, the

strategies they use and their final success rates, and if no one test of comprehension is clearly "better" than all the rest, then it might be a good idea to re-assess some of the apparently successful techniques offered by the research, but tested under conditions like those actually used by teachers, and these suggestions form the basis for the experiment described in the next chapter.



CHAPTER V

SUMMARY

The first half of the chapter surveys research into two techniques for improving prose comprehension and learning - the use of imagery instructions and adjunct questions. A number of criticisms have been made, particularly of adjunct question studies, and these criticisms and some unanswered questions lie behind the design of the experiment described in the second half of the chapter.

The experiment compared imagery-instructed and adjunct question groups with control groups under two test conditions - with and without the text available, and on a delayed recall test. All reading times were taken and another group was instructed to read the text twice to assess the importance of reading time as a causal factor, while a final control group answered the questions without reading the passage. Literal and inferential comprehension were tested, and the children's reported strategies were also analysed and related to test success.

V 1 TWO TECHNIQUES FOR IMPROVING PASSAGE COMPREHENSION

Apart from suggesting numerous factors that will influence comprehension, the experimental research has offered several techniques for trying to improve it, and two of these techniques form the major treatment conditions in the second main experiment to be described. Imagery can be seen as an underlying theme in the present research, and the first technique is simply to instruct subjects to form images as they read. Several experiments, again mainly American, have suggested that even such simple instructions facilitate comprehension, but there are several important qualifications and the method remains to be tried under conditions more like those actually found in schools. The second technique involves the use of "adjunct questions" - questions inserted in the text to be answered during reading. This approach is rather more complex, and more research has been done on it, but again it has rarely been examined under conditions like those used by teachers to test comprehension, and the two techniques do not appear ever to have been directly compared using children as subjects. This section can be seen as parallel to the first section of chapter three in that it attempts to survey the specific research studies directly related to the experiment to be described.

(a) Imagery instructions and Passage comprehension

The previous general discussion of factors affecting comprehension has already supplied a background for looking at research into the effects of imagery instructions on children's

passage comprehension. To begin with, it will be clear that not all subjects will benefit equally from instructions to form images: age is an important factor, for example, and younger and less able children are less likely to improve their comprehension of a passage after simple imagery instructions (e.g. Shimron, 1974; Guttman, Levin & Pressley, 1977). However, the provision of external pictures as "examples of good images" can significantly improve the comprehension of even quite young children (e.g. Pressley, 1976) or ESN subjects (e.g. Riding & Shore, 1974) and may be particularly useful to "culturally disadvantaged" children (e.g. Rohwer & Harris, 1975). Pressley (1976) reports that twenty minutes' instruction was all that was needed to help his 8 year-old subjects, though rather longer training periods were used by Lesgold, McCormick and Golinkoff (1975) to get children of similar age to draw cartoons to improve their reading comprehension, while Lesgold, Levin, Shimron and Guttman (1975) extended this technique, and were able to improve the listening comprehension even of 6 year-olds by getting them to manipulate suitable cardboard cut-outs, thus avoiding the need for any form of drawing skills. The age dimension is certainly an important one, and there are great differences in the value of provided pictures and imagery instructions for subjects in different age groups when they are directly compared on prose learning (Rasco, Tennyson & Boutwell, 1975).

However, there are several problems in the use of pictures to improve image generation, apart from the obvious one that the

use of external pictures cannot be directly equated with the production of mental images, as already discussed in the second chapter with reference to mnemonics. Not all pictures may be equally helpful, and factors like the type of picture (e.g. line drawing or photograph) or the importance of relatedness, which have also already been discussed with regard to the learning of individual items, will probably affect the use of pictures as external prompts or organisers in passage comprehension (Rohwer & Matz, 1975). Bransford and Johnson (1972 a,b) for example, found that their "partial context" picture, which depicted the same items as their "appropriate context" picture, but rearranged, was of little help in recall and did not improve subjects' comprehensibility ratings of their passage. However, a similar study by Arnold and Brooks (1976) claims to show that an "Unintegrated" picture did not aid children's comprehension of a story, though closer examination shows that the supposedly "unintegrated" items are in fact depicted in a thoroughly misleading way, and, like Bransford and Johnson's, form an inappropriate rather than an unintegrated context. This is particularly important in view of evidence that children's recall of a passage actually declines if accompanying pictures contradict the text, (Pressley, 1977). Conversely, there is the problem that pictures may themselves contain or suggest the answers to the questions, though an ingenious study by Guttman, Levin and Pressley (1977) suggests that pictures from which the answers cannot be derived still facilitate children's prose learning, and they in fact conclude that the value of pictures as prompts in children's prose learning is now firmly established.

The effectiveness of simple imagery instructions alone will also involve various individual differences like those mentioned in the general discussion. For example, if breakdown in comprehension can occur at different levels, thus suggesting different models of reading difficulty (Weiner & Cromer, 1967; Cromer, 1970), then instructions to form images will not necessarily help all poor readers. Thus Rohwer (1973) found that "a visual image organisational strategy" improved the comprehension of "difference" poor readers by inducing them "to attend to semantic characteristics and relationships", but was much less useful for "deficit" poor readers who have vocabulary and decoding problems. Rohwer also noted that children who were good at remembering visual items rather than verbal ones benefited more from imagery instructions, and such a finding is very much in line with Levin's "learner types" (e.g. Levin, Rohwer & Cleary, 1971; Levin, Divine-Hawkins, Kerst & Guttman, 1974), as well as with the other research on individual differences in imaging and their effects on learning that were discussed in the first chapter (e.g. Reinert, 1976). It is also clear that the preferred strategies of the reader will interact with the nature of the materials and presentation - Klee and Eysenck for example (1973) found this with sentences, and Riding and Taylor (1976) found a highly significant interaction between 7 year-olds' imagery performance and their passage comprehension - good imagers scored better on a concrete passage than on a more abstract passage, while poor imagers did just the opposite.

It would thus appear that imagery instructions may help some children improve their understanding of connected discourse, but certainly not all. Pressley (1977) feels that there are grounds for "cautious optimism" and concludes -

Imagery may prove to be a valuable tool for overcoming some intellectual handicaps, a tool allowing some disadvantaged children to derive as much meaning from prose learning experiences as their more fortunate peers, but induced imagery does not facilitate the learning of all types of poor learners, nor does it always increase the prose learning of average and good readers.

In fact, imagery instructions may appear to have no effect at all on the comprehension scores of older and more successful readers. Anderson and Kulhavy (1972) found no significant differences between instructed and control groups of high school subjects, but when they analysed the students' responses to a questionnaire, they found that about a third of the "imagery" group had failed to use imagery, while half the control group had done so spontaneously! In fact, further analysis showed that those who had used imagery had done significantly better than those who had not. However, this evidence is only correlational - subjects who fail to follow imagery instructions may also be poor comprehenders for a variety of reasons. The experiment does re-emphasise the problems of getting older subjects to follow mediation instructions, however (cf. chapter two) and highlights the value of trying to find out what strategies are actually used by subjects, including those in control groups left to their own devices. This procedure is still surprisingly rare in the research into children's comprehension, and was an important point in both the experiments described here.

In fact, a later experiment by Kulhavy and Swenson (1975) is probably the closest to the present study, though it leaves a number of important questions unanswered. Kulhavy and Swenson gave 10-11 year-olds a 20 paragraph passage about an imaginary island, with either no specific instructions (control group) or instructions to form mental images after reading each paragraph. All subjects also had to fill in the answer to a question at the end of each paragraph, and at the end of reading, which was timed to the nearest 10 seconds, half the subjects in each group received a set of questions which were either verbatim repetitions or paraphrases of the ones they had already done, and half the subjects received a placebo test. Finally, a surprise re-test was given to all subjects one week later. On both tests, subjects in the image group performed better than control subjects, though the difference was significant only for the delayed tests, and was particularly strong with paraphrase rather than verbatim questions, supporting the authors' predictions that forming images would aid "semantic" recall and that the effect would increase with time. Retention effects were not due to the presence of immediate tests, and reading times did not differ between the groups, so that Kulhavy and Swenson conclude -

Grade school students remember more from a text if they try to form mental images while reading . . . image instructions act to increase the amount of text-content available over time . . . The implications for school learning are obvious, at least with younger students. If one can supply learners with an efficient

strategy, it is likely that more will be remembered from the study of the text. It seems probable that such procedures may furnish the teacher with yet another tool for facilitating the learning of text-based materials.

The study in fact raises several questions which will be discussed as the introduction to the present experiment, but the suggestion is strongly made by Kulhavy and Swenson, and elsewhere in the research, that while younger subjects may need more explicit prompts and older ones may not benefit at all from imagery instructions, children roughly in the 10-11 age group will improve their learning of school-like materials if they are simply told to form images as they read.

Finally, if imagery really is helpful in this way, several rather more theoretical approaches which have already been mentioned need to be briefly re-stated if the use of images, like mnemonics, is to be seen as anything more than a gimmick, and four theoretical contributions can be seen as relevant. First, at the most general level, it is certainly not necessary to take an "image-as-meaning" position, which involves a number of problems suggested in the first chapter. On the other hand, it may be equally unnecessary to take a "propositional" view (e.g. Pylyshyn, 1973) and insist that all knowledge is expressed

in underlying propositions regardless of its source. A recent paper by Kieras (1978) does attack the limitations of Dual Coding theory and supports a propositional view, but rejects its more extreme forms and suggests that there could be both semantic and perceptual forms of representation differing in both form and content, so that there is perhaps some common ground between the two theories. Similarly, Paivio (e.g. 1971) certainly does not equate imagery with comprehension, but suggests that they are closely related and that images are involved particularly in the comprehension of concrete material, while Kosslyn and Pomerantz (1977) point out that since images display distinct and unique emergent properties, there is no reason to discount their functional importance in cognition, whether or not one accepts the existence of "deeper" forms of knowledge. Moreover, if comprehension takes place on a number of levels, and is affected by a wide range of factors, then its relationship with imagery may well vary considerably according to the subjects, materials and conditions involved. In brief, the suggestion that imagery instructions might facilitate comprehension implies neither the acceptance of a naive "image-as-meaning" position nor the necessary rejection of possible underlying forms of knowledge.

On a slightly more specific level, two more approaches offer ways of relating the experimenter's instructions and materials to the strategies used by the subject. First, Rohwer's elaboration hypothesis (Rohwer, 1973), while claiming to have only "heuristic value" and not the predictive power of a full-

blown theory, emphasises stimulus attributes rather than the modality of internal representations which are the central concern of Paivio's Dual Coding hypothesis. Rohwer proposes a hierarchy of prompts, described in the second chapter, which are relevant to a wide variety of experiments, including the effect of imagery instructions on children's passage comprehension. Simple instructions to form images while reading constitute an "explicit prompt" in which a strategy is offered, but not augmented by the provision of an external aid like the pictures in the vocabulary experiment. The effectiveness of such a prompt will depend on the existing strategies the subject brings to bear on the task: vocabulary learning was a rather novel task for the children, so that any form of provided strategy was helpful, but 11 year-olds are much more used to answering questions on what they read, and may already form images spontaneously anyway so that instructions to do so may be of little help. Conversely, those who lack effective spontaneous strategies might benefit more from augmented prompts in the form of pictures, for example, as they did in the vocabulary study. At any rate, Rohwer's hypothesis would at least suggest the importance of studying subjects' existing strategies as well as the effectiveness of those provided by the experimenter, and this was to be an important part of the experiment.

The other relevant approach is that of levels-of-processing, which has been suggested as particularly applicable to comprehension. Such an approach, while again clearly differing from Dual Coding theory, need not be entirely incompatible with it: Paivio

(1975b) certainly accepts that there are different levels of processing, though his proposals are rather different from those of Craik and Lockhart (1972), since he sees the deeper levels in terms of associations between the verbal and visual coding systems, or within each one. In reading comprehension, images may well be involved at the shallowest level of getting perceptual information from the page - Goodman (1967) claims that the graphic cues picked up by the reader generate images of the text which are "partly what he sees and partly what he expected to see", and there is a possible link here with Neisser's (1976) view of images as "perceptual anticipations, preparations for picking up certain kinds of information", (which would also extend to the effects of images on deeper levels of comprehension). An experiment by Klee and Eysenck (1973) concluded that "imaginal processes constitute an important element of comprehension" though they in fact tested it by getting their subjects to decide whether sentences were meaningful or anomalous, and such a task appears to involve only a very shallow level of comprehension - Mistler - Lackman (1972), for example, showed that subjects could decide whether or not a sentence was meaningful without even realising that it was also ambiguous, and Dooling (1972) also considers that the task involves only shallow comprehension. At any rate images may be important in various ways at the shallower levels, but it does appear that imagery, whether induced by simple instructions or by pictures, may be even more helpful in the deeper levels of comprehension. Kulhavy and Swenson (1975) found

that imagery instructions were particularly beneficial in "semantic recall", though they tested this only at the level of paraphrase questions, a possible shortcoming that will be discussed later. However, Rohwer and Matz (1975) found that their "disadvantaged" subjects improved their comprehension particularly on inference items when pictures were provided, and simple imagery instructions enabled retarded children to draw more inferences from sentences in an experiment by Paris, Mahoney and Buckhalt (1974). Indeed, Aaronson (1976) considers that the formation of images should actually lead to poorer verbatim recall since they transform and recode the original verbal information in a passage. Since it was intended to ask the children in the present study both literal and inferential comprehension questions, the relative value of imagery in answering each type was another point of interest.

The final theoretical approach which seems particularly relevant to the contribution of imagery in comprehension is the idea of "surrogate structures". Pompi and Lachman (1967) suggested that when subjects read a passage, they do store a few lexical items, but they also generate "surrogate structures" - huge "chunks" containing the general themes of the passage. A related research area concerns the idea of "semantic integration", also previously discussed, which has used the "false recognition" paradigm to suggest that subjects tend to store the general idea behind a group of sentences (e.g. Bransford & Franks 1971; Paris Mahoney & Buckhalt, 1974). Further investigations of surrogate

structures supported the idea and emphasised the importance of passage concreteness (e.g. Philip Chalk, 1972; De Villiers, 1974) and Yuille and Paivio (1969) were quick to suggest that "imagery may play an important role in the storage of a theme. Perhaps imaginal representation of the theme is the basis of the surrogate processes suggested by Pompi and Lachman". Whether or not they need necessarily be in image form - and supporters of a propositional theory of knowledge would probably reject the idea - surrogate processes would appear to be important in the recall of passage information: Dooling and Lachman (1971) conclude that "an abstract representation of a passage's central meaning is used as a mnemonic device in the retention of prose". Imagery's value as a mediating link in specific mnemonics has already been discussed, and the idea of surrogate structures, coupled with these other research approaches, can perhaps be seen as extending the usefulness of imagery to the comprehension and recall of connected discourse.

Coupled with what has already been said about imagery in previous chapters, these theoretical views and a number of experiments combine to suggest that instructions to form images might well help children to improve their understanding and recall of passages like those used in schools. However, other techniques have also been proposed, and one which has received particular attention is the use of adjunct questions. Though the main concern of the present study was with imagery, it was felt that it

might be interesting to compare the effect of imagery instructions with another technique in which the use of images is only incidental, particularly since almost all previous research has limited itself to comparing a single chosen technique with a control group, so that very little appears to be known about the relative values of imagery and adjunct questions. The next section, therefore, looks briefly at research into the use of adjunct questions to improve passage comprehension and recall.

(b) Adjunct Questions and passage comprehension

At first sight, the body of research on the use of Adjunct Questions (AQs) - questions inserted in a text to be answered as the subject reads - appears to be rather self-contained, probably because it is almost all derived directly or indirectly from Rothkopf's concept of "Mathemagenic behaviours". Rothkopf's initial interest had been in programmed instruction, and his early definition of mathemagenic behaviors as "those behaviors that produce learning" (Rothkopf, 1965) has since given way to a more cognitive emphasis (Anderson & Biddle, 1975) so that a later definition is of mathemagenic activities as "those student activities that are relevant to the achievement of specified instructional objectives in specified situations or places". (Rothkopf, 1970).

He proposed three types of mathemagenic behaviour - Class I "orientation" which includes getting students to the classroom and

keeping them there, Class II "object acquisition" which involves "selecting and procuring appropriate instructional objects", and Class III "translation and processing" which involves all the student's unobservable cognitive activities, and the bulk of the research has been concerned with this third class. Rothkopf also re-emphasises the important distinction between the "nominal stimulus", or the materials, and the "effective stimulus", which is the way the materials are perceived by the subject (a distinction that has already been discussed in chapter two with reference to the use of pictures in modern language learning, as an attempt to avoid native language).

Although only the nominal stimulus is directly under the control of the teacher, he can try to intervene in the learning process either before it begins (cf. Ausubel's "advance organisers" and the widely-researched importance of preparatory set) or actually during the student's learning, and against this background the insertion of questions in a text can be seen as related not only to the principles of programmed instruction, but also to a more cognitive approach to learning. Thus Rothkopf (1970) suggests that the insertion of questions will encourage the reader to process the text more thoroughly, and that mathemagenic behavior which results in improved learning will be reinforced, and therefore be more likely to re-occur. However, in the same paper, he also suggests that theoretical formulations about the nature of mathemagenic behavior would be premature, and other writers agree

that the work lacks a firm theoretical basis (e.g. Carver, 1972; Wittrock, 1975). Taking a more practical approach, the previous chapter suggested that a large number of variables are involved in passage comprehension, and Carver cheerfully acknowledges that "the retention of prose materials is probably dependent upon an infinite number of variables". Table 32 offers a way of categorising some of these, and forms the structure for the rest of this discussion.

6) Subject variables have already been discussed in some detail, and only a few extra points need to be made here. One criticism of AQ research is that it has almost always used student subjects, and has still rarely been extended to children's reading (Faw & Waller, 1976). A notable exception is that of Swenson and Kulhavy (1974) who in fact conclude that "control of inspection behaviours is a markedly different task with young learners". The age variable has already been shown as very important with regard to the relative effectiveness of subjects' spontaneous strategies and those provided by the experimenter, so that when one notices, for example, that the subjects in a frequently-cited experiment by Rothkopf (1966) were not only adults but were also volunteers paid for their services, it is clear that there may be dangers in generalising from their results to children's normal classroom learning, and this alone could justify an experiment looking at children's comprehension and recall of passages with and without adjunct questions.

Extending this argument, AQ research has also generally

	(i) SUBJECTS	(ii) INSTRUCTIONS	(iii) PASSAGE	(iv) AQS	(v) TEST
N A T U R E	Individual differences (e.g. age, SES). Perception of and expectations about other variables. Implied strategies.	"Vague hortatory" or Specific? Reading aims given? Use of AQS specified?	Text Variables (e.g. length, difficulty, structure) Interest level? Length between questions?	Level (e.g. Verbatim, paraphrase inference) Type (e.g. multiple choice, Short answer).	Level and Type of Questions. AQS Repeated?
P R E S E N T A T I O N		Mode and Source?	Mode and Source? Back Referring Allowed? Time Limit?	Mode and Source? Overt Response? Before or After? All together or Separate? Answers given?	Mode and Source? Immediate or Delayed? Expected Text Available?

Table 32 Examples of Variables Affecting Value of Adjunct Questions.

failed to assess the subjects' perceptions of the task and the strategies they consequently adopt; Carver (1972) points out that the instructions about the use and possible later re-appearance of the questions have often been so vague that subjects may have adopted a variety of what he calls "programs" (i.e. strategies or plans). Put in another way, it should not be forgotten that instructions and questions, and not just the passage, represent only "nominal" stimuli: Mathemagenic behaviors already exist in the reader, and are not simply the result of adjunct questions placed in the text by the experimenter (Frase, 1975). Indeed, one of Carver's most scathing criticisms is that AQ research ignores the cognitive contributions of the subject, and he concludes that -

The predominant behavioristic influence from the past is probably responsible for the current conduct and reporting of research which treats an S as though he were a rat, and excludes the subjective aspects of the experiment which happen to be extremely important in research with humans.

One particularly relevant point is that it may be that subjects not given adjunct questions but expecting a test later might well ask themselves "likely" questions as they read. Flavell (1970) suggests that children develop several "mnemonic mediation strategies", including self-testing, and this does seem to be related to advanced reading techniques like SQ3R. Bernstein (1973) found that children could be taught to generate questions for themselves which resulted in test scores just as good as those obtained by children who had answered questions asked by the teacher, though a less sympathetic conclusion might be that neither was much help, and Wright (1972) proposes that the strategy of guessing what later questions might be and trying to answer them is in fact disruptive because it hinders the reader

in getting "the overall idea" of the passage.

At any rate, it is clear that there is room for some experimental work on the success of adjunct questions in children's passage comprehension and recall, especially if their spontaneous strategies are also investigated and instructions about the use of the questions are made clear, and these points form part of the basis for the second main experiment here.

(ii) Instructional Set, as suggested above, interacts with the subject's existing strategies and perception of the task, so that from a cognitive point of view the two cannot be separated as easily as Table 32 seems to suggest. Rothkopf (1966) considered that "vague hortatory directions in intent affect S's mathemagenic activities sufficiently to elevate postreading test performance", though Carver's criticism of vague instructions has already been noted, and Frase (1975) suggests that readers learn significantly more if specific goals are first stated. There would seem to be a parallel again with Rohwer's (1973) "elaboration hypothesis" in which "vague hortatory directions" would correspond with "minimally explicit prompts" like "read this carefully", and which would therefore be expected to be differentially effective according to the age of the subject. Hiller (1974) criticises existing AQ studies for not explaining the purpose of the adjunct questions or their relationship with the final test, and Carver (1972) certainly feels that the value of adjunct questions has so far been found only when the instructions were so vague that the

subjects did not know whether they would be re-tested on the same questions and were given no reason to regard them simply as learning aids.

The placing of questions before rather than after a section of the passage can also act as a kind of set, directing the subject to look for certain pieces of information rather than others, so that the effect on overall comprehension or learning may be damaging rather than useful (cf. Wright, 1972, above). Rothkopf's "Law of minimum learning" proposes that subjects will make the minimum amount of effort necessary to perform a specific task, and there is a clear similarity here with Anderson's work on attention (Anderson, 1970) which also suggests that a task which is too easy results in only poor learning. The similarity with programmed learning research is again apparent here, and Anderson also describes research on the value of providing an obvious answer (as in programmed instruction), a less obvious answer, or no answer at all to questions inserted in the text. The nature of the questions may also have quite a subtle effect - Rothkopf and Bibiscos (1967) found that students consistently given questions of a certain type (e.g. always requiring a measured quantity as an answer) improved their scores on previously un-asked questions of the same type in the postreading test, and that this effect was stronger on the second half of the passage, when their behaviour had presumably been "shaped".

Finally, it is possible that the method of presentation of instructions, and other forms of establishing a set, may have some

effect. This does not appear to have been studied with the instructions themselves, but there is some evidence that the mode of presentation of the questions could be a factor in classroom learning: Rothkopf and Bloom (1970) found that orally presented questions resulted in both longer reading times and higher test scores than written questions did, though the two factors were not separated - an important point to be discussed later.

(iii) The Passage. As with the section on subject variables, previous discussion has already indicated a large number of factors that will affect passage comprehension, ranging from individual word properties to the overall structure of the passage and the reader's interest in it. AQ research does not appear to have manipulated passage variables, although two studies described by Anderson (1970) did try to compare the effects of adjunct questions on an "interesting" and a "boring" film, and suggested that the questions slightly improved attention where it might otherwise have been lacking. Similarly, Hiller (1974) reasoned that students needed most help with materials they find hard or uninteresting, so selected a passage which was not only difficult, but also deliberately boring! Most interest is in the questions, however, and the reading materials are controlled by being kept constant, rather than manipulated, while it is realised that text characteristics do obviously modify mathemagenic behavior (e.g. Frase, 1970).

One variable that could be considered under this heading, however, is the length of passage between questions; on the one

hand, it appears that the closer a question is to the relevant section of text, the more likely it is that subjects will get that same question right on a later test (Anderson & Biddle, 1975), though there may be an "optimum level" here since Anderson's own research suggests that questions which are very easy do not facilitate recall (Anderson, 1970). On the other hand, McGaw and Grotelueschen (1972) suggest that adjunct questions can affect the subject's reading even several pages later because they increase his general level of attention. The value of long or short sections of passage between questions thus probably interacts with the reading goals and the nature of the test, and such a conclusion is supported by the finding of Swenson and Kulhavy (1974) that short sections resulted in superior immediate recall of specific information, while longer sections gave better scores on a delayed recall test.

Turning to the presentation of the passage, there are two issues that appear to be of central importance in AQ research, the first concerning the opportunity to refer back to the passage when answering the questions, and the second the control or measurement of reading time. The first issue in fact involves the whole area of mode of presentation: differences between listening and reading comprehension have already been emphasised, and one important factor here is that in listening comprehension studies, subjects obviously cannot refer back to earlier parts of the passage as they normally can when they read. Several of Carver's (1972) criticisms of the research have already been

noted, but one of his strongest is that the experiments are lacking in educational importance for this very reason, a criticism echoed by Hiller (1974). Carver concludes -

In no case has any of the recent research allowed Ss to refer back to previously covered materials or questions. In most practical situations where Ss are given instructional materials, they are allowed to refer back to a previous section while they are reading or studying . . . the recent research results must be regarded as irrelevant to most practical situations.

Although neither Carver nor Hiller give any evidence for this claim, the results of the second part of the questionnaire (chapter four) would certainly suggest that, at least in English lessons, children are normally allowed to look back over the text when doing the "test" questions, though a distinction should be made between testing "passage comprehension" and "prose learning" in spite of the fact that the research itself often fails to make such a distinction, and both were to be studied in the present experiment. Carver also emphasises recent research because one of the earliest studies in the field (Washburne, 1929) in fact did allow back referral, and disagreed with the results of most of the research since, with regard to the relative effects of certain types of question (see below). Even when reading comprehension is studied, experiments have consistently prevented back-referring, again possibly because of the original basis for the approach in programmed learning research. For example, Rothkopf and Bloom (1970) presented their 16,000 word text on 108 slides, and studies using booklets have always instructed subjects not to look back (e.g. Rothkopf, 1966; Rothkopf & Bibiscos, 1967). In a

sense, this is to falsify the whole reading process - Carroll (1974) in his defence of the printed word against the "misleading and possibly fascist ideas" of its opponents, points out that one of print's great advantages is that it "gives us more options for selecting a strategy for comprehension", including the ability to re-assess what we have already read. In fact, Rothkopf has more recently accepted this criticism of the research (Rothkopf, 1974), but he still appears to offer no evidence to support the value of adjunct questions when back-referring is allowed, and merely concedes that they may not work under all conditions. Because of this one of the aims of the present study was to see if the reported facilitation of adjunct questions was still apparent when subjects were allowed to look back at the text while answering.

The second issue centres on the importance of reading time, and again certain criticisms have been made of the research. Time is clearly an important factor in prose comprehension or learning - Riding and Shore (1974), for example, improved the listening comprehension of ESN children simply by slowing down the rate of presentation, and Carver (1972) cites several pieces of evidence which relate reading time to prose learning. Yet few AQ studies have seriously taken reading time into account (Faw & Waller, 1976) and Carver claims that the frequently reported superiority of AQ groups is simply due to the fact that they spend more time reading the material, while the small gains, under limited conditions only, make the use of adjunct questions an inefficient form of improving learning. Faw and Waller (1976)

are equally critical; they too point out that because reading time has not been adequately studied, the small improvements apparently due to adjunct questions fail to take account of the longer reading time required. They conclude -

We do not know of a single study involving insertion of questions in written prose in which efficiency scores of experimental groups have been demonstrated to be significantly higher than those of control groups on incidental learning.

In defence, Rothkopf (1974) agrees that the time factor is important, but claims that increased reading time involves "more processing" which is a direct result of the adjunct questions.

The actual control of reading time by means of slides (e.g. Rothkopf & Bloom, 1970) or pacing by means of recordings as the subject reads (e.g. Neville, 1975) is one possibility, but it tends to alter the whole reading process even more drastically than the prevention of back-referring does, so that its practical usefulness should be only as a remedial technique designed to change inefficient reading. On the other hand, the measurement of reading time should be of interest and this has in fact been done in several AQ studies, though there seems to have been no attempt to go beyond finding the mean reading times for the various experimental groups, probably in the hope that there would be no significant differences. The relationship between individual reading times and test scores does not appear to have been studied, though research on cognitive styles, particularly Kagan's work on "reflection/impulsivity" might suggest that the relationship between learning success and time spent on the task is not a simple

one (e.g. Kagan, 1971). Some sophisticated techniques have also been developed for studying the time taken by an individual reader to read any section of text, as well as the incidence of back-referring, skimming and different kinds of "read" (Augstein & Thomas, 1973), but the use of machine presentation makes the reading process a little artificial, and there are obviously practical disadvantages for its use in group studies. In the present experiment, it was therefore decided not to attempt to control reading time in any way, but to measure the time taken by each child to read the text, while at the same time deliberately encouraging one of the experimental groups to spend a lot longer reading the passage, without giving them either adjunct questions or imagery instructions, to see what effect this had on their scores. A final point which does not seem to have been mentioned in the research is that it was also decided not to warn the children that they would be timed, since even this might affect their reading, especially in a group situation. (In fact, at the end of the study, when the children realised their reading times had been taken, they appeared just as interested in them as they were in their test scores, and just as pleased with a short time as with a high score).

(iv) Adjunct Question variables themselves have obviously been of central interest in the research, though the nature of the questions has perhaps been studied rather less than different ways of placing and presenting them. It is also difficult to separate adjunct and test questions, since in much of the research

the two have been at least partly the same. One finding is that an overt response is valuable (e.g. Kulhavy & Swenson, 1975) and that the effects of adjunct questions seem stronger if "short answer" questions rather than multiple choice items, are used (Anderson & Biddle, 1975). As previous discussion has suggested, it is also generally accepted that there are different levels of questions, and that a basic division can be made between "literal" comprehension questions, where the answer is contained explicitly in the text, and "inference" questions where the reader has to "go beyond the information given". At first sight, AQ research seems to have recognised the distinction and to have studied the relative effects of both types of question: Swenson & Kulhavy's (1974) experiment on 11 year-olds discussed Anderson's "phonological" and "semantic" encoding, and devised questions to assess both, as did Hiller (1974), while Anderson and Biddle (1975) review several other experiments which have also made this distinction.

However, closer inspection reveals that, while phonological encoding has been assessed by the use of verbatim questions, semantic encoding questions have been restricted to the level of paraphrase. For example, Swenson and Kulhavy used the following two adjunct and test items for the sentence "The Islanders construct their clothes from palm leaves".

The islanders construct their clothes from . . .
(verbatim item)

Things the natives wear are made with . . .
(paraphrase item).

Similarly, Anderson and Biddle's (1975) review is restricted to studies on the effects of these two types of question, their conclusion being that verbatim items generally facilitate short-term recall better than paraphrase questions do, but that paraphrase questions give better delayed recall scores. However, it will be clear from previous discussion in chapter four that Anderson's classification ignores other possible levels of comprehension like those suggested by Barratt (Clymer, 1972) and that, while paraphrase questions certainly are "deeper" than verbatim ones, they still require little inferential comprehension. Consequently, it was decided to devise test questions that would hopefully test this deeper level of understanding, as well as more conventional ones which could be answered from information explicit in the text.

Turning to the presentation of adjunct questions, one issue seems to have dominated the research. Anderson and Biddle (1975) claim that questions spaced throughout the text are more helpful than grouped questions, though it has already been suggested that the value of long or short sections of text between questions depends on the nature of the learning required on the test. Given adjunct questions spaced throughout the text however, much of the research has concentrated on the issue of whether the question should come before or after the section of text containing the answer, and the general trend of findings is clear. On the whole, adjunct questions before the text give better scores if the same questions are asked again on the test, but actually depress scores on questions not previously asked. In contrast, questions

after the text appear to improve test performance on both old and new questions, thus giving both a specific and a general facilitation (e.g. Rothkopf, 1966; Rothkopf & Bibiscos, 1967; Frase, 1970). Writing in 1970, Anderson's survey of the research concluded that -

Inserting questions in reading material, either before or after the section of the passage on which they are based, sharply improves performance relative to a no-question group when the same questions are repeated again on the criterion test. These studies also confirm that performance on nonpracticed criterion test questions improves only when questions are inserted in reading material after the relevant passages.

An interesting exception to this general finding was discovered by Rickards (1976) who also pointed out that much of the previous research had limited itself to rather shallow questions, and went on to compare "conceptual" questions placed before and after each segment. Contrary to the usual trend, Rickards found that conceptual questions placed before the section of text gave better delayed recall scores than either verbatim or conceptual post-questions, and explained the results in terms of preparatory set. In general, however, the superiority of post-questions has been confirmed (e.g. McGaw & Grotelueschen, 1972; Swenson & Kulhavy, 1974) to the extent that, by 1975, Anderson and Biddle were able to draw up a table showing how many studies had found positive, negative or non-significant effects of the two question positions on the test. Although there are distinct dangers in drawing up such a table (ranging from the pre-conceived notions of journal editors to a reluctance to submit or publish non-significant results), the trend was clear,

and by way of explanation, Rothkopf's "Law of minimum learning" and Anderson's work on attention both suggest that, given a specific task like finding the answer to a question, subjects will limit their attention and effort to answering that question only, at the expense of other information.

However, criticisms of vague instructions leaving the use of the adjunct questions and the nature of the final test in doubt must be borne in mind, as must Washburne's (1929) finding that, when back-referring was allowed, adjunct questions after the text did not facilitate test scores on new questions. Carver's (1972) criticisms of both these points have already been discussed, and it was hoped that this was one of the issues that could be re-examined in the present study.

(v) Test variables. Because of the strong interactions which exist between variables, most of the points concerning the criterion test have already been covered (e.g. the relationship between "old" and "new" test items and the placing of adjunct questions, or subjects' expectations about the test). Inferential comprehension seems to have been largely ignored in test items as well as adjunct questions, and another serious criticism is that the majority of studies have tested only immediate recall, again making their educational relevance dubious. (Faw & Waller, 1976). This is an important point, and it has already been suggested that a distinction should be made between comprehension and prose learning. When "comprehension" is tested in English lessons, the questionnaire strongly suggested that children are usually allowed

to refer back to the text, and if teachers of English or other subjects want to test children's "prose learning" without the text present, they are surely more interested in long-term retention than in immediate recall. It would thus appear that the exclusive use of immediate recall tests in a number of experiments may make them irrelevant both to "comprehension" and "prose learning" as they are tested in schools, especially if they also limit themselves to verbatim or paraphrase questions only.

A number of criticisms can thus be made about existing research into the use of adjunct questions, ranging from the limited choice of subjects and a lack of interest in their strategies to the artificial effects of not allowing back-referral and a widespread use of immediate recall tests as measures of only the more shallow levels of comprehension. Carver (1972) criticises both the research, which he regards as irrelevant to education, and the whole concept of mathemagenic behavior, which he regards theoretically as a "barbarism that contributes more to confusion of thinking than to clarity of thinking". More sympathetically, Faw and Waller (1976) propose that future AQ research should (a) collect time data, (b) use and study control groups, (c) be more "realistic" by creating conditions more like those found in school and (d) study long-term effects. The second experiment

aimed to do these four things and to look at some of the other critical points that have been raised, as well as to compare the effectiveness of adjunct questions and imagery instructions in children's prose comprehension and recall.

V 2 AN EXPERIMENT IN CHILDREN'S PASSAGE COMPREHENSION AND
LEARNING

(a) Summary Two hundred and two 11 year-olds were given an 800 word passage to read with either imagery instructions, adjunct questions, or left to their own devices (control group). Each group was randomly assigned to one of two conditions, one being allowed to retain the text while answering literal and inferential comprehension questions on it, and one not. Another group was simply told to read the passage twice , and a measure of the "passage dependency" of the questions was obtained by getting a final group to answer them without seeing the text. All subjects had reading times taken, and all took an identical but unexpected recall test a week later. Group results showed that, while text availability strongly affected both immediate and delayed scores, imagery and re-reading instructions had not significantly improved scores beyond those of the control group, while adjunct questions had facilitated only recall on these same questions when they were asked as part of the delayed test. However, analysis of the children's reported strategies not only suggested that different strategies had different effects and were themselves differently affected by test conditions, but also offered an explanation for the lack of significant differences between the group results, which was borne out when some of them were re-examined.

(b) Introduction - design of the experiment

In view of the discussion of comprehension in chapter four and the review of imagery and adjunct question research in the previous section, a number of points were to be considered, and

this introductory section relates what has already been said to the actual design of the experiment.

(i) Subjects. AQ research has rarely used children as subjects, so that a study of any age group should be of interest. Imagery research suggests that by the age of eleven or so, children should be able to make use of an "explicit prompt", in the form of simple instructions to form images, to improve their passage comprehension. Chapter one also suggested that, although techniques may form an important part of the primary school curriculum, the subject-centred emphasis of secondary schools means that the "products" or content of learning may sometimes be emphasised at the expense of "processes" or learning strategies, and help with study techniques may be rarely given - the questionnaire results supported this view with regard to prose comprehension and learning, in English lessons at least. Consequently, it was felt that children in the first year of secondary school could profitably be used as subjects again, to see what spontaneous strategies they use in some hopefully "realistic" reading tasks, and to judge the effectiveness of giving them adjunct questions or imagery instructions.

(ii) AQ and Imagery groups were to be directly compared, largely because this does not ever seem to have been done, at least with children as subjects. The closest published study is perhaps an experiment by André and Sola (1976), who got students to rate sentences either for image vividness or pronounceability and to learn them, being tested during acquisition by either paraphrase

or verbatim questions. They found that image rating did not lead to better recall on the test, while the use of paraphrase questions did, but this is a far cry from the deliberate use of instructions to examine children's comprehension and recall of connected discourse. An unpublished experiment by D'Amico (1976) did also get 36 postgraduate students to read a 21 page passage from Carl Rogers, using four treatment groups given either imagery instructions or adjunct questions, each with or without feedback. He found no significant differences, but did not use any control groups and did not try to find out if the subjects had followed the instructions - an important point with student subjects, if supposedly less necessary with children (e.g. Wittrock, 1975).

Surprisingly, one apparent opportunity to compare the two strategies was not taken; although they used exactly the same passage and children in the same age group, Kulhavy and Swenson (1975) do not even mention Swenson and Kulhavy (1974)! The first study used paraphrase and verbatim adjunct questions before and after each section of text, and in fact found that, contrary to the usual results of experiments on student subjects, postquestions did not improve scores on new items in the criterion test. The second study used the same passage to compare control and imagery instructed groups, but all subjects also answered adjunct questions as in the first experiment, so that the effects of the two techniques were not separated. Closer inspection reveals

that the adjunct and test questions were changed slightly for the second experiment (though the writers do not explain or comment on this), thus making comparison between the two sets of results impossible.

In this experiment, then, groups given AQ and imagery instructions were to be compared directly on several tests, both with each other and with control groups left to their own devices.

(iii) AQ Instructions. Particular care was taken with the AQ instructions to emphasise that the adjunct questions were simply for the children to check on their own reading, and would definitely not appear in the test which all groups were told they would be given afterwards. Since the research strongly suggests that only postquestions improve scores on "new" test items, each question was placed after the appropriate section of text, as alternating pages of a booklet, and the children would be told not to look ahead at the question. However, in view of the criticisms of "unreal" conditions in almost all AQ research, the children would also be told that, once they had read the page and turned over to read the question, they could look back to check on what they had read. They could also look back to any page at any time during reading, and these two conditions were seen as possibly important divergences from previous research, hopefully making the present study much more like the normal reading of a text.

(iv) Imagery instructions were to be limited to a simple

instruction to "try to picture the story in your mind as you read it". However, since there may be problems in forming images while reading, as opposed to listening, the children would also be instructed to stop at the end of each page and to try to picture in their minds the events they had just read. As with the AQ groups, they could look back at previous pages at any time during reading. Finally, it was felt that a check on how far they had followed the instructions would also prove useful.

(v) Control groups would simply be told to read the passage carefully, with added instructions about being able to look back whenever they wished. As in the vocabulary study, they would later be asked to say what spontaneous strategies they had used, since very little appears to be known about what children do on such tasks when left to their own devices. Pressley (1976), for example, found that 8 year-olds given imagery instructions and picture prompts did only slightly better than subjects given neither, and suggested that one possible reason could be that control group subjects used imagery spontaneously, though he did not try to find out if this was the case. In fact, the general research finding is that facilitation in prose learning produced by imagery instructions tends to be only slight (Pressley, 1977) and it was felt that more information about what control subjects actually do might throw some light on this trend. It was also hoped to relate any reported spontaneous use of imagery to success in the tests, as in the vocabulary experiment, and not just to

compare control and experimental group means.

(vi) Reading time was to be measured for all subjects, though they would not be warned that this would be done. Some writers, particularly those critical of AQ research (e.g. Carver, 1972) suggest that reading time is in fact the most important effective variable, and in an attempt to assess this criticism it was decided to run an extra group to be instructed exactly as the control group except that they would be told to read each page twice. Longer reading times were, therefore, expected, but no deeper processing strategies would have been offered. (This group is referred to as the "Repeat" group). Superior scores by this group would add strength to the argument that reading time alone is a highly effective variable.

(vii) The passage was to be the first 800 words or so of a "reader" actually used in the school by second year pupils. It was thus "real" in that many of the children would eventually read it, but the effects of prior knowledge were kept as small as possible by the use of a fictional text. It has already been suggested that specially-constructed passages may seem artificial and may emphasise certain text attributes at the expense of others. Pressley (1976) can again be criticised here, for although he claims that the test passage he used was similar to ones children are expected to read and remember in school, closer examination shows that it involved sentences describing, amongst other things,

a cow bouncing a basketball and a rat using a skipping-rope. In view of what has already been said about image bizarreness, the pictures and images involved may well have been more memorable than ones generated for real school texts, thus exaggerating any apparent effects of imagery.

(viii) The Questions. In view of the questionnaire results and the conclusions of Anderson and Biddle (1975) it was felt that "short-answer" questions would be more "real", and in the case of the adjunct questions, more beneficial than multiple choice items. In fact, all the questions were discussed with several experienced English teachers before use. If the nature of the questions used strongly affects subjects' strategies and scoring levels (e.g. Aaronson, 1976; Levin, Ghatala, Guttman, Subkoviak & Bender, 1978), then it was felt that the disadvantages of "short answer" questions would be outweighed by their much greater similarity to actual classroom practice.

It has also been suggested that much of the research which appeared relevant had limited its testing of comprehension to shallower verbatim and paraphrase questions. An attempt was thus made to generate some questions which would also require inferences from the text, particularly in view of the possibility that imagery might prove especially useful in answering such questions.

(ix) Passage dependency of all questions was to be assessed by the use of a small group of subjects who would try to answer them

without reading the passage ("Questions Only" group). The scores of this group were expected to provide a base-line which has been lacking in a number of studies, both of adjunct questions and imagery instructions.

(x) Test conditions were to form one of the most important variables in the experiment. The widespread use of immediate recall tests as a measure of comprehension is clearly at odds with normal practice in English lessons, though Tuinman and Farr (1972) claim that for all practical purposes, the presence or absence of text makes no difference when children answer questions on it, with both long and short passages and with questions both high and low in passage dependency. It was therefore decided to run two groups for each of the control, AQ and imagery treatments, one allowed to use the text when answering the questions, and the other not, each with appropriate prior warning. If various strategies, whether spontaneous or in response to instruction, were differently affected by having the text available or not, then the claims of direct educational reference made by much of the research into comprehension could be called into question.

Finally, if teachers are interested in "prose learning", their concern will be less with immediate recall than with long-term retention, which has again been studied less frequently in the research. Consequently, all subjects were to be given a recall test, without the passage available, a week later. It is well established that any form of adjunct question facilitates

performance on the same question repeated in the criterion test, so that there seemed little point in including the adjunct questions in the immediate test here. However, less is known about the long-term effects, and it was therefore decided to include the adjunct questions in the delayed test, to see if any specific facilitation would persist.

These were the main points considered in the design of the experiment now to be described.

(c) Subjects. The subjects were the same children who had taken part in the vocabulary experiment - the whole first year of a local comprehensive school. They were taught English, like French, in their mixed-ability form groups, and previous analysis had suggested that these groups were similar, in reading age at least. The study was carried out about three months after the vocabulary experiment, so that the mean age of the children by then was 11 years 11 months. In all, 202 children took part in this second study.

(d) Materials.

(i) The text. In common with practice in many other schools, the Head of the English Department had drawn up lists of "readers" to be used with each year group. Teachers were free to choose from these lists, but the year restriction meant that teachers were not

likely to pick a book only to find that some of their class had read it already. It was thus possible to choose a book for the experiment from the second-year list which very few of the first year pupils would have read, and which was at the same time "real" material in the sense that many of them probably would read it the following year. All the children were in fact asked if they had read the story before they took part in the experiment, and only one pupil out of the whole year group said that he had.

The book chosen was Lee McGiffin's On the trail to Sacramento, set in the California Gold Rush of the 1850s and based on the diaries of an actual expedition. The plot is, at least according to the publishers, "a first-rate adventure story, dramatic, informative, and crisply written", though since the main characters are a fifteen year-old boy and his dog, it might be thought to appeal rather more to boys than girls. The first fifteen paragraphs were used, without alteration, as the text, giving a length roughly similar to that found in several "traditional" comprehension books used in the first year. (About 800 words). The full text is given in Appendix D.

The passage was re-printed in booklet form, as seven pages of text. (This was necessary for the Adjunct Question experimental group procedure but it was obviously important to present the material in the same form to all groups, even though this was not exactly the way the passage appeared in the book itself).

(11) The Questions. On the basis of several years' experience in writing comprehension questions, E felt reasonably confident in drawing up a set of draft questions on the passage which would give a fairly wide/spread of scores. These draft questions were then discussed with the seven English teachers, but only a few minor alterations had to be made. In keeping with the research literature and with questions found in a number of comprehension passages in books used in the first year, a distinction was made between "literal" and "inference" questions, and this too proved acceptable to all the teachers, though one of them did make the point that it was not always easy to distinguish between the two in practice and that inference questions often tended to come only after the literal ones in textbooks. To this extent, then, it was felt by E, and apparently agreed by the teachers, that the exercise was quite "real" in the sense that they themselves might set similar work. The only adverse comment, made by one teacher, was a prediction that several of her children would find some of the questions difficult, and that she would normally include some very easy questions in order to encourage the less able children. However, she did agree that the work could be regarded as an acceptable "test" similar to that given to all the children at the end of the first year.

There were three questions for each page of the booklet. One literal and one inference question appeared on the question sheet, and another literal one appeared directly after each page of text, inserted for the adjunct question group only. In fact, two literal questions were produced for each page, and E decided by

tossing a coin which one would be used as the test question, and which the adjunct question. Care was taken to base the questions evenly on information at the beginning, middle and end of each page as far as possible, and in random order.

(e) Procedure. All the children were tested in their normal English classes, which were exactly the same groups used for French. They were randomly assigned to the various experimental conditions, and were first assured that this second experiment did not involve French but was about reading a story and answering questions on it. Each class was taught by a different teacher, but according to the teachers, all seven were quite familiar with comprehension exercises. They were all briefly shown a copy of the book, and told that they might well have to read it next year. They were not told any details of the story. Response sheets were given out, and instructions given to the various treatment groups as below. The children were always told to put their name but not their class group on the answer sheet (so that answers could be marked "blind"). Finally, they were assured that spelling and punctuation mistakes would not count. For each of the main control and experimental conditions, two groups were run, one allowed to refer back to the text, as is the usual practice in school comprehension work, and one not allowed to refer back, as is the usual practice in experimental research.

Control groups were instructed as follows:

"In a moment I'm going to give you all part of the story, printed in a booklet like this". (E. held up one of the booklets).

"I'll put it face down on your desk, and you mustn't start till I tell you to. When I say so, turn the booklet over and start to read. Read each page carefully, and then go on to the next page when you're ready. You can stop or look back if you want to. When you think you're ready to answer the questions on the story, put your hand up and one of us will come round and give you a sheet with the questions on it". (E briefly held up one of the question sheets). "Some of the questions are quite difficult, so don't worry if you think you're not doing very well - I don't expect anybody to get full marks". The children in the control group were in fact the same ones who had been in the vocabulary experiment control group, with one class now randomly assigned to each of two conditions. The first (control, with text) were then told that they would be allowed to keep the story and look at it while answering the questions. The other class (control, no text) were told that the story would be taken away so that they could not look at it while answering the questions. After a reminder that this was a test, and that there should be no talking or copying, the booklets containing the passage were distributed face downwards. E finally made sure that all the children understood what to do, and told them to begin.

With the help of the class teacher, question sheets were distributed as the children raised their hands, and a note made of the time taken for reading (to the nearest quarter minute) on each child's answer sheet. They had deliberately not been warned that they would be timed, since this is obviously not normal

practice in a comprehension lesson, and might have affected their reading. The timing procedure was based on that of Swenson and Kulhavy (1974) and it was made clear that there was no time limit either on reading the passage or answering the questions.

Imagery Groups (Im) were given identical instructions except that, instead of being told to "read each page carefully", they were told to "try to picture the story in your mind as you read it. Stop for a moment at the end of each page and try to picture what you've just read". It has been suggested that reading might interfere with the formation of images, so the children were also asked to try to form images after reaching each page, following the procedure suggested by Pressley (1976). Apart from this, the whole procedure was identical to that of the control group, with one class being allowed to keep the text while answering, and the other not, with suitable advance warning given to each group.

Adjunct Question Groups (AQ) received identical booklets except that a page was inserted after each page of text, with a single verbatim question, relating to the previous page. This was explained to the children before they began to read, and they were told not to look ahead to the question before reading the page of text. However, it was made clear that they could look back at the page after reading the question, and refer back to any previous pages at any time they wished. It was also emphasised that the

questions were there only to help them check on their own reading - they were not to write any answers down, and these same questions would definitely not be part of the test they would do after completing the passage. All adjunct questions referred to the previous page, never the following one, and this too was pointed out. Two classes were randomly assigned to the "with text" and "no text" conditions, with appropriate warning.

"Repetition" Group were instructed exactly as the control group except that they were told to read each page twice before going on to the next. Only one type of test was given - questions without text available.

Questions only. Finally, a small group was given only the questions (adjunct first, then the question sheet) in an attempt to assess the "passage dependency" of the questions used. No timing was carried out with this group.

When all the children had finished answering the questions, those in the control groups were asked to say what strategies they had used. Since the materials here were in the form of connected discourse, unlike the ~~discrete~~ items in the vocabulary study, it was felt that the children could be asked what strategies they had used generally, rather than splitting up the passage into separate pages, or even into paragraphs. In the vocabulary study, the strategy descriptions had been collected only after the delayed test, so that E could go through each item separately without affecting later test scores by the

additional rehearsal. In the comprehension study, however, the children were not shown the passage again, but were simply asked what strategies they had used, and it was thus felt safe to ask them this after the immediate tests, rather than wait a week till all the tests were over. Five categories were offered, and the children were asked to copy down the label of any they had used - thus being encouraged to put more than one if necessary. (This was felt important especially in view of the very few reports of combined strategies to learn the separate vocabulary items).

Guessed at Questions - you guessed at what the questions might be as you read, and tried to find the answers.

Picture - you tried to picture the story in your mind as you read it.

Read again - you stopped and read through parts (or all) of the story again.

These first three categories are clearly similar, at least in some ways, to the strategies actually given to the AQ, Image and Repeat groups respectively, though they are obviously not exactly the same. The final two categories were -

Don't know what you did, just read the story and answered the questions.

Other - subjects were encouraged to write down any other methods they had tried to use.

As before, E was careful not to suggest that any category was "better" than any other, and that some people might well not know exactly how they went about the task.

Similarly, a check was carried out to see how far the image groups had in fact used images. E suggested that some of the children might have found it hard to picture the story all the time, and that nobody would be annoyed if they admitted that they had not been able to do so. Five categories were written on the blackboard, and the children were asked simply to copy down the phrase that was closest to what they had done. The five possibilities were -

I pictured the story in my mind:

All the time.
Most of the time.
Some of the time.
A bit of the time.
Not at all.

Although the first and last categories are not ambiguous, it was possible that the children might differ in their interpretation of "most", "some" and "a bit". In an attempt to minimise this, E reminded them that there had been seven pages, and said that "most of the time" would mean you had managed to form pictures for roughly five or six pages, "some of the time" would mean about three or four pages, and "a bit" one or two pages. The decision is still obviously very subjective, and perhaps not too much weight should be put on it, but it might be useful to know how far the children felt they had been able to follow the instructions, and if their reported compliance was related to success in the tests.

Scoring. No scoring was done until all the response sheets had been collected in and thoroughly shuffled. Twenty sheets were then drawn out at random, and checked against the draft

marking scheme drawn up by E. Several changes were made, almost always because of answers not considered by E but which seemed to deserve some credit. Up to two marks were awarded for each question, ignoring spelling and punctuation.

Although such a marking scheme may appear very subjective, three points can be briefly made in its defence:

- (i) Most important, it is probably close to the sort of comprehension exercise actually set by teachers themselves, as the questionnaire suggested.
- (ii) The revised marking scheme in fact presented surprisingly few problems. After revision, "any other answer" apart from the alternatives given was awarded no marks, for the sake of objectivity.
- (iii) All the answer sheets were marked "blind" - E could not tell which treatment group a subject belonged to, since all the sheets were shuffled before marking and only re-sorted by name into class groups when all the marking was complete.

All the questions and final marking schemes appear in Appendices E and F.

Delayed Test. Exactly one week later, all available subjects were given an unexpected test consisting of the original adjunct

questions, followed by the same questions used the week before. They were not provided with the text, but simply told to "try to remember the story you read last week". It was emphasised that this was a difficult task, and that nobody was expected to do really well. Marking was exactly as before, with the same technique used to revise the draft scheme for the adjunct questions. Again, no marking was done until all the sheets had been collected and shuffled, so that all answers were again marked "blind".

(f) Results. As in the vocabulary study, the results are presented in two parts, one dealing with the differences between groups, and the other relating reported strategies to success in the tests.

(i) Group Results. Since there were eight groups involved, seven of whom took five tests, the results are rather complex and are therefore presented graphically rather than in tables. Figures 8-12 show the mean scores and figure 13 the reading times in minutes for each group; scores in each case are out of 14, and in each table the following abbreviations are used:

C	Control Group, no text
CT	Control group, with text
I	Image group, no text
IT	Image group, with text
AQ	Adjunct question group, no text
AQT	Adjunct question group, with text
R	Repeat group
QO	Questions only group

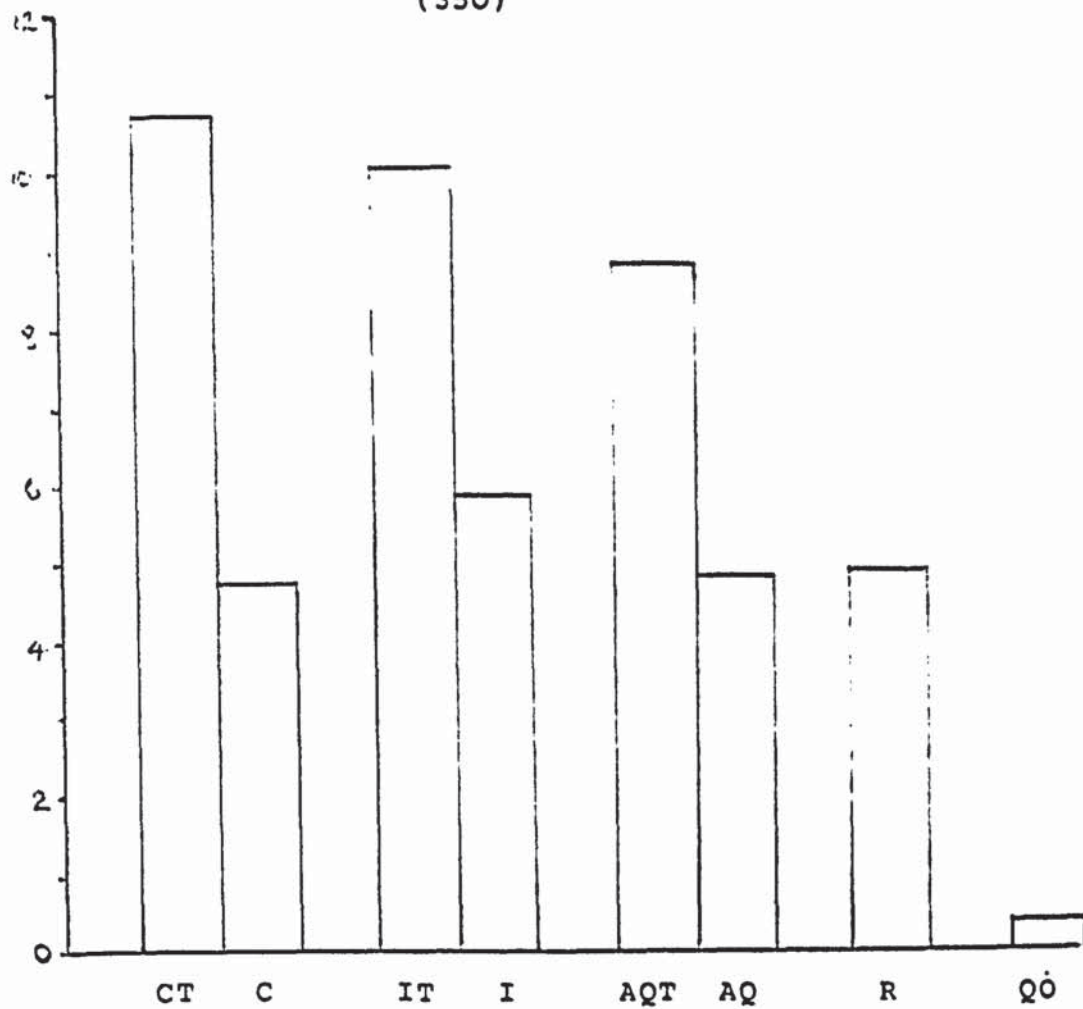


Figure 8 - Mean scores for each group, immediate literal questions.

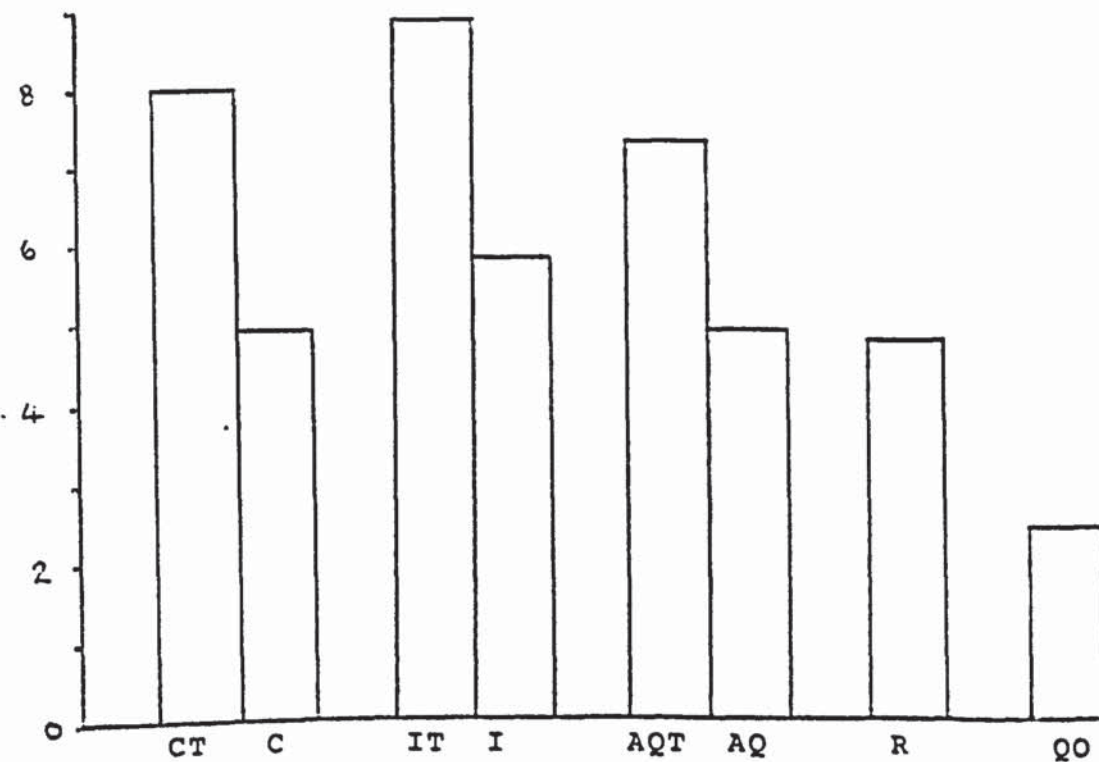


Figure 9 - Mean scores for each group, immediate inference questions.

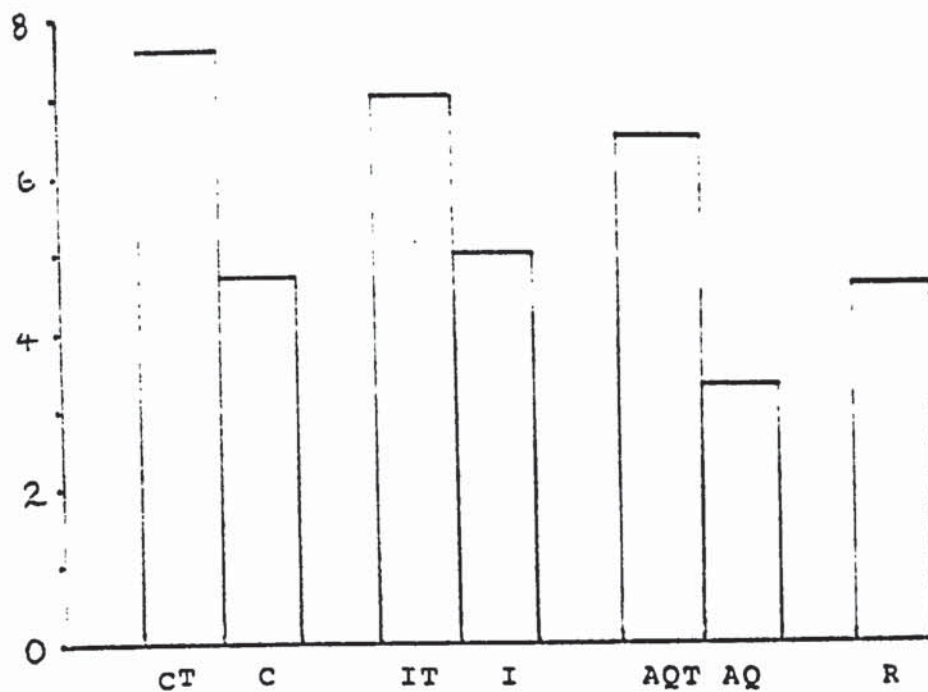


Figure 10 - Mean scores for each group, delayed literal questions.

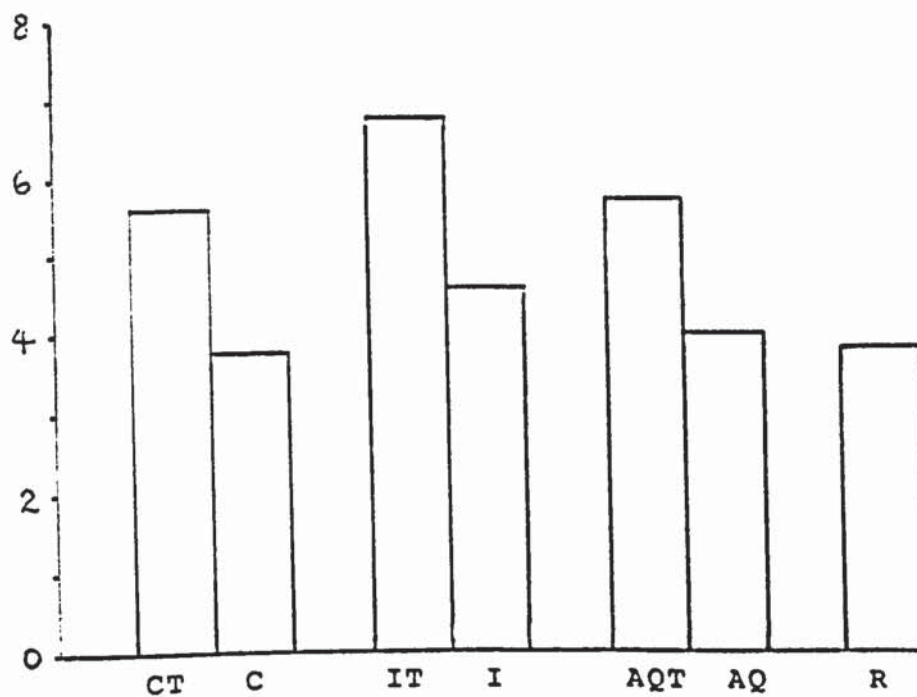


Figure 11 - Mean scores for each group, delayed inference questions.

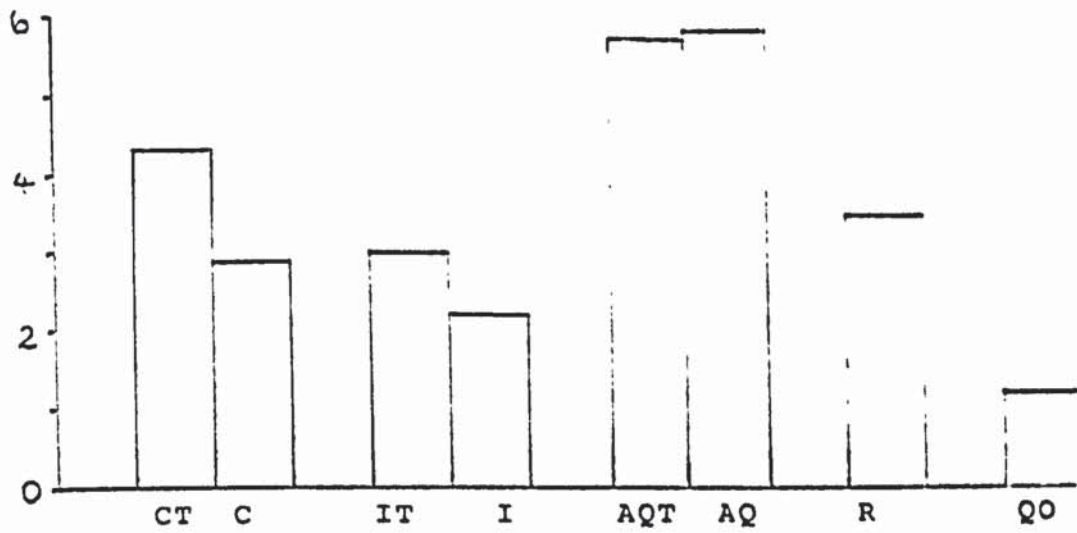


Figure 12 - Mean scores for each group, adjunct questions on delayed test.

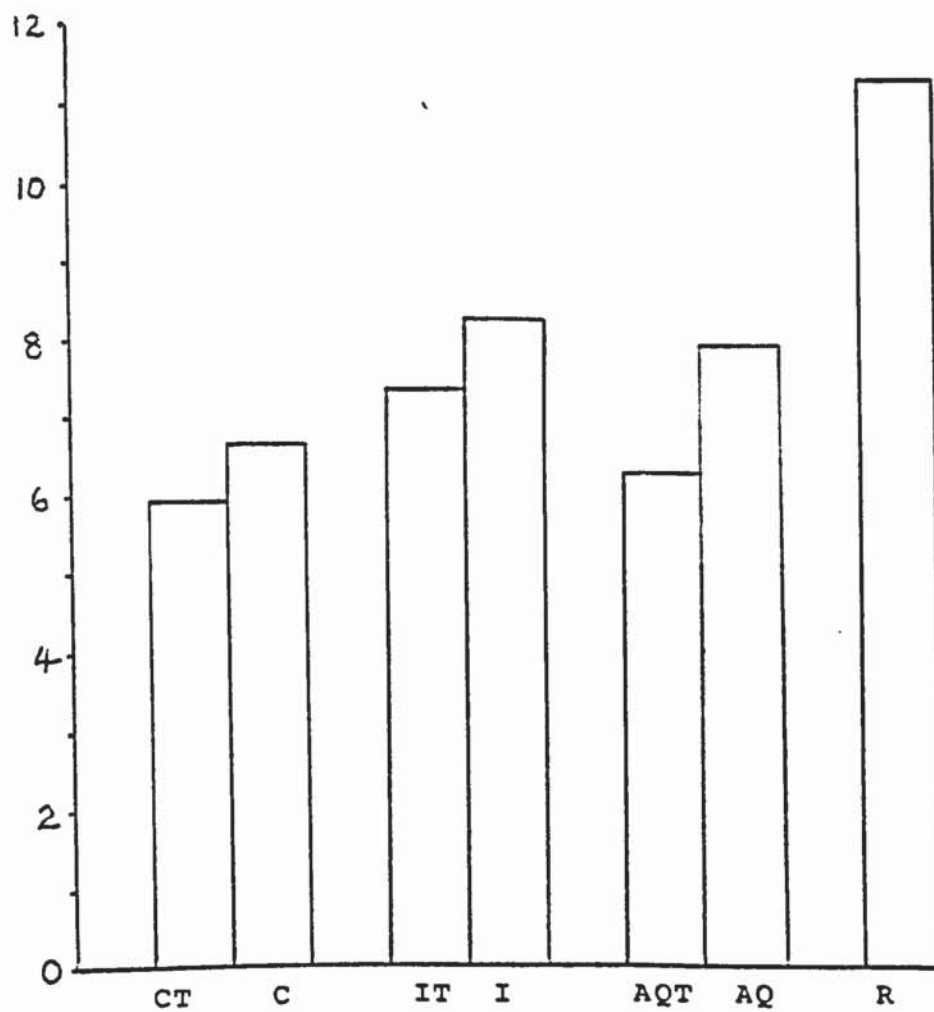


Figure 13 - Mean reading times in minutes for each group.

The small group of 18 subjects given only the questions without ever seeing the passage, to give an idea of their "passage dependency", and to act as a base-line, scored a mean of 0.4 out of 14 on the literal questions, 1.2 on the adjunct questions, and 2.4 on the inference questions, representing scores of 2.8%, 8.5% and 16.1% respectively. This last score is rather high, and analysis showed that all but two of the 18 subjects had correctly guessed the answer to inference question 12 ("How did Joe know that many wagons had already passed that way?"). If this question is excluded, their mean score would drop to 0.7 out of 12 (5.8%) so that, with this one exception, it would seem that the questions were reasonably high in passage dependency, and certainly more so than the "passage dependent" questions used by Tuinman and Farr (1972) whose subjects managed to guess the answers to just over 20% of the questions without seeing the passage (as opposed to 58% on their passage-independent questions).

The rest of the analysis was similar to that used in the vocabulary study. Analysis of variance was used to examine the scores of each group on each type of question, and the results are shown in tables 33 to 37.

The results are clear: in every test, there were significant differences between the various group means at well beyond the .01 level. However, two important points must be made. First, the results of the analysis of the two immediate tests and the adjunct questions could well have been boosted by the inclusion of the Questions Only group, even though there are still highly signifi-

Condition	Mean Score	Standard Deviation
Control, and text	10.7	2.8
Control, no text	4.7	2.5
Im and text	10.0	2.6
Im, no text	5.8	2.2
AQ and text	8.8	3.5
AQ, no text	4.8	3.2
Repeat	4.9	3.6
Test only	0.4	1.0

Source	s.s.	d.f.	m.s.	F	sig.
Between groups	1766.0	7	252.3	29.1	.01
Within groups	1595.5	184	8.7		
Total	3361.5	191			

Table 33 Scores on immediate literal comprehension questions
by condition, with analysis of variance, all groups.

Condition	Mean Score	Standard Deviation
Control and text	8.0	3.5
Control, no text	4.9	2.4
Im and text	8.9	3.2
Im, no text	5.8	3.3
AQ and text	7.3	3.5
AQ, no text	4.9	2.8
Repeat	4.7	2.5
Test only	2.4	1.2

Source	s.s.	d.f.	m.s.	F	sig.
Between groups	699.7	7	99.9	11.5	.01
Within groups	1593.2	184	8.7		
Total	2292.9	191			

Table 34 Scores on immediate inference comprehension questions by condition, with analysis of variance, all groups.

Condition	Mean Score	Standard Deviation
Control and text	4.3	1.8
Control, no text	2.9	1.7
Im and text	3.0	1.7
Im, no text	2.2	1.7
AQ and text	5.7	2.1
AQ, no text	5.8	2.8
Repeat	3.5	1.8
Test only	1.2	1.2

Source	s.s.	d.f.	m.s.	F	sig.
Between groups	397.9	7	56.8	15.2	.01
Within groups	640.8	171	3.7		
Total	1038.6	178			

Table 35 Scores on delayed adjunct questions by condition,
with analysis of variance, all groups.

Condition	Mean Score	Standard Deviation
Control and text	7.6	3.1
Control, no text	4.7	2.2
Im and text	7.0	2.4
Im, no text	5.0	3.0
AQ and text	6.5	3.0
AQ, no text	3.3	2.1
Repeat	4.6	3.4

Source	s.s.	d.f.	m.s.	F	sig.
Between groups	306.5	6	51.1	6.7	.01
Within groups	1177.4	154	7.6		
Total	1484.0	160			

Table 36 Scores on delayed literal comprehension questions by condition, with analysis of variance, all groups.

Condition	Mean Score	Standard Deviation
Control and text	5.6	3.3
Control, no text	3.8	1.8
Im and text	6.8	2.3
Im, no text	4.6	2.6
AQ and text	5.7	2.5
AQ, no text	4.0	2.0
Repeat	3.8	2.4

Source	s.s.	d.f.	m.s.	F	sig.
Between groups	177.9	6	29.7	5.1	.01
Within groups	899.2	154	5.8		
Total	1077.1	160			

Table 37 Scores on delayed inference questions by condition,
with analysis of variance, all groups.

cant differences in the analysis of the other two delayed tests, which the Questions Only group obviously did not re-take. An even more important point is that each condition in fact involved two variables - the instructional set given, and the presence or absence of the text while answering the immediate test questions, either or both of which could have caused the different mean scores. (NB All groups actually answered the delayed questions without the text present, but half of them had previously answered the immediate questions with the text available, as it normally would be in a lesson). Because of these two points, further analysis of variance was carried out on the six main groups only - Control, Image and Adjunct Questions, each with and without text originally available. Tables 38-42 show the results: in each case "Text presence" refers to the presence or absence of the text during the two immediate tests, and "Condition" refers to the use of Control, Image, or Adjunct Question instructions before initial reading.

Source	s.s.	d.f.	m.s.	F	sig.
Text presence	634.8	1	634.8	76.5	.01
Condition	31.3	2	15.7	1.9	ns
Interaction	21.9	2	11.0	1.3	ns
Between groups	688.0	5	137.6	16.6	.01
Within groups	947.7	114	8.3		
Total	1635.7	119			

Table 38 Further analysis of variance, main groups, immediate literal comprehension questions.

Source	s.s.	d.f.	m.s.	F	sig.
Text presence	235.3	1	235.3	23.5	.01
Condition	42.3	2	21.2	2.1	ns
Interaction	9.1	2	4.6	0.5	ns
Between groups	286.7	5	57.3	5.7	.01
Within groups	1137.0	114	10.0		
Total	1423.7	119			

Table 39 Further analysis of variance, main groups, immediate inference comprehension questions.

Source	s.s.	d.f.	m.s.	F	sig.
Text presence	9.1	1	9.1	2.2	ns
Condition	216.4	2	108.2	26.4	.01
Interaction	18.5	2	9.3	2.3	ns
Between groups	244.0	5	48.8	11.9	.01
Within groups	466.9	114	4.1		
Total	710.9	119			

Table 40 Further analysis of variance, main groups, delayed adjunct questions.

Source	s.s.	d.f.	m.s.	F	sig.
Text presence	202.8	1	202.8	29.8	.01
Condition	28.8	2	14.4	2.1	ns
Interaction	8.4	2	4.2	0.6	ns
Between groups	240.2	5	48.0	7.1	.01
Within groups	780.2	114	6.8		
Total	1020.4	119			

Table 41 Further analysis of variance, main groups, delayed literal comprehension questions.

Source	s.s.	d.f.	m.s.	F	sig.
Text presence	91.9	1	91.9	16.4	.01
Condition	21.0	2	10.5	1.9	ns
Interaction	3.1	2	1.6	0.3	ns
Between groups	116.0	5	23.2	4.1	.01
Within groups	642.6	114	5.6		
Total	758.6	119			

Table 42 Further analysis of variance, main groups, delayed inference comprehension questions.

On both immediate and delayed tests using both literal and inference questions, the significant differences between the group means remain, but the only significant factor is the original presence or absence of the text while answering the questions immediately after reading. On the adjunct questions given a week after reading, however, a different pattern emerges (Table 40). Here, the significant factor is the treatment condition, and it is clear from Table 35 and Figure 12 that the Adjunct Question groups easily out-performed the rest here.

The final part of this stage of the analysis looks at the reading times, since it had always been considered that differences between the various treatment conditions could be at least partly due to the different amounts of time each group spent reading the passage, given their treatment instructions and advance warning that they either would or would not be allowed to refer back to the text when answering the test questions. Consequently, analysis of variance was carried out on the reading times of all the Control, Image and Adjunct Question groups, suggesting that there were significant differences present, which were further analysed for possible interaction. Table 43 shows the results.

Source	s.s.	d.f.	m.s.	F	sig.
Text presence	29.1	1	29.1	10.0	.01
Condition	55.2	2	27.6	9.5	.01
Interaction	4.4	2	2.2	0.8	ns
Between groups	88.7	5	17.7	6.1	.01
Within groups	324.9	114	2.9		
Total	413.6	119			

Table 43 Analysis of variance-reading times of main groups.

The results suggest that the control, image or adjunct question treatments, and the knowledge that the text would or would not be available, both had a significant effect on reading time, though there was no significant interaction between them. In fact, as figure 13 suggests, reading time was always longer if the children knew that the text would not be available when it came to answering the questions, regardless of the treatment instructions, although the difference was significant only between the two Adjunct Question groups ($t = 4.9$, with 38 d.f., $p < .01$) and not for the Control or Image groups ($t = 1.3$ with 38 d.f., $p > .05$, and $t = 1.6$ with 38 d.f., $p > .05$ respectively). Separate t tests were also used to compare the treatment conditions, showing that imagery instructions had resulted in significantly longer reading times than those of the control groups, both with

and without text ($t = 3.0$ with 38 d.f., $p < .01$, and $t = 2.7$ with 38 d.f., $p < .01$ respectively), but not significantly longer than the AQ groups. In turn, the AQ group reading times were also longer than those of the control group, but the difference was significant only in the "no text" condition ($t=2.5$ with 38 d.f., $p < .05$).

Discussion. The results of the first stage of analysis, at the level of group comparisons, were not really as had been expected. Although there were significant differences between the group means on all five tests, further analysis showed that they were almost entirely due to the availability of the text, and that the different treatment conditions had little effect. Moreover, the facilitation produced by text availability persisted on the delayed tests, probably because of the superior rehearsal it offered (cf. Yuille 1973). The results thus strongly disagreed with those of Tuinman and Farr (1972), who found no effect of letting 14 year-olds look at the passage or not as they answered questions. However, Tuinman and Farr's "long" passages ranged in length from only 166 to 356 words, with a mean of 262, and they tested only immediate recall, so that their experimental task was perhaps not very relevant to actual classroom demands. If this is the case, then the results of the present study suggest that there are important differences between typical tests of comprehension used in school and those used in the laboratory, seriously questioning the more direct claims of educational implications that have been made by a number of researchers.

Immediate recall may well be a psychologically valid test of some aspects of comprehension, but its educational relevance is indirect rather than direct, and the various factors which have been thought to affect children's passage comprehension, and proposed techniques for improving it, perhaps need to be reassessed: variables like passage dependency or concreteness may be far less important if comprehension is tested as it is in schools, and the previous discussion in chapter four suggested that the cognitive demands of the two tasks may be very different (e.g. Aaronson, 1976).

The only time when treatment condition had a significant effect was in the AQ test, where children who had already seen the questions answered them significantly better than children who had not, even a week later. This result agrees with the general trend of AQ research, and extends the finding of specific facilitation of "old" items to children's delayed recall, though such results are open to the criticism that they are simply due to rehearsal effects (cf Yuille, 1973). Interestingly, text availability made no difference to the two AQ groups' scores here - the AQ group with text in fact scored 5.7 to the AQ without text's 5.8, but both were clearly superior to even the highest-scoring of the rest of the groups ($t = 2.5$ with 42 d.f., $p < .05$). However, the other results disagree with the usual research finding that postquestions also facilitate scores on "new" items, and definitely agree with what appears to be the only comparable experiment using children - Swenson and Kulhavy's (1974) finding that such adjunct questions did not lead to any general facilitation. In fact, the two AQ groups performed worse than

the corresponding imagery groups on all other tests, and worse than the corresponding control groups on all tests except the delayed inference questions, where their superiority was negligible (5.7 as against 5.6 for the groups with text, and 4.0 as against 3.8 for the groups without text). Differences were often small, but the conclusion is clear: children given adjunct questions while reading did not perform better than children given imagery instructions or left to their own devices on "new" questions, whether immediate or delayed, literal or inferential.

This disagreement with other research could be because all subjects here were allowed to refer back to the text in all conditions, or because this study used children rather than adults. It is difficult to choose between these two possibilities, however, because of the nature of the few other studies which also found no general facilitation with postquestions. On the one hand, Washburne (1929) and Hiller (1974) allowed back-referring during reading, but both used student subjects, while on the other hand, Swenson and Kulhavy (1974) did not allow back-referring, but used children as subjects. The explanation could perhaps be due to some other factor, but if the use of either back-referring or children was involved, the present results again cast doubts on the direct educational relevance of much AQ research: using children given a more "real" task, postquestions appeared to have no general facilitative effect whatsoever.

Turning to the imagery groups, the overall results are again disappointing, especially in view of the large effects produced by "explicit" prompts using imagery in the vocabulary experiment,

and the suggestions in the research literature that simple imagery instructions should be enough to improve the comprehension scores of 11 year-olds. Closer examination of the group means showed that the two imagery groups had always scored better than the equivalent control groups on the inference questions, and in the "no text" condition on both the immediate and delayed verbatim tests, suggesting that imagery might have been more useful for the more difficult test conditions. This optimistic suggestion was tried out by separate t-tests, but not one of the differences between control and image groups was significant, whether with or without the text available, on any of the five tests. These results therefore do not support the suggestion that imagery might be particularly useful in inferential comprehension (e.g. Paris, Mahoney & Buckhalt, 1974). In fact, it seems that in this case image instructions, like the use of adjunct questions, did not result in significantly better scores than leaving the children to their own devices, on any of the tests.

In view of these findings, it is difficult to believe that any group differences in reading time alone could have had a direct effect on scores. Imagery instructions did result in significantly longer reading times than control group instructions, with no significant gain in scores, so that imagery could be seen as an "inefficient" strategy, as could the use of adjunct questions (cf. Carver, 1972). However, reading time alone may have little causal significance because the results of the

"repeat" group told to read each page twice were almost identical to those of the control group on both immediate and delayed, literal and inference questions (where all differences were only 0.2 or less), and rather better, though not significantly so, on only the adjunct questions (scores are shown in tables 33-37). Yet their mean reading time was 11.2 minutes, as opposed to 6.6 minutes in the control group, and though this is, interestingly, not twice as long, it means that these children clearly spent far longer reading the passage, with no significant improvement as a result. This point could be verified by a study controlling reading time, and exposing one group to the materials for much longer than another, but there would be no guarantee that the "short" time had been adequate or that the children had spent all the "long" time reading. The present procedure had the advantage of letting the children read for as long as they wished, and it strongly suggests that reading time alone is a less important causal variable than Carver proposes, though if a child spends longer on a passage because instructions have made him adopt different strategies, increased reading times could then be involved in any improvement (cf. Rothkopf, 1974).

As for sex differences, when boys' and girls' mean scores are considered separately for each of the five sets of questions answered by the six main groups, it appears that the girls had better mean scores than the boys on 15 of the sets of results, and worse on 15, and no clear trend was apparent with regard to the types of test at which either appeared to be better. Similarly

the overall mean score was 5.6 for the boys, and 5.7 for the girls, so that sex differences do not appear to have been very important, possibly because the topic was of interest to the boys (cf. Asher & Markell, 1974).

One final point is that, as with any study reporting non-significant differences between groups, the explanation could simply be that the tests were not sensitive enough to distinguish between the treatments. However, examination of individual scores shows that they range from 0-14 (out of 14) on the literal questions, and 0-13 on the inference questions, with reasonably large standard deviations suggesting roughly normal distribution (tables 33, 34, 36 and 37), and these questions also clearly distinguished between the text available and unavailable conditions. The adjunct questions were used only on the obviously more difficult unexpected delayed recall test, with no text available, so scores range only from 0-11, but these questions clearly distinguished between the AQ groups and the rest. Passage dependency, assessed by means of the "Questions only" group, was high for all three types of question, so that the answers could not easily be guessed (with the exception of the one inference item, question 12, discussed earlier). It would thus appear that the questions were generally able to discriminate between individual subjects and between at least some of the treatment conditions, and test scores did also correlate significantly with reading age when the treatment groups were considered separately.

In conclusion, the analysis at the level of group mean scores strongly suggests that the most important variable was text availability, with all the implications for claims of direct educational relevance made by previous research. Imagery instructions, adjunct questions, or simply spending a lot longer reading the text did not lead to scores significantly better than those of the equivalent control groups. The next stage of analysis therefore looks more closely at the scores in the light of the children's reports about what they actually did.

(ii) Reported Strategies

Analysis of the control groups' reported strategies offers an immediate, and perhaps even surprising, possible explanation of the imagery and adjunct question groups' disappointing results. Although the children had reported using imagery spontaneously less than 7% of the time in the vocabulary study (26 out of 52 of them never reported using it, and of the rest only 5 said they did so for more than 2 of the 17 words), a very different picture emerged when these very same children described their reading strategies. In the control group who knew they would be allowed to retain the text, 13 out of 24 subjects said they had tried to picture the story in their minds as they read it, and in the "no text" condition, 19 out of 28 reported using imagery. A chi-square test suggested that the proportion of subjects reporting imagery was not significantly different in the two conditions, but significant differences were found when each group was sub-divided into subjects scoring above and below the median (combined test scores), as in table 44. Since the

frequencies in some cells were very small, Fisher exact probability tests were used, and showed that significantly more subjects scoring above the median had reported using imagery, in both conditions of text availability ($p < .05$, with text; $p < .01$ without text).

		Scored above median	Scored below median
Control group with text	Reported imagery	8	2
	No imagery	3	7
Control group, no text	Reported imagery	11	4
	No imagery	1	7

Table 44 Reported use of imagery, and test scores, control groups.

When the children's reports that they "guessed at the questions" are similarly analysed, an interesting difference emerges. In the group who knew they would be allowed to retain the text, 11 out of 24 children reported guessing at the questions, but in the group who knew the text would not be available, the proportion rises significantly to 22 out of 28 ($\chi^2 = 6.0$, $p < .05$). Table 45 shows the results when both control groups were sub-divided as before, but Fisher exact probability tests showed that the proportion of those guessing at the questions did not differ significantly according to scores, for either condition of text availability.

		Scored above median	Scored below median
Control group, with text	Reported guessing	7	3
	No guessing	4	6
Control group, no text	Reported guessing	11	8
	No guessing	1	3

Table 45 Reported guessing at questions, and test scores, control groups.

The third suggested possible strategy had been to re-read parts (or all) of the passage, and reports of doing this were similarly analysed. With the text available only 5 out of 24 subjects reported re-reading, and without it, 13 out of 28 did so, though the difference was not significant at the 5% level ($\chi^2 = 2.7$). Table 46 shows the sub-groups' results, but again the differences were not significant for either control group.

		Scored above median	Scored below median
Control group, with text	Reported re-reading	3	2
	No re-reading	8	7
Control group, no text	Reported re-reading	8	4
	No re-reading	4	7

Table 46 Reported re-reading, and test scores, control groups.

In sum, the three spontaneous strategies appear rather different: spontaneous use of imagery was not affected by text availability, but was related to superior test scores when it was used; in contrast, guessing at the questions was significantly affected by text availability, but was not related to better scores; re-reading was neither affected by text availability nor related to better performance.

Reports of "don't know" (i.e. "You don't know what you did, just read the story and answered the questions") were very rare - only 4 children out of 52 used this response, and other responses were also rare - a few children said they had stopped at the end of each page, and one boy in the "no-text" group (the fifth highest scorer) said, "Anything interesting I said over and over again to memorise". A few wrote rather vague comments - "I thought about the sentences", for example, but two particularly interesting ones were "I thought I was there", and, from the girl with the highest score overall (54 out of a possible 56 on the immediate and delayed literal and inference questions) - "Tried to remember important things. When I read I sort of see a moving picture behind the words on the page".

The ability to make comments like this last one could well be related to the ability to do well on the tests, so that little weight can be put on such spontaneous reports. However, it does seem from the analysis of high and low scorers' reports that the use of imagery, but not of "guessing questions" and re-reading, was related to test success, so further analysis was carried out only on the specific effects of reported imagery.

Taking the two conditions of text availability separately, subjects reporting using imagery had done better than subjects not reporting imagery on every single test except the adjunct questions in the "no text" condition, (where the difference was not significant.) However, although image users scored better than those who reported none on nine out of the ten sets of scores, their superiority was certainly not always significant: separate t-tests in fact showed that the differences were not significant for any of the literal comprehension tests, or for the "with text" condition on the inference questions. In the "without text" group, those reporting imagery scored significantly higher than those not reporting it, on both immediate and delayed inference tests ($t = 3.3$ with 26 d.f., $p < .01$; $t = 1.9$ with 25 d.f., $p < .05$ respectively). So although the frequency results had indicated a significant general trend, the actual superiority in the scores of imagery users was significant only on the more difficult inference questions, under the more difficult "without text" conditions. This result gives some support to the suggestion made by Sheehan (1971, 1972) that imagery might be particularly useful in an unexpected recall task.

The imagery-instructed groups had also been asked how much they had actually used imagery, on a five-point scale ranging from "All the time" to "not at all", with other categories roughly related to the number of pages for which they had formed images as instructed. There are obvious problems with such

subjective ratings, including a reluctance to use either of the extreme categories, especially the one suggesting that they had completely failed to follow instructions, but some interesting results do emerge. Taking the Image group, with text, 19 out of 28 subjects said they had used images all or most of the time, 7 some of the time, and 2 little or none. When they are split into two groups of the 19 who tended to use imagery a lot, and the 9 who did not, the former had superior scores on all five sets of questions, and their total mean score of 40.5 was significantly better than the 29.4 scored by the other sub-group ($t = 3.9$ with 21 d.f. $p < .001$). Reported use of imagery in the Image, no text group, showed a rather different pattern, with most subjects (13 out of 24) saying they formed images "some of the time", and the rest equally divided between "all or most" and "little or none" of the time. Mean total scores for these three sub-groups increase as the reported use of imagery increases - 12.5 in the "little or none" group, 22.9 in the middle group, and 28.7 for those who reported most imagery.

The final stage of the analysis involved a return to group comparisons. Previous results had suggested that spontaneous "guessing at the questions" or re-reading were not related to test success, but that spontaneous imagery was. A comparison was therefore made between control group subjects who had not reported spontaneous imagery, and imagery-instructed subjects who had followed the instructions to form images for every page (or

at least, most pages). Table 47 shows the results for the two conditions of text availability, with levels of significance for the difference between Imagery and control group subjects in each case, as shown by separate t-tests.

		Immediate Tests		Delayed Tests		
		Literal	Infer.	Literal	Infer.	AQ
With Text	S's following imagery instructions	10.5	10.6	8.1	7.5	3.8
	Control group, non-imagers	10.0	7.2	7.4	5.4	4.1
	Significance of difference	n.s.	.01	n.s.	.05	n.s.
No Text	S's following imagery instructions	7.3	7.7	6.0	6.0	1.7
	Control group, non-imagers	3.9	3.2	4.0	2.9	3.6
	Significance of difference	.05	.01	n.s.	.05	n.s.

Table 47 Differences between control and imagery groups' means scores, according to reported strategies.

(g) General Discussion. The children's reported strategies thus offer a possible explanation for the lack of significant differences between control and experimental groups as a whole, at least in the case of imagery, so that when children in the imagery instructed groups who followed the instructions are compared with control group subjects who did not use imagery spontaneously anyway, some significant differences do appear, particularly on the inferential questions, thus lending some support to the earlier suggestion that images might be particularly beneficial at this level of comprehension. (e.g. Paris, Mahoney & Buckhalt, 1974; Rohwer & Matz, 1975; Aaronson, 1976). However, there is the important qualification, also discussed in chapter three, that this evidence is merely correlational in nature. Children who follow imagery instructions may do better than children who do not use imagery spontaneously for a variety of reasons, and the attempt to manipulate the variables by group instructions proved much less satisfactory than in the vocabulary experiment, where the highly significant group differences meant that the causal importance of imagery could be more confidently inferred.

Text availability while answering the questions emerged throughout as a highly influential variable, emphasising the difference between comprehension as it is tested in the classroom and in the laboratory. Not only did having the text available lead to clearly superior scores, as one might expect in spite of Tuinman and Farr's (1972) results, but it had different effects on

the strategies which the children reported: their use of imagery and re-reading were not significantly affected, but they tended to spend longer reading the text and guessed more at what the questions might be, so that the two tasks can be seen as cognitively rather different. Another way of looking at this could be to say that forming images and re-reading are general strategies that children tend to use (or not to use) regardless of test conditions, while guessing at the questions is a more specific strategy which is useful only under certain conditions.

Most of the other points for discussion have already been covered, but there are certainly several possible weaknesses in the study, some of which might be resolved by further research. The widespread use of spontaneous strategies by these subjects was almost certainly involved in the lack of differences in group scores, but the same might not be the case for younger children who, at least according to Rohwer's hypothesis discussed above, might be more likely to benefit from imagery instructions. However, it might be advisable to check on the prose comprehension and learning demands actually made of, say, 8 year-olds before trying to evaluate techniques for improving them. Apart from using younger subjects, different levels of prompt could also be used, and in particular, the effects of augmented prompts in the form of pictures could be further investigated: would they be as useful in answering questions with the text present as they appear to be in the widely-reported research using recall tests, for example? If not, the direct educational relevance of this

research can again be questioned.

Another major problem, in common with all research into unobservable strategies, is the need to rely either on inferences made by the experimenter, or on introspective reports made by the subjects, which may be even more unreliable with children. On the other hand, significant observable differences in test scores did appear to be related to the reports that were obtained, and helped to throw some light on the common finding that imagery instructions, for example, generally provide only slight facilitation in prose learning and comprehension (Pressley, 1977). Perhaps more experiments could actually ask children how they go about a variety of tasks, instead of simply assuming that they follow mediation instructions completely, and the present study is very much in agreement with Anderson and Kulhavy's (1972) finding that student subjects frequently failed to use imagery as instructed, or used it spontaneously anyway, resulting in non-significant differences between control and experimental groups, so that it would now appear likely that their findings can be extended to normal 11 year-olds. A further extension would be the collection of more detailed information on the strategies used by all subjects, to find out, for example, if AQ subjects had ignored the questions, or had used imagery, or if students in the "Repeat" group had followed instructions, or used a different kind of "read" each time (cf. Augstein & Thomas, 1973). Subjects could also be asked how much they had used various strategies, as they were in the imagery group, perhaps using some sort of grid to be filled in, though care would

have to be taken that this did not become too complex for children to use.

A slightly different approach would be the use of various techniques to ensure that children followed instructions, but there are problems here. An overt response to be filled in for AQ groups, as used by Swenson and Kulhavy (1974), for example, does not ensure that children process the text more thoroughly to begin with, unless one adds the artificial condition of not allowing back-referring, and there is also the problem of what to do about wrong answers unless the questions are made very easy, in which case the reader will probably benefit little from them (Anderson, 1970). An interesting technique developed by Riding and Taylor (1976) tested children's use of imagery as they read by asking them questions whose answers had not actually been stated in the text, but which could be answered if the child was forming images. For example, given the sentence "Giacco knocked at the door", children would be asked, "What colour was the door?" However, their suggestion that the use of "verbal" strategies could be examined in a similar way is less satisfactory: given "The little boy saw a cat chase the mouse", the suggested question "Is a cat an animal?" could be answered without the passage having been attended to at all, and both sorts of questions involve the problems of being totally lacking in passage dependency. The problems of checking on use of various strategies still remain to be solved, though the technique of inserting adjunct questions to test imagery is obviously very relevant to the present study.

Another area which could be more thoroughly examined is the relationship between strategies adopted in compliance with instructions, and spontaneous strategies. These might be quite similar in the case of imagery, where only a simple instruction was provided anyway, but in the case of adjunct questions, it cannot be assumed that the spontaneous strategy of guessing what questions might be asked is cognitively similar to the processing brought about by actual adjunct questions, which may work only when the subject knows he cannot refer back to the passage and not in the more "normal" reading task used here (Carver, 1972). In the same way, there may be only limited similarities between subjects' spontaneous re-reading and their behaviour when instructed to read the passage twice, and again this could perhaps be examined by further research.

Carver's other objection to much of the research, that reading time is the crucial factor, has been at least partly answered. It appears that increased reading time alone is not in itself a cause of superior performance, or the results of the "Repeat" group would have been much better. In fact, it is more tempting, in view of the fact that their results were almost identical to those of the control group, to see simple re-reading as similar to the shallow processing of rote repetition discussed in the vocabulary experiment. What surely matters is what the reader does during the time: for example, it is perhaps significant that taking two girls with identical reading times in the control group, both almost the first to finish, one had easily the highest total test score, and the other easily the lowest. Further

research could relate the time taken to do "real" school tasks to cognitive styles, particularly Kagan's "reflective" and "impulsive" dimension(e.g. Kagan, 1971), as mentioned earlier. (In fact, in the present study, control group subjects with high reading ages took less time to read the passage, though the difference was significant only in the "no text" condition).

One final problem, in common with most research, is the degree to which the results can be generalised. It was felt that the sample, covering a wide range of ability and social background, as described in chapter three, was reasonably typical of children in their first year at secondary school, but different results might well be obtained with younger or retarded children who lack spontaneous strategies. Ideally, the findings should also be tested with a variety of materials, and a compromise would have to be worked out between using "real" texts and the value of being able to manipulate variables like passage concreteness. Similarly, a wider variety of both test and adjunct questions could be used (cf. Hiller, 1974), especially since much previous research has usually used only paraphrase questions to test the "deeper" levels of comprehension. This study certainly makes no claims to have sorted out the complex variables involved in comprehension, but it does suggest the potential dangers in generalising from experiments using only recall tests, and highlights the value of studying children's reported strategies when examining the effects of various techniques for trying to improve their passage comprehension and learning.

CHAPTER VI

CONCLUSIONS

The final chapter briefly sums up the main conclusions that can be drawn from the study and suggests some underlying themes relating the two experiments to each other and to the wider contexts of pure and applied research.

Taking the questionnaire first, the possible criticism of sample bias was at least partly answered in the second chapter, and though further studies could obviously involve a much bigger sample, it was felt that the method adopted did allow children from a number of schools to be asked, outside school yet still in reasonably controlled circumstances, about their perceptions of what went on in their French and English lessons. This latter point is important since it is possible that the children tended to overestimate the frequency of tasks they did not enjoy, or not to recognise when help in learning techniques was being given, and it would certainly be interesting to compare the teachers' own responses with those given by the children they taught.

In spite of these criticisms, however, the results were quite definite: both vocabulary learning and

comprehension work were reported as being widespread, though little advice appears to be given as to how they should be carried out, and each still takes a rather "traditional" form. Most children are still given vocabulary tests of individual words and use books of isolated comprehension exercises, in spite of the frequent opposition to such activities from those who write in the specialist subject journals. In the case of vocabulary learning, such criticisms can be countered by recent research on strategies which strongly suggests that native language equivalents are still going to be covertly attached to new vocabulary by learners and that attempts to avoid this happening by changing the "nominal stimulus" by using pictures and avoiding overt native language may only lead to more problems, as some of the more recent articles are beginning to agree. In the case of comprehension work, opposition comes from supporters of a "personal growth" rather than a "skills" model of English, but is tempered by the recognition that comprehension in the widest sense is a central responsibility of all subject teachers, so that disagreement tends to be centred on the form the work should take rather than on its actual inclusion in the curriculum. Both debates continue, but the questionnaire strongly suggested that many teachers,

rightly or wrongly, still set both sorts of work and, if nothing else, this exploratory survey did enable the two experiments to be related to actual classroom practice.

In the vocabulary experiment itself, the provision of a mnemonic strategy proved very effective, particularly if an augmented prompt in the form of an interactive line drawing was also included, and the significant group differences allowed a causal relationship to be inferred, going beyond the equally significant correlational data of the success of the children's reported strategies. These reports strongly suggested that successful and unsuccessful children approached the same task rather differently, and that the differences could not be attributed to the superior ability of the high scorers to remember or describe what strategies they had used. The keyword method thus appears to be highly effective, not just with students learning the English equivalents of carefully selected items, but with children, forming their own images as well as using provided pictures, learning "real" vocabulary, both French into English and English into French, on both immediate and delayed tests. However, there are obviously several weaknesses in the experiment, some of which could be investigated by

further research.

One criticism is that facilitation could simply have been due to the extra cues involved in the keyword groups' presentation. However, care was taken that the Rote group heard each pair just as often as the keyword groups did, and the possible value of extra cues is seriously undermined by Pressley (1977) who exposed control group subjects to pictures or keywords without any instruction in the method, and with no significant improvement in performance. Another possible factor could be the novelty effects of the method, and although an explanation entirely in terms of motivation or subjects' enjoyment was discussed and found inadequate, further research, perhaps using the keyword method over an extended period, could see what happened as the novelty wore off. A longitudinal approach could also be used to examine another very important question - could children eventually be taught to generate both acoustic and image links for themselves? This is particularly important in view of the feeling often expressed in the literature that subject-generated mediators are preferable to experimenter-provided ones, though previous discussion in Chapter Two did suggest that this was not always the

case.

Further research could also examine what sorts of words are best suited to the method, since the Stanford researchers, who have done most work in the area, feel that the selection of the most suitable vocabulary and keywords is still empirical rather than theoretical (e.g. Atkinson, 1975). However, it does appear that the acoustic and image links are independent, and that if each is learned separately by different groups, the probability of learning one correctly, multiplied by the probability of learning the other, gives a good indication of the success rate of the keyword method on that item (Atkinson and Raugh, 1975).

Although it was not a major concern of the present study, where the words were deliberately not specially selected, some analysis was in fact attempted by comparing the items found easiest and hardest by the Image group. The results are shown in Appendix G. Thorndike-Lorge frequency values were available for all the English words and all but one of the keywords, but no clear differences emerged between easy and difficult items. Similarly, Image values (Van Der Veur, 1975) were available, at least for all but one of the English words,

but a Mann-Whitney U test showed that they did not differ significantly between easy and difficult items. It had also been considered, in view of the discussion in the third chapter, that the bizarreness of the image link might be a factor, and so six adult judges rated the image descriptions given to the children (see Appendix C) in random order, using a simple four point scale ranging from "Really ridiculous or bizarre" to "Normal, ordinary". Mean ratings are also shown in Appendix G, but a Mann-Whitney U test again showed that easy and difficult items did not differ significantly in rated bizarreness. The question of what makes a good keyword link thus needs further research in which several possible variables could be manipulated, but what does emerge is that it almost certainly depends on a combination of variables affecting both the English word and the keyword, as well as the link between them, and for the time being, Anderson's (1975) recommendation of using a group of people to select them remains a useful, practical proposal. It should also be remembered that "easy" and "difficult" are relative terms, and that even the comparatively poor keywords achieved a success rate similar to that of the words which the children in the control group learned best of all.

Finally, further research could look at the possibility of extending the use of mnemonics to other areas of the curriculum. It is quite possible that vocabulary learning is a very unusual activity, and that mnemonics would be of little use elsewhere, but further research should perhaps check on the actual classroom demands being made of children before dismissing mnemonics as amusing tricks, with very limited usefulness. Bower's claims (1973a, b) that we still do expect children to master a considerable amount of arbitrary information would mean that mnemonics might be more widely used, and the questionnaire data suggested that it may well not be enough simply to consult the specialist subject journals to find out what actually goes on in classrooms. New materials should obviously be taught in a meaningful way whenever possible, but if some information is inherently and unavoidably arbitrary in nature, then mnemonics offer a way of extending the principles of meaningful learning and avoiding the shallow and ineffective processing of rote repetition.

Turning to the passage comprehension and learning study, the major weaknesses and some corresponding suggestions for further research were discussed at the

end of the previous chapter, so that little needs to be added here. Previous criticisms, particularly of much of the AQ research, were considered in the design of the experiment, and the questionnaire data also allowed at least some of the conditions to be similar to those found in schools. The use of an actual text rather than a specially-constructed one, the subjects' freedom to refer back to the passage and to take as long over reading as they wished, coupled with the use of short-answer rather than multiple choice items and the "with-text" condition as well as the usual recall test, all combined to make the present experiment hopefully more "real", while giving ample reasons for its various disagreements with previous studies. Several weaknesses were suggested, however, and further research could certainly involve a much wider range of subjects, passages, adjunct and test questions, as well as perhaps checking more thoroughly on the strategies reported by all subjects.

On the more positive side, the experiment did suggest that text availability is a vital factor, affecting both the strategies subjects are likely to use, and their scores even on delayed recall tests, where the original test probably gave them more opportunity to find and rehearse

the correct answers (cf. Yuille, 1973). The implications of this finding for the claims of direct educational relevance made by experimenters using recall tests to study comprehension have already been discussed. The results also disagreed with the common AQ research finding that postquestions facilitate performance on "new" test items as well as "old" ones, though this disagreement could be either because of the freedom to refer back to the text, or the use of children as subjects. Another disagreement was with Carver's (1972) proposal that reading time is an important variable in its own right, since the results of the "Repeat" group were almost identical to the equivalent control group, and it was felt that while reading time might be affected by adopted strategies, it was not alone an adequate explanation for success or failure in the tests. Finally, the study again emphasised the value of asking subjects about what they actually did, suggesting not only that their various strategies differed in effectiveness and were differently affected by test conditions, but also offering an explanation for the small differences in group scores (cf. Pressley, 1977), and showing that the use of imagery at least was related to passage comprehension and learning.

Comparing the two experiments, the much less striking results of the comprehension study could be due to a number of factors, though a likely explanation can be offered in terms of levels of processing and research into strategies and the particular values of imagery, which formed underlying themes throughout. It has already been suggested that one fundamental difference between French and English as school subjects is that in French, the teacher is the possessor of knowledge not shared by his pupils, while the English teacher is expected to begin by accepting the existing language of his pupils as a basis, particularly if he is more interested in their "personal growth" rather than developing their skills. This general difference is reflected in the two specific tasks used here and in the spontaneous strategies reported by the children. In learning from a passage or answering comprehension questions on it, these 11-year-olds reported that they spontaneously used several strategies, and only a small minority (less than 8%) said they did not really know how they went about the task. It is likely that comprehension and prose learning are tasks that they have been used to doing for some time, and the questionnaire showed that even the junior school children regularly did comprehension work,

though further research could enlarge on this point. In contrast, junior school French has been severely criticised and was taught at only one of the junior schools, so that the task of learning French vocabulary was almost certainly quite a new one to most of the children in the experimental sample. This is reflected in their reported strategies - much greater use of the "don't know" category and rote learning, particularly among the less successful learners, and certainly very little spontaneous use of imagery.

It would thus appear that the children's existing strategies for the two experimental tasks were rather different, but it is also possible that, in the case of the "Image" groups, the nature of the provided prompt was not quite the same. Rohwer's (1973) hierarchy of prompts would describe both the simple instruction to form images (as in the comprehension experiment) and the actual description of images (as in the vocabulary experiment) as "explicit prompts", while the pictures would supply "augmented explicit prompts". However, in the vocabulary study a specific prompt was presented for each pair of words, and the image was actually described, whereas in the comprehension study

the children had much more complex materials and were simply told at the beginning to generate images as they read and at the end of every page. Many of them in fact failed to do this, and a large proportion of the control group subjects used some imagery, so that the actual differences between the strategies of control and image groups were slight, and the instructions given to the image group in particular may in fact have been little more use than a "minimally explicit prompt" like being simply told to read the passage for a test.

Flavell (1970) provides a slightly different way of relating the sorts of strategies required in the two experiments, through his description of mnemonic mediation as "a planful cognitive act" which can be either enactive, iconic or symbolic, and which may be general or specific. Imagery has been a common theme throughout and although both tasks were obviously highly verbal, the nature of imagery as discussed in the first chapter means that it can often be a very effective mediator for both visual and verbal materials. However, it might be possible to distinguish between the two tasks and the provided strategies used here by seeing vocabulary learning as demanding highly specific mnemonic

mediators, both verbal (in the acoustic link) and visual (in the image link), which the children were unlikely to generate for themselves because of their lack of experience at the task. In contrast, comprehension passages demand more general strategies like trying to image the main events and looking for underlying themes, possibly in the form of "surrogate structures", so that images may be part of an approach of "general planfulness" in reading any text, rather than a fixed part of a sequence used in a specific technique. In both experiments, the strategy of simply repeating or prolonging the stimulus either by rote repetition in the vocabulary study, or straightforward re-reading in the comprehension study, appeared to be less useful than deeper forms of processing in which the learner actively manipulates or transforms the materials, and imagery emerged as an apparently effective way of doing this. On the other hand, the differences between the spontaneous strategies reported by the children, combined with the rather different cognitive demands and provided strategies in each experiment, do suggest an explanation for the much more dramatic results of the vocabulary study.

Explanations in terms of strategies can obviously

be applied much more generally to children's learning failure and success: Paris, Mahoney and Buckhalt (1974) for example, feel that the "mental storehouse" view of learning has been over-emphasised in research on retardation, while Howe (1976) concludes:-

Where differences are found in measured learning ability it is usually possible to identify differences in what individuals actually do as they attempt to learn... In instances where less successful learners can be helped or encouraged to undertake the kinds of procedures habitually carried out by good learners, their performance may be brought closer. It follows that if we are to help poor learners, a major sub-goal should be to identify and to describe as precisely as we can how successful learners cope with actual learning situations. We can then often help poorer learners by aiding them to acquire procedures and strategies for learning.

Such an attempt both to describe and manipulate learners' strategies was at the basis of the two studies described here, and can be seen as underlying much of the cognitive research described in earlier chapters, but if such research is to be educationally as well as psychologically valid, it must take account of the actual classroom demands made on children as well as the strategies that they adopt. Several writers (e.g. Kintsch, 1974; Phillips, 1976) have suggested that recent cognitive

research has had little direct impact on education, and it may well be that teachers tend understandably to emphasize the observable products of learning at the expense of its processes in the form of pupils' strategies and ways of improving them (e.g. Herritt, 1973; Olson and Bruner, 1974; Ott, Blake and Butler, 1976). Conversely, other writers have accepted that experimenters themselves should be more willing to apply the techniques and findings of basic research to the classroom, not only in special education (e.g. Wambold and Hayden, 1975), the traditional concern of educational psychology, but also to the education of normal children, carrying out the sorts of tasks that are actually expected of them in schools (e.g. Palermo, 1970).

On a more general level, Neisser (1976) sees a widespread failure of cognitive psychology to apply itself to the "real" world, and concludes-

Cognitive psychologists must make a greater effort to understand cognition as it occurs in the ordinary environment and in the context of natural observable activity. This would not mean an end to laboratory experiments, but a commitment to the study of variables that are ecologically important rather than those that are easily manageable.

Cognitive psychology offers several exciting and rapidly expanding areas of research: attempts to look beyond the "nominal stimulus" and to describe and manipulate learners' strategies not only suggest an explanation for the differences between poor and successful learners, but also allow for optimism about narrowing these differences. In spite of disagreements about its nature and functions, imagery, long neglected, does seem to be an important if elusive factor in many verbal tasks, including the two extremes of learning individual words and understanding a long passage. These approaches certainly seem as though they should be very relevant to education, but if Neisser's criticisms are valid, then there is perhaps more room for research which, while retaining the experiment's advantages of control and manipulation, is also aware of the nature and conditions of classroom tasks which children are actually expected to perform.

REFERENCES

- AARONSON, D. Performance Theories for Sentence Coding: Some Qualitative Observations. Journal of Experimental Psychology, 1976, 2, 42-55.
- ABBS, P. English Broadsheets - Extensions, 1970, Heinemann.
- ABERCROMBIE, M.L.J. The Anatomy of Judgement, 1960, Hutchinson.
- AMMON, R. Generating Expectancies to Enhance Comprehension. The Reading Teacher, 1975, 245-249.
- ANDERSON, B.F. Cognitive Psychology, 1975, Academic Press.
- ANDERSON, J.R., and BOWER, G.H. Human Associative Memory, 1973, Winston.
- ANDERSON, R.C. Control of Student Mediating Processes During Verbal Learning and Instruction. Review of Educational Research, 1970, 40, 349-369.
- ANDERSON, R.C. How to Construct Achievement Tests to Assess Comprehension. Review of Educational Research, 1972, 42, 145-170.
- ANDERSON, R.C. Substance Recall of Sentences. Journal of Experimental Psychology, 1974, 26, 530-541.
- ANDERSON, R.C., and BIDDLE, W.B. On Asking People Questions About what They are Reading. In BOWER, G.H. The Psychology of Learning and Motivation, 1975, 9, 89-132.
- ANDERSON, R.C. and HEDDE, J.L. Imagery and Sentence Learning. Journal of Educational Psychology, 1970, 62, 526-530
- ANDERSON, R.C., and MCGAW, B. On the Representation of Meanings of General Terms. Journal of Experimental Psychology, 1973, 101, 301-306.
- ANDRE, T., and SOLA, J. Imagery, Verbatim and Paraphrased Questions and Retention of Meaningful Sentences. Journal of Educational Psychology, 1976, 68, 661-669.
- ANDREOFF, G.R. and YARMEY, A.D. Bizarre Imagery and Associative Learning: a Confirmation. Perceptual and Motor Skills, 1976, 43, 143-148.

- ANTHONY, W.S. Activity in the Learning of Piagetian Operational Thinking. British Journal of Educational Psychology, 1977, 47, 18 - 24.
- ARNHEIM, R. Visual Thinking, 1970, Faber and Faber.
- ARNOLD, D.J., and BROOKS, P.H. Influence of Contextual Organizing Material on Children's Listening Comprehension. Journal of Educational Psychology, 1976, 68, 711 - 716.
- ASHER, J.J., KUSUDO, J.A., and de la TORRE, R. Learning a Second Language through Commands: the Second Field Test. Modern Language Journal, 1974, 58, 24 - 32.
- ASHER, S.R., and MARKELL, R.A. Sex Differences in the Comprehension of High and Low Interest Reading Material. Journal of Educational Psychology, 1974, 66, 680 - 687.
- ATKINSON, R.C. Ingredients for a Theory of Instruction. American Psychologist, 1972, 27, 921 - 931.
- ATKINSON, R.C. Mnemotechnics in Second-Language Learning. American Psychologist, 1975, 30, 821 - 828.
- ATKINSON, R.C. and RAUGH, M.R. An Application of the Mnemonic Keyword Method to the Acquisition of a Russian Vocabulary. Journal of Experimental Psychology, 1975, 104, 126 - 133.
- ATWOOD, G.E. An Experimental Study of Visual Imagination and Memory. Cognitive Psychology, 1971, 2, 290 - 299.
- AUGSTEIN, S. and THOMAS, L. Developing Your Own Reading. In MOYLE, D. (Ed.) Reading Development, 1973, Open University Press.
- AUSUBEL, D.P. Educational Psychology - a Cognitive View, 1968, Holt, Rinehart and Winston.
- BANNISTER, D. On the Absurdity of "Being a Psychologist". Bulletin of the B.P.S., 1977, 30, 211.
- BEGG, I. Imagery and Integration in The Recall of Words. Canadian Journal of Psychology, 1973, 27, 159 - 167.
- BEGG, I., and PAIVIO, A. Concreteness and Imagery in Sentence Meaning. J.V.L. V.B., 1969, 8, 821 - 827.

- BEECH, J.R. Image Scanning: More Problems for Pure Propositional Accounts of Visual Imagery. Bulletin of the B.P.S., 1978, 31, 162 (Abstract of 1978 B.P.S. Conference Paper).
- BEGG, I., and ROBERTSON, R. Imagery and Long Term Retention. JVLVB, 1973, 12, 689 - 700.
- BELCH, P.J. An investigation of the Effects of Different Questioning Strategies on the Reading Comprehension Scores of Secondary Level EMR Students. Unpub. Ed. D. Thesis, 1974, West Virginia University.
- BERECZ, J. Treatment of Smoking with Cognitive Conditioning Therapy: A Self-Administered Aversion Technique. Behavior Therapy, 1976, 7, 641 - 648.
- BERLYNE, D.E. Curiosity and Exploration. Science, 1966, 153, 25 - 33.
- BERNSTEIN, S.L. The Effects of Children's Question asking Behavior on Problem Solution and Comprehension of Written Material. Unpub. PhD Thesis, 1973, Columbia University.
- BINYON, M. Some Scouts Unhappy over Reforms. T.E.S. 3.7.1970, P.8.
- BLAKEMORE, C. Madness and Morality. The Listener, 16/12/76, 779 - 782. (From the 1976 Reith Lectures).
- BLICK, K.A., and BOLTWOOD, C.E. Mnemonic Techniques Used by College Students in Paired-Associates Learning. Psychological Reports, 1972, 31, 459-462.
- BOBROW, S.A. and BOWER, G.H., Comprehension and Recall of Sentences. Journal of Experimental Psychology, 1969, 80, 455 - 461.
- BORG, W.R. Educational Research, 1963, Longmans.
- BORMUTH, J.R., CARR, J., MANNING, J., and PEARSON, D. Children's Comprehension of Between - and Within - Sentence Syntactic Structures. Journal of Educational Psychology, 1970, 61, 349 - 357.

- BOWER, G.H. Analysis of a Mnemonic Device. American Scientist, 1970, 58, 496-510 (a).
- BOWER, G.H. Imagery as a Relational Organiser in Associative Learning. JVLVB, 1970, 9, 529-533 (b).
- BOWER, G.H. Mental Imagery and Associative Learning. In GREGG, L.W. (Ed.) Cognition in Learning and Memory, 1972 Wiley and Sons.
- BOWER, G.H. Educational Applications of Mnemonic Devices. In DOYLE, K.O. (Ed.) Interactions: Readings in Human Psychology, 1973, D.C. Heath and Co. (a).
- BOWER, G.H. How to.....wh.....remember! Psychology Today (U.S.A.), Oct. 1973, 63-70 (b).
- BOWER, G.H. Experiments on Story Understanding and Recall. Quarterly Journal of Experimental Psychology, 1976, 28, 511-534.
- BOWER, G.H. and REITMAN, J.S. Mnemonic Elaboration in Multilist Learning. JVLVB, 1972, 11, 478-485.
- BRANSFORD, J.D., and FRANKS, J.J. The Abstraction of Linguistic Ideas. Cognitive Psychology, 1971, 2, 331-350.
- BRANSFORD, J.D., and JOHNSON, M.K. Contextual pre-requisites for Understanding: Some Investigations of Comprehension and Recall. JVLVB, 1972, 11, 717-726 (a)
- BRANSFORD, J.D., and JOHNSON, M.K. Consideration of Some Problems in Comprehension. In CHASE, W.G. (Ed.) Visual Information Processing, 1972, (b).
- BRANSFORD, J.D., and McCARRELL, N.S. A Sketch of a Cognitive Approach to Comprehension: Some thoughts about Understanding What it means to Comprehend. In WEIMER, W.B., and PALERMO, D.S. (Eds.) Cognition and the Symbolic Processes, 1974, Lawrence Erlbaum.
- BREWER, W.F. The Problem of Meaning and the Interrelations of the Higher Mental Processes. In WEIMER, W.B. and PALERMO, D.S. (Eds.) Cognition and the Symbolic Processes, 1974, Lawrence Erlbaum.

- BRIGGS, G.G., HAWKINS, S., and CROVITZ, H.F. Bizarre Images in Artificial Memory. Psychonomic Science, 1970, 19, 353-354.
- BRITTON, J. What is "English"? Trends in Education, 1974, 36, 24-28.
- BROADBENT, D.E. Cognitive Psychology and Education. British Journal of Educational Psychology, 1975, 45, 162-176.
- BROOKS, A. New Emphases in Scouts and Guides. T.E.S. 8.12.1972 p.32.
- BROOKS, L.R. Spatial and Verbal Components of the Act of Recall. Canadian Journal of Psychology, 1968, 22, 349-368.
- BROWN, A.L. Mnemonic Elaboration and Recency Judgements in Children. Cognitive Psychology, 1973, 5, 233-248.
- BROWN, A.L., and SMILEY, S.S. Rating the Importance of Structural Units of Prose Passages: a Problem of Meta-Cognitive Development. Child Development, 1977, 48, 1-8.
- BROWN, R. A First Language: The Early Stages, 1973, George Allen and Unwin.
- BRUNER, J.S. Learning and Thinking. Harvard Educational Review, 1959, 29, 184-192.
- BRUNER, J.S. The Process of Education, 1960, Harvard University Press.
- BRUNER, J.S. The Course of Cognitive Growth. American Psychologist, 1964, 19, 1-15.
- BRUNER, J.S. Toward a Theory of Instruction. 1966, Belknap
- BRUNER, J.S. Beyond the Information Given, 1974, George Allen and Unwin.
- BRUNER, J.S., Olver, R.R., and GREENFIELD, P.M. Studies in Cognitive Growth, 1966, Wiley.
- BUGELSKI, B.R. Presentation Time, Total Time, and Mediation in Paired-Associate Learning. Journal of Experimental Psychology, 1962, 63, 409-412.

- BUGELSKI, B.R. Words and Things and Images. American Psychologist, 1970, 25, 1002-1012.
- BUGELSKI, B.R. The Definition of the Image. In SEGAL, S.J. Imagery - Current Cognitive Approaches, 1971.
- BUGGIE, S.E. Imagery and Relational Variety in Associative Learning. Unpub. PhD Thesis, Oregon University, 1974.
- BULL, B.L. and Wittrock, M.C. Imagery in the Learning of Verbal Definitions. British Journal of Educational Psychology, 1973, 43, 289-293.
- BULLOCK COMMITTEE A Language for Life, 1975, H.M.S.O.
- BUONASSISSI, J.V., BLICK, K.A., and KIBLER, J.L. Evaluation of Subject - Generated Descriptive Story Mnemonics in a Free-Recall Task. Psychological Reports, 1972, 30, 648.
- BURSTALL, C. Factors Affecting Foreign Language Learning: a Consideration of Some Recent Research Findings. Language Teaching and Linguistics: Abstracts, 1975, 8, 5-25.
- BURSTALL, C., JAMIESON, M., COHEN, S., and HARGREAVES, M. Primary French in the Balance, 1974, NFER.
- BUZAN, T. Use your Head, 1974, BBC Publications.
- CARROLL, J.B. Research on Teaching Foreign Languages. In GAGE, N.L. (Ed.) Handbook of Research on Teaching, 1963, Rand McNally.
- CARROLL, J.B. Basic and Applied Research in Education: Definitions, Distinctions and Implications. Harvard Educational Review, 1968, 38, 263-276.
- CARROLL, J.B., Defining Language Comprehension Some Speculations. In FREEDLE, R.O., and CARROLL, J.B. (Eds.) Language Comprehension and the Acquisition of Knowledge, 1972, Winston.
- CARROLL, J.B. The Potentials and Limitations of Print as a Medium of Instruction. In OLSON, D.R. (Ed.) Media and Symbols The forms of Expression, Communication and Education, 1974 University of Chicago Press.

- CARVER, R.P. A critical Review of Mathemagenic Behaviors and the Effect of Questions Upon the Retention of Prose Materials. Journal of Reading Behavior, 1972, 4, 93-119.
- CHAFE, W.L. Discourse Structure and Human Knowledge. In CARROLL, J.B. and FREEDLE, R.O. (Eds.) Language Comprehension and the Acquisition of Knowledge, 1972, V.H. Winston.
- CLARK, H.H. Bridging. In JOHNSON-LAIRD, P.N., and WASON, P.C. (Eds.) Thinking: Readings in Cognitive Science, 1977, Cambridge University Press.
- CLARKSON, T.A., HAGGITH, P.A., TIERNEY, M.C., and KOBASIGAWA, A. Relative effectiveness of Imagery Instructions and Pictorial Interaction on Children's Paired-Associate Learning. Child Development, 1973, 44, 179-181.
- CLELAND, D.L. A construct of Comprehension. Reading and Inquiry, I.R.A. Proceedings, 1965, 10, 59-64.
- CLYMER, T. What is "reading"? Some Current Concepts. In MELNIK, A., and MERRITT, J. (Eds.) Reading: Today and Tomorrow, 1972, University of London Press.
- COCKRIEL, I.W. Sociometric Status Scores and Reading Achievement. Handbook of International Sociometry, 1973, 7, 84-87.
- COLE, L.R. Relationships between Visual Representations and Linguistic Items in Second-Language Teaching. I.R.A.L., 1976, 14, 339-350.
- COLLINS, A.M. and QUILLIAN, R. Experiments on Semantic Memory and Language Comprehension in GREGG, L.W. (Ed.) Cognition in Learning and Memory, 1972, Wiley and Sons.
- COLLYER, S.C., JONIDES, J., and BEVAN, W. Images as Memory Aids: is Bizarreness Helpful? American Journal of Psychology, 1972, 85, 31-38.
- COLTHEART, M., HULL, E., and SLATER, D. Sex Differences in Imagery and Reading. Nature, 1975, 253, 438-440.
- CRAIK, F.I.M., and LOCKHART, R.S. Levels of Processing: a Framework for Memory Research. J.V.L.V.B., 1972, 11, 671-684.
- CROMER, W. The Difference Model: a New Explanation for Some Reading Difficulties. Journal of Educational Psychology, 1970, 61, 471-483.
- CROVITZ, H.F. The Capacity of Memory Loci in Artificial Memory. Psychonomic Science, 1971, 24, 187-188.

- D'AMICO, M.L. Effects of Interspersed Questions, Imagery Instructions and Feedback on Acquisition and Retention of Written Counselor Material. Unpub. Ed. D. Thesis, 1976, West Virginia University.
- D'ARCY, P. Reading for Meaning - 2 The Reader's Response 1973, Schools Council.
- DAVIDSON, K. Objective Text. Use of English, 1974, 26, 12-18.
- DAVIDSON, R.E., and ADAMS, J.F. Verbal and Imagery Processes In Children's Paired-Associate Learning. Journal of Experimental Child Psychology, 1970, 9, 429-435.
- DELIN, P.S. Success in Recall as a Function of Success in Implementation of Mnemonic Instructions Psychonomic Science, 1968, 12, 153-154.
- DE VILLIERS, P.A. Imagery and Theme in the Recall of Connected Discourse. Journal of Experimental Psychology, 1974, 103, 263-268.
- DIXON, J. Current Concerns and the Future of NATE. English in Education, Summer 1972, 6, 71-83.
- DODWELL, P.C. New Horizons in Psychology: 2, 1972, Penguin.
- DOOB, L.W. The Ubiquitous Appearance of Images. In SHEEHAN, P.W. (Ed.) The Function and Nature of Imagery, 1972, Academic Press.
- DOOLING, D.J. Some Context Effects in the Speeded Comprehension of Sentences. Journal of Experimental Psychology, 1972, 93, 56-62.
- DOUGLAS, W. An American View of the Failure of Curriculum Reform and the Way Ahead. English in Education, Summer 1972, 5-18.
- DUCKWORTH, D., and ENTWISTLE, N.J. Attitudes to School Subjects: a Repertory Grid Technique. British Journal of Educational Psychology, (1974), 44, 76-83.
- DWYER, F.M. Adapting Visual Illustrations for Effective Learning. Harvard Educational Review, 1967, 37, 250-263.

- EDFELDT, A.W. Concerning Differences Between Good and Poor Readers. Journal of Experimental Education, 1975, 43, 90-93.
- EDMUNSON, E.D., and NELSON, D.L. Anxiety, Imagery and Sensory Interference. Bulletin of the Psychonomic Society, 1976, 8, 319-322.
- EVANS, J. St. B. T. On the Absurdity of Being an Anti-Psychologist: A reply to Bannister. Bulletin of the B.P.S., 1977, 341-342.
- EYSENCK, M.W. Levels of Processing: A Critique. British Journal of Psychology, 1978, 69, 157-169.
- FARNES, N.C. Reading Purposes, Comprehension and the Use of Context, 1973. (PE 261, Units 2 and 3) Open University Press.
- FARNHAM-DIGGORY, S. Cognitive Processes in Education, 1972, Harper and Row.
- FAW, H.W., and WALLER, T.G. Mathemagenic Behaviors and Efficiency in Learning from Prose Materials: Review, Critique and Recommendations. Review of Educational Research, 1976, 46, 691-720
- FLAVELL, J. Developmental Studies of Mediated Memory. In REESE, H.W., and LIPSITT, L.P. (Eds.) Advances in Child Development and Behavior, 1970, Vol. 5.
- FORSTER, K.I., and OLBREI, I. Semantic Heuristics and Syntactic Analysis. Cognition, 1973, 2, 319-347.
- FRANKS, J.J. Toward Understanding Understanding. In WEIMER, W.B., and PALERMO, D.S. (Eds.) Cognition and the Symbolic Processes, 1974, Lawrence Erlbaum.
- FRASE, L.T. Boundary Conditions for Mathemagenic Behaviors. Review of Educational Research, 1970, 40, 337-347.
- FRASE, L.T. Prose Processing. In BOWER, G. H. (Ed.) The Psychology of Learning and Motivation, 1975, 9, 1-47.
- FRIEDMAN, A., and BOURNE, L.E. Encoding the Levels of Information in Pictures and Words. Journal of Experimental Psychology, 1976, 105, 169-190.

- GARTEN, J.A., and BLICK, K.A. Retention of Word-Pairs for Experimenter - Supplied and Subject - Originated Mnemonics Psychological Reports, 1974, 35, 1099 - 1104.
- GATTENGNO, C. The Problem of Reading is Solved. Harvard Educational Review, 1970, 40, 283-286.
- GIBSON, E.J. Principles of Perceptual Learning and Development, 1967.
- GLUCKSBERG, S., TRABASSO, T., and WALD, J. Linguistic Structures and Mental Operations. Cognitive Psychology, 1973, 5, 338-370.
- GOLDMAN, R.D. Effects of a Logical Versus a Mnemonic Learning Strategy on Performance in Two Undergraduate Psychology Classes. Journal of Educational Psychology, 1972, 63, 347-352.
- GOLLA, F., HUTTON, E.L., and GREY WALTER, W. The Objective Study of Mental Imagery. I Physiological Concomitants. Journal of Mental Science, 1943, 89, 216-223.
- GOMBRICH, E.H. The Visual Image. Scientific American, 1972, 227, 82-96.
- GOODMAN, K.S. Reading - Psycholinguistic Guessing Game. Journal of the Reading Specialist, 1967, 266-271.
- GORDON, D.A. and BAUMEISTER, A.A., The use of Verbal Mediation in the Retarded as a Function of Developmental Level and Response Availability Journal of Experimental Child Psychology, 1971, 12, 95-105.
- GORDON, R. An Investigation into Some of the Factors that Favour the Formation of Stereotyped Images. British Journal of Psychology, 1949, 39, 156-167.
- GRANGER, R.C. Children's Noun-Pair Learning as a Function of Subject - Generated or Experimenter - Provided Strings and Response Measure. Unpub. Ed. D. Thesis, University of Massachusetts, 1973.
- GRIFFITH, D. The Attentional Demands of Mnemonic Control Processes. Memory and Cognition, 1976, 4, 103-108.

- GRONINGER, L.D. Imagery and Subjective Categorization Effects on Long-Term Recognition and Retrieval. Bulletin of the Psychonomic Society, 1974, 3, 261-263.
- 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.
- HABER, R.N. Where are the Visions in Visual Perception? In SEGAL, S.J. (Ed.) Imagery - Current Cognitive Approaches, 1971.
- HAGEN, J.W. Strategies for Remembering. In FARNHAM-DIGGORY, S. (Ed.) Information - Processing in Children, 1972, Academic Press.
- HAKATU, K., and CANCINO, H. Trends in Second Language Research. Harvard Educational Review, 1977, 47, 282-293.
- HARRINGTON, C.V. Some Thoughts and Research on Vocabulary Teaching. E.L.E.C. Bulletin, 1971, 34, 21-30.
- HARRIS, M.E.R. Imagery in Children - Assessment and Correlates. Unpub. PhD Thesis, Syracuse University, 1976.
- HAUCK, P.D., WALSH, C.C., and KROLL, N.E.A. Visual Imagery Mnemonics: Common Vs Bizarre Mental Images. Bulletin of the Psychonomic Society, 1976, 7, 160-162.
- HAVILAND, S.E., and CLARK, H.H. What's new? Acquiring new Information as a Process in Comprehension. JVLVB, 1974, 13, 512-521.
- HICKS, R.E., and YOUNG, R.K. Part - Whole Transfer in Free Recall as a Function of Word Class and Imagery. Journal of Experimental Psychology, 1973, 101, 100-104.
- HEIEN, L.G. Language Teaching in Different Keys. Modern Language Journal, 1973, 57, 185-189.
- HILGARD, E.R., and BOWER, G.H. Theories of Learning (4th Edition), 1975, Appleton - Century - Crofts.
- HILLER, J.H. Learning from Prose Text: Effects of Readability Level, Inserted Question Difficulty, and Individual Differences Journal of Educational Psychology, 1974, 66, 202-211

- HIRST, P.H. The Curriculum. Schools Council Working Paper No. 12 (1967) H.M.S.O.
- HITCH, G., and BADDELEY, A. Memories are Made of This. Psychology Today, 1977, 3, (2), 11-15.
- H.M.I. Report. Modern Languages in Comprehensive Schools (1977) H.M.S.O.
- H.M.I. Report. Secondary School English - The First two Years. Trends in Education, 1981, 21, 6-12.
- HOLLENBERG, C.K. Functions of Visual Imagery in the Learning and Concept Formation of Children. Child Development, 1970, 41, 1003-1015.
- HOLT, R.R. Imagery: The Return of the Ostracised. American Psychologist, 1964, 19 254-264.
- HOLT, J. How Children Fail, 1965, Pitman.
- HOLYOAK, K., HOGETERP, H., and YUILLE, J.C. A Developmental Comparison of Verbal and Pictorial Mnemonics in Paired-Associate Learning. Journal of Experimental Child Psychology, 1972, 14, 53-65.
- HONEY FORD, R. Against Objective Testing. Use of English, 1973, 25, 17-20 and 26.
- HORVITZ, J.M., and LEVIN, J.R. Semantic and Imaginal Factors in Learning as Related to Age. Journal of Experimental Child Psychology, 1972, 14, 11-20.
- HOWE, M.J.A. Good Learners and Poor Learners. Bulletin of the B.P.S., 1976, 29, 16-19.
- HUNTER, I.M.L. Mnemonic Systems and Devices. Science News 1956, 39, 75-97.
- HUNTER, I.M.L. Memory, 1957, Penguin.
- HUNTER, I.M.L. How to Play Twenty Questions. New Society, 26/12/74, 810-811.
- HUNTER, I.M.L. Memory: Theory and Application. In VARMA, V.P., and WILLIAMS, P. (Eds.) Piaget, Psychology and Education, 1976, Hodder and Stoughton.

- HUTTENLOCHER, J. Constructing Spatial Images: a Strategy in Reasoning. In JOHNSON-LAIRD, P.N. and WASON, P.C. (Eds.) Thinking Readings in Cognitive Science, 1977, Cambridge University.
- ISAKSON, R.L., and MILLER, J.W. Sensitivity to Syntactic and Semantic Cues in Good and Poor Comprehenders. Journal of Educational Psychology, 1976, 68, 787-792.
- JABLONSKI, E.M. Cued and Uncued Multitrial Free Recall as a Function of Age, Mnemonic Instructions and Sex. Unpub. PhD Thesis, Missouri University, 1972.
- JAMES, W. Psychology (Briefer Course), 1892, MacMillan and Co.
- JANSSEN, W.H. Selective Interference During the Retrieval of Visual Images. Quarterly Journal of Experimental Psychology, 1976, 28, 535-539.
- JARMEN, R.F. Level I and Level II Abilities: Some Theoretical Reinterpretations. British Journal of Psychology, 1978, 69, 257-269.
- JENKINS, J.R., NEALE, D.C., and DENO, S.L. Differential Memory for Pictures and Word Stimuli. Journal of Educational Psychology, 1967, 58, 303-307.
- JENSEN, A.R. How Much can We Boost IQ and Scholastic Achievement? Harvard Educational Review, 1969, 39, 1-123.
- JENSEN, A.R., and ROHWER, W.D. Verbal Mediation in Paired-Associate and Serial Learning. J.V.L.V.B., 1963, 1, 346-352.
- JERVIS, C.S. Aspects of English. T.E.S. 20/9/1974.
- JOHNSON, M.K., BRANSFORD, J.D., NYBERG, S.E. and CLEARY, J.J. Comprehension factors in Interpreting Memory for Abstract and Concrete Sentences. J.V.L.V.B., 1972, 11, 451-454.
- JOHNSON-LAIRD, P.M., and WASON, P.C. Thinking: Readings in Cognitive Science, 1977, Cambridge University.
- JONES, B. Response Bias in the Recognition of Pictures and Names by Children. Journal of Experimental Psychology, 1974, 103, 1214-1215.
- JONES, R. From Courses to Resources. T.E.S., 20/9/1974.
- JONES, S. Imagery in a Short Term Memory Task. Psychonomic Science, 1972, 29, 76-77.

- JOYNSON, R.B. The breakdown of Modern Psychology. Bulletin of the B.P.S., 1970, 23.
- KAGAN, J. Developmental Studies in Reflection and Analysis. In CASHDAN, A., and WHITEHEAD, J. Personality Growth and Learning, 1971, Longmans.
- KATZ, S., ATKESON, B., and LEE, J. The Bransford - Franks Effect: Integration or Artefact? Memory and Cognition, 1974, 2, 709-713.
- KATZ, L. and WICKLUND, D.A. Letter Scanning Rate for Good and Poor Readers in Grades Two and Six. Journal of Educational Psychology, 1972, 63, 363-367.
- KENNEDY, J.M. A Psychology of Picture Perception, 1974, Jossey-Bass.
- KERST, S.M. Interactive Visual Imagery and Memory Search for Words and Pictures. Memory and Cognition, 1976, 4, 573-580.
- KERST, S.M. and LEVIN, J.R. A Comparison of Experimenter - Provided and Subject - Generated Strategies in Children's Paired-Associate Learning. Journal of Educational Psychology, 1973, 65, 300-303.
- KIBLER, J.L., and BLICK, K.A. Evaluation of Experimenter - Supplied and Subject - Generated First Letter Mnemonic in a Free Recall Task. Psychological Reports, 1972, 30, 307-313.
- KIERAS, D. Beyond Pictures and Words: Alternative Information - Processing Models for Imagery Effects in Verbal Memory. Psychological Bulletin, 1978, 85, 532-554.
- KING, D.L. An Image Theory of Classical Conditioning. Psychological Reports, 1973, 33, 403-411.
- KING, D.J., and RUSSELL, G.W. A Comparison of Rote and Meaningful Learning of Connected Meaningful Material. J.V.L.V.B., 1966, 5, 478-483.
- KINTSCH, W. The Representation of Meaning in Memory, 1974, John Wiley.
- KINTSCH, W., and BATES, E. Recognition Memory for Statements From a Classroom Lecture. Journal of Experimental Psychology, 1977, 106
- KIRBY, J.R., and DAS, J.P. Comments on Paivio's Imagery Theory. Canadian Psychological Review, 1976, 17, 66-68.

- KLEE, H., and EYSENCK, M.W. Comprehension of Concrete and Abstract Sentences. J.V.L.V.B., 1973, 12, 522-529.
- KNIGHT, R. Watch that Language! T.E.S., 6/1/1978.
- KOESTLER, A. The Act of Creation, 1964, Hutchinson.
- KOSSLYN, S.M. Constructing Visual Images: an Exercise in Neo-Mentalism Unpub. PhD Thesis, Stanford University, 1974.
- KOSSLYN, S.M., and ALPER, S.N. On the Pictorial Properties of Visual Images: Effects of Image Size on Memory for Words. Canadian Journal of Psychology, 1977, 31, 32-40.
- KOSSLYN, S.M., HOLYOAK, K.J., and HUFFMAN, C.S. A Processing Approach to the Dual Coding Hypothesis. Journal of Experimental Psychology, 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.
- KRASHEN, S.D. The Left Hemisphere. UCLA Educator, 1975, 17, 17-23.
- KRETSCHMER, J.C. Toward a Piagetian Theory of Reading Comprehension. Reading World, 1975, 14, 180-188.
- KULHAVY, R.W., and HEINEN, J.R. Imaginal Attributes in Learning Sentence - Embedded Noun Pairs. Psychological Reports, 1974, 26, 487-490.
- KULHAVY, R.W., and SWENSON, I. Imagery Instructions and the Comprehension of Text. British Journal of Educational Psychology, 1975, 45, 47-51.
- LEBRATO, M.T., and ELLIS, N.R. Imagery Mediation in Paired-Associate Learning by Retarded and Non-Retarded Subjects. American Journal of Mental Deficiency, 1974, 78, 704-713.
- LEE, W. Supra-Paragraph Prose Structure: its Specification, Perception and Effects on Learning. Psychological Reports, 1965, 17, 135-144.

- LESGOLD, A.M., and GOLDMAN, S.R. Encoding Uniqueness and the Imagery Mnemonic in Associative Learning. J.V.L.V.B., 1973, 12, 193-202.
- LESGOLD, A.M., McCORMICK, C., and GOLINKOFF, M. Imagery Training and Children's Prose Learning. Journal of Educational Psychology, 1975, 67, 663-667.
- LESGOLD, A.M., LEVIN, J.R., SHIMRON, J., and GUTTMAN, J. Pictures and Young Children's Learning from Oral Prose. Journal of Educational Psychology, 1975, 67, 636-642.
- LEVIN, J.R. Inducing Comprehension in Poor Readers: a Test of a Recent Model. Journal of Educational Psychology, 65, 19-24.
- LEVIN, J.R., DAVIDSON, R.E., WOLFF, P., and CITRON, M. A comparison of Induced Imagery and Sentence Strategies in Children's Paired-Associate Learning. Journal of Educational Psychology, 1973, 64, 306-309.
- LEVIN, J.R., DIVINE-HAWKINS, P., KERST, S.M., and GUTTMAN, J. Individual Differences in Learning from Pictures and Words: the Development and Application of an Instrument. Journal of Educational Psychology, 1974, 66, 296-303.
- LEVIN, J.R., GHATALA, E.S., GUTTMAN, J., SUBKOVIK, M., and BENDER, B.G. Frequency Interference in Children's Recognition of Sentence Information. American Educational Research Journal, 1978, 15, 39-51.
- LEVIN, J.R., HORVITZ, J.M., and KAPLAN, S.A. Verbal Facilitation of Paired-Associate Learning: a Limited Generalization. Journal of Educational Psychology, 1971, 62, 439-444.
- LEVIN, J.R., and KAPLAN, S.A. Imaginal Facilitation of Paired-Associate Learning: a Limited Generalization? Journal of Educational Psychology, 1972, 63, 429-432.
- LEVIN, J.R., ROHWER, W.D., and CLEARY, A. Individual Differences in the Learning of Verbally and Pictorially Presented Paired-Associates. American Educational Research Journal, 1971, 8, 11-26.

- LINDAUER, M.S. The Sensory Attributes and Functions of Imagery and Imagery - Evoking Stimuli. In SHEEHAN, P.W. The Function and Nature of Imagery, 1972, Academic Press.
- LOCKHART, R.S., and CRAIK, F.I.M. Levels of Processing: a Reply to Eysenck. British Journal of Psychology, 1978, 69, 171-175.
- LOWRY, D.H. The Effects of Mnemonic Learning Strategies on Transfer, Interference and 48 Hour Retention. Journal of Experimental Psychology, 1974, 103, 16-20.
- LUBKE, D. The Role of the Mother Tongue in Learning Vocabulary. Language Teaching Abstracts, 1972, 5, 40.
- LURIA, A.R. The Mind of a Mnemonist, 1969, Jonathon Cape.
- MARCEL, T. The Effective Visual Field and the Use of Context in Fast and Slow Readers of Two Ages. British Journal of Psychology, 1974, 65, 479-492.
- MARKS, D.F. Individual Differences of Visual Imagery and Their Effect on Functions. In SHEEHAN, P.W. (Ed.) The Function and Nature of Imagery, 1972, Academic Press.
- MARLAND, M. English in the Classroom. Trends in Education, 1968, 12, 33-38.
- MASON, G. Where will All the English Go? T.E.S. 2/6/1978.
- McClure, D.R. The Determining Effect of Verbal and Non-Verbal Cues Upon the Comprehension of Fourth Grade Readers. Unpub. EdD Thesis, 1974, Georgia University.
- MCDONALD, F.J. The Teacher and the Improvement of Educational Practice. In DE CECCO, J.P., Human Learning in the School, 1963, Holt, Rinehart and Winston.
- McGaw, B., and GROTELUESCHEN, A. Direction of the Effect of Questions in Prose Material. Journal of Educational Psychology, 1972, 63, 580-588.
- McKELLAR, P. Imagery From the Standpoint of Introspection. In SHEEHAN, P.W. (Ed.) The Function and Nature of Imagery (1972).

- McMAHON, T.P. Imagery and Verbal Mediation in Learning Sight Words with Second Graders. Unpub. EdD Thesis, 1975, North Colorado University.
- McNULTY, J.A. The Effects of Instructions to Mediate Upon P.A. Learning. Psychonomic Science, 1966, 4, 61-62.
- MELTON, A.W. The Science of Learning and the Technology of Educational Methods. In DE CECCO, J.P. (Ed.) Human Learning in the School, 1963, Holt, Rinehart and Winston.
- MERRICK, B. English Course Books for Secondary Schools. English in Education, Spring 1974, 8, 24-34.
- MERRITT, J. Perspectives on Reading (PE 261, Unit 1), 1973, Open University.
- MERRY, R., and GRAHAM, N.C. Imagery Bizarreness in Children's Recall of Sentences. British Journal of Psychology, 1978, 315-321.
- MESSENGER, T. Language Across the Curriculum. In RAGGETT, M., and CLARKSON, M. Curriculum Skills: Classroom, Language and Reading, 1974, Ward Lock International.
- MILGRAM, N.A. Retention of Mediation Set in Paired-Associate Learning of Normal Children and Retardates. Journal of Experimental Child Psychology, 1967, 5, 341-349 (a).
- MILGRAM, N.A. Verbal Content Versus Visual Compound in Paired-Associate Learning by Children. Journal of Experimental Child Psychology, 1967, 5, 597-603 (b).
- MILLER, G.A., GALANTER, E., and PRIBRAM, K.H. Plans and the Structure of Behaviour, 1960, Holt.
- MILLER, G.A. Psychology, the Science of Mental Life, 1962, Hutchinson.
- MILLER, G.A., and ISARD, S. Some Perceptual Consequences of Linguistic Rules J.V.L.V.B., 1963, 2, 217-228.
- MISTLER-LACHMAN, J.L. Levels of Comprehension in the Processing of Normal and Ambiguous Sentences. JVLVB, 1972, 11, 614-623.

- MISTLER-LACHMAN, J.L. Depth of Comprehension and Sentence Memory. JVLVB, 1974, 13, 98-106.
- MOESER, S.D. Memory for Meaning and Wording in Concrete and Abstract Sentences. J.V.L.V.B., 1974, 13, 682-697.
- MOESER, S.D. Memory for Language Organisation in Concrete and Abstract Sentences. Memory and Cognition, 1975, 3, 560-568.
- MONTAGUE, W.E. Elaborative Strategies in Verbal Learning and Memory. In BOWER, G.H. (Ed.) The Psychology of Learning and Motivation, 1972, 6, Academic Press, N.Y.
- MOORE, S., and ANTROBUS, A.L. Longman's French Course: Teacher's Book (1973) Longmans.
- MORRIS, P.E. Imagery Reviewed: Propositions, Analogies and Anticipations. Bulletin of the B.P.S., 1978, 31, 182, (Abstract of 1978 B.P.S. Conference Paper).
- MORRIS, P.E., and COOK, N. When do First Letter Mnemonics Aid Recall? British Journal of Educational Psychology, 1978, 48.
- MORRIS, P.E., and STEVENS, R. Linking Images and Free Recall. JVLVB, 1974, 13, 310-315.
- MOWBRAY, C.T., and LURIA, Z. Effects of Labelling on Children's Visual Imagery. Developmental Psychology, 1973, 9, 1-8.
- MUELLER, J.H., and JABLONSKI, E.M. Instructions, Noun Imagery and Priority in Free Recall. Psychological Reports, 1970, 27, 559-556.
- MUELLER, T. Another Look at how to Teach Listening and Reading Comprehension. Modern Language Journal, 1974, 58, 19-23.
- MURRAY, D.J., and ROBERTS, B. Visual and Auditory Presentation, Presentation Rate and Short-Term Memory in Children. British Journal of Psychology, 1968, 59, 119-125.
- MURRAY, F.S., and HOUSE, C. Imagery in Paired-Associate Learning in Five-Year-Old Children. Bulletin of the Psychonomic Society, 1976, 8, 135-138.

- NAPPE, G.W., and WOLLEN, K.A. Effects of Instructions to Form Common and Bizarre Mental Images on Retention. Journal of Experimental Psychology, 1973, 100, 6-8.
- NEISSER, U. Cognitive Psychology, 1967, Appleton.
- NEISSER, U. Changing Conceptions of Imagery. In SHEEHAN, P.W. (Ed.) The Function and Nature of Imagery, 1972, Academic Press.
- NEISSER, U. Cognition and Reality, 1976, W.H. Freeman.
- NELSON, D.L., REED, V.S., and WALLING, J.R. Pictorial Superiority Effect. Journal of Experimental Psychology, 1976, 2 523-528.
- NEVILLE, M.H. Effectiveness of Rate of Aural Message on Reading and Listening. Educational Research, 1975, 18, 37-43.
- NEVILLE, M.H., and PUGH, A.K. Context in Reading and Listening: a Comparison of Children's Errors in Cloze Tests. British Journal of Educational Psychology, 1974, 44, 224-232.
- NEWMAN, S.E. and SALTZ, E. Effects of Contextual Cues on Learning from Connected Discourse. American Journal of Psychology, 1960, 73, 587-592.
- NEWSOM, J. et al. Half Our Future. (The Newsom Report), 1963, H.M.S.O.
- N.F.E.R. Integrated Innovation. Educational Research News No. 30, Spring 1978.
- NISBET, J.D., and ENTWISTLE, N.T. Educational Research Methods, 1970, London University Press.
- NORMAN, D.A. Memory and Attention: Introduction to Human Information Processing, 1969, Wiley.
- NUFFIELD FOREIGN LANGUAGES TEACHING MATERIALS PROJECT, En Avant (Teachers' Book Stage 1A), 1972 (4th Edition), E.J. Arnold and Son Ltd.
- ODOM, P.B., and NESBITT, N.H. Some Processes in Children's Comprehension of Linguistically and Visually Depicted Relationships. Journal of Experimental Child Psychology, 1974, 17, 399-408.

- ODOM-BROOKS, P.H., and ARNOLD, D.J. Cognitive Development in Mental Subnormality. In HAMILTON, V., and VERNON, M.D. (Eds.) The Development of Cognitive Processes, 1976, Academic Press.
- OGDEN, C.K., and RICHARDS, I.A. The Meaning of Meaning, 1923, R.K.P.
- OLSON, D.R. From Utterance to Text: the Bias of Language in Speech and Writing. Harvard Educational Review, 1977, 47, 257-281.
- OLSON, D.R., and BRUNER, J.S. Learning through Experience and Learning through Media. In OLSON, D.R. (Ed.) Media and Symbols: The Forms of Expression, Communication and Education, 1974, University of Chicago Press.
- OTT, C.E., BLAKE, R.S. and BUTLER, D.C. Implications of Mental Elaboration for the Acquisition of Foreign Language Vocabulary. I.R.A.L., 1976, 14, 37-48.
- OTT, C.E., BUTLER, D.C., BLAKE, R.S., and BALL, J.P. The Effect of Interactive-Image Elaboration on the Acquisition of Foreign Language Vocabulary. Language Learning, 1973, 23, 197 -206.
- PAGE, B. Graded Tests in Language Proficiency. Annual Conference of the Joint Council of Language Associations, January 1978.
- PAIVIO, A. Mental Imagery in Associative Learning and Memory. Psychological Review, 1969, 76, 241-263.
- PAIVIO, A. On the Functional Significance of Imagery. Psychological Bulletin, 1970, 73, 385-392.
- PAIVIO, A. Imagery and Verbal Processes, 1971, Holt, Rinehart and Winston.
- PAIVIO, A. A Theoretical Analysis of the Role of Imagery in Learning and Memory. In SHEEHAN, P.W. (Ed.) The Function and Nature of Imagery, 1972, Academic Press.
- PAIVIO, A. Language and Knowledge of the World. Educational Researcher, 1974, 3, 5-12.
- PAIVIO, A. Imagery and Synchronic Thinking. Canadian Psychological Review, 1975, 16, 147-163 (a).
- PAIVIO, A. Coding Distinctions and Repetition Effects in Memory. In BOWER, G.H. (Ed.) The Psychology of Learning and Motivation, 1975, 9, 179-214 (b).

- PAIVIO, A., and CSAPQ K. Picture Superiority in Free Recall: Imagery or Dual Coding? Cognitive Psychology, 1973, 5, 176-206.
- PAIVIO, A., and Foth, D. Imaginal and Verbal Mediators and Noun Concreteness in Paired-Associate Learning: The Elusive Interaction JVLVB, 1970, 9, 384-390.
- PAIVIO, A., and YUILLE, J.C. Changes in Associative Strategies and Paired Associate Learning over Trials as a Function of Word Imagery and Type of Learning Set. Journal of Experimental Psychology, 1969, 79, 458-463.
- PALERMO, D.S. Imagery in Children's Learning: Discussion. Psychological Bulletin, 1970, 73, 415-421.
- PARIS, S.G., and LINDAUER, B.K. The Role of Interference in Children's Comprehension and Memory For Sentences. Cognitive Psychology, 1976, 8, 217-227.
- PARIS, S.G., MAHONEY, G.J., and BUCKHALT, J.A. Facilitation of Semantic Integration in Sentence Memory of Retarded Children. American Journal of Mental Deficiency, 1974, 78, 714-720.
- PARIS, S.G., and UPTON, L.R. Children's Memory for Inferential Relationships in Prose. Child Development, 1976, 47, 660-668.
- PEEL, E.A. The Analysis of Comprehension and Judgement From Textual Material. Educational Review, 1974, 27, 100-113.
- PENDLEBURY, B.J. A Grammar School English Course, 1958, Nelson.
- PERENSKY, J.J., and SENTER, R.J. An Investigation of Bizarre" Imagery as a Mnemonic Device Psychological Record, 1970, 20, 145-150 (a).
- PERENSKY, J.J., and SENTER, R.J. The Effect of Subjects' Conforming to Mnemonic Instructions. Journal of Psychology, 1970, 74, 15-20 (b).
- PESKIN, M.E. Interaction of Dialect, SES and Ethnicity Upon the Listening and Reading Comprehension of Fifth Graders. Unpub. PhD Thesis, Rutgers University, 1973.

- PETTIT, N.T., and COCKRIEL, W. A Factor Study of the Literal Reading Comprehension Test and the Inferential Reading Comprehension Test. Journal of Reading Behavior, 1974, 6, 63-75.
- PHILIPCHALK, R.P. Thematicity, Abstractness and the Long-Term Recall of Connected Discourse. Psychonomic Science, 1972, 27, 361-362.
- PHILLIPS, S. The Contribution of Psychology to Education. Oxford Review of Education, 1976, 2, 179-196.
- PIAGET, J., and INHELDER, B. Mental Imagery in the Child, 1971 (Original Version 1966).
- PIMSLEUR, P., and QUINN, T. The Psychology of Second Language Learning, 1971, Cambridge University Press.
- PINES, M.B., and BLICK, K.A. Experimenter - Supplied and Subject - Generated Mnemonics in the Retention of Word Pairs. Psychological Reports, 1974, 34, 99-106.
- POMPI, K.F., and LACHMAN, R. Surrogate Processes in the Short-Term Retention of Connected Discourse. Journal of Experimental Psychology, 1967, 75, 143-150.
- POSTMAN, L. The Effects of Language Habits on the Acquisition and Retention of Verbal Associations. Journal of Experimental Psychology, 1962, 64, 7-19.
- POSTMAN, L. Verbal Learning and Memory. Annual Review of Psychology, 1975, 26, 291-335.
- POTTER, F.W. Representational Modes of Thinking Employed by Children Aged 13 and 14, and Their Relationship To Performance in Standardised Tests of Ability, Measures of Creativity and Personality. Unpub. PhD Thesis, Aston University, 1977.
- PREIBUSCH, W. and ZANDER, H. Vocabulary Teaching - a Search for an Analytical Model. I.R.A.L., 1971, 9, 131-145.
- PRESSLEY, G.M. Mental Imagery Helps Eight-Year-Olds Remember What they Read. Journal of Educational Psychology, 1976, 68, 355-359.
- PRESSLEY, G.M. Imagery and Children's Learning: Putting the Picture in Developmental Perspective. Review of Educational Research, 1977, 47, 585-622.

- PRESSLEY, G.M. Children's Use of the Keyword Method to Learn Simple Spanish Vocabulary Words. Journal of Educational Psychology, 1977, 69, 465-472.
- PRYTULAK, L.S. Natural Language Mediation. Cognitive Psychology, 1971, 2, 1-56.
- PUTNAM, L.R. Improving the Teaching of Comprehension. In MOYLE, D. (Ed.) Reading: What of the Future?, 1975, Ward Lock International.
- PYLYSHAN, Z.W. What the Mind's Eye Tells The Mind's Brain: A Critique of Mental Imagery. Psychological Bulletin, 1973, 80, 1-24.
- RADFORD, J. Verbalization Effects in a "Non-Verbal" Intelligence Test. British Journal of Educational Psychology, 1966, 36.
- RAITT, J.G. The Bullock Report and the Schools Council. Dialogue, 1976, 22, 12-13.
- RANKLIN, J.E., and BARBER, P.J. Peripheral Visual Cues During Reading. Bulletin of the B.P.S., 1978, 31, 186. (Abstract of 1978 B.P.S. Conference Paper).
- 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.
- RAUGH, M.R. and ATKINSON, R.C. A Mnemonic Method for Learning Second-Language Vocabulary. Journal of Educational Psychology, 1975, 67, 1-16.
- REE, H. A Licence to Learn Languages. Address to A.V.L.A. Conference, December 1972.
- REESE, H.W. Imagery and Contextual Meaning. Psychological Bulletin, 1970, 73, 404-414.
- REESE, H.W., and PORGES, S.W. The Development of Learning Processes. In HAMILTON, V., and VERNON, M.D. (Eds.) The Development of Cognitive Processes, 1976, Academic Press.
- REINERT, H. "One Picture is Worth a Thousand Words? Not Necessarily!" Modern Language Journal, 1976, 60, 160-168.

- RICHARDS, I. A. Design for Escape, 1968, Harcourt Brace and World Inc.
- RICHARDSON, A. Mental Imagery, 1969, Springer.
- RICHARDSON, A. Voluntary Control of the Memory Image. In SHEEHAN, P.W. (Ed.) The Function and Nature of Imagery, 1972, Academic Press.
- RICHARDSON, A. The Meaning and Measurement of Memory Imagery. British Journal of Psychology, 1977, 68, 29-43.
- RICKARDS, J.P. Interaction of Position and Conceptual Level of Adjunct Questions on Immediate and Delayed Retention of Text. Journal of Educational Psychology, 1976, 68, 210-217.
- RIDING, R.J. School Learning: Mechanisms and Processes, 1977, Open Books.
- RIDING, R.J., and SHORE, J.M. A Comparison of Two Methods of Improving Comprehension in Educationally Subnormal Children. British Journal of Educational Psychology, 1974, 300-304.
- RIVERS, W.M. Teaching Foreign Language Skills, 1968, University of Chicago Press.
- ROBBINS, D., BRAY, J.F., IRVIN, J.R., and WISE, P.S. Memorial Strategy and Imagery: an Interaction between Instructions and Rated Imagery. Journal of Experimental Psychology, 1974, 102, 706-709.
- ROHWER, W.D., and HARRIS, W.J. Media Effects on Prose Learning in Two Populations of Children. Journal of Educational Psychology, 1975, 67, 651-657.
- ROHWER, W.D. Images and Pictures in Children's Learning: Research Results and Educational Implications Psychological Bulletin, 1970, 73, 393-403.
- ROHWER, W.D., LYNCH, S., LEVIN, J.R., and SUZUKI, N. Pictorial and Verbal Factors in the Efficient Learning of Paired-Associates. Journal of Educational Psychology, 1966, 58, 278-284.
- ROHWER, W.D. Prime Time for Education - Early Childhood or Adolescence? Harvard Educational Review, 1971, 41, 316-341.
- ROHWER, W.D., Elaboration and Learning in Childhood and Adolescence in REESE, H.W. (Ed.) Advances in Child Development and Behavior, 1973, 8, 1-57.

- ROHWER, W.F., and MATZ, R.D. Improving Aural Comprehension in White and in Black Children. Pictures Versus Print. Journal of Experimental Child Psychology, 1975, 19, 23-36.
- ROSENBERG, S. and SIMON, H.A. Modelling Semantic Memory: Effects of Presenting Semantic Information in Different Modalities. Cognitive Psychology, 1977, 9, 293-325.
- ROSS, J., and LAWRENCE, K.A. Some Observations on Memory Artifice. Psychonomic Science, 1968, 13, 107-108.
- ROTHKOPF, E.Z. Some Theoretical and Experimental Approaches to Problems in Written Instruction. In KRUMBOLTZ, J.D. (Ed.) Learning and the Educational Process, 1965, Rand McNally.
- ROTHKOPF, E.Z. Learning From Written Instructive Material: An Exploration of the Control of Inspection Behavior by Test-Like Events. American Educational Research Journal, 1966, 3, 241-249.
- ROTHKOPF, E.Z. The Concept of Mathemagenic Activities. Review of Educational Research, 1970, 40, 325-336.
- ROTHKOPF, E.Z. Barbarism and Mathemagenic Activities. Journal of Reading Behavior, 1974, 6, 3-8.
- ROTHKOPF, E.Z., and BIBISCOS, E.E. Selective Facilitative Effects of Interspersed Questions on Learning from Written Materials. Journal of Educational Psychology, 1967, 58, 56-61.
- ROTHKOPF, E.Z. and BLOOM, R.D. Effects of Interpersonal Interaction on the Instructional Value of Adjunct Questions in Learning from Written Material. Journal of Educational Psychology, 1970, 61, 417-422.
- ROWE, E.J. Discrimination Learning of Pictures and Words: a Replication of Picture Superiority. Journal of Experimental Child Psychology, 1972, 14, 323-328.
- ROWE, E.J., and PAIVIO, A. Discrimination Learning of Pictures and Words. Psychonomic Science, 1971, 22, 87-88.
- SASSON, R.Y., and FRAISSE, P. Images in Memory for Concrete and Abstract Sentences. Journal of Experimental Psychology, 1972, 94, 149-155.

- SCHOOLS COUNCIL. English: a Programme for Research and Development in English Teaching, 1965, HMSO.
- SCHOOLS COUNCIL. Writing Across the Curriculum Project. Why Write?, 1973, Evans/Methnen Educational.
- SEAMON, J.G., and GAZZANIGA, M.S. Coding Strategies and Cerebral laterality Effects. Cognitive Psychology, 1973, 5, 249-256.
- SEMEONOFF, B. (Ed.) Personality Assessment, 1966, Penguin.
- SETER, R.J., and HOFFMAN, R.F. Bizarreness as a Nonessential Variable in Mnemonic Imagery: A Confirmation. Bulletin of The Psychonomic Society, 1976, 7, 163-164.
- SHEEHAN, P.W. Individual Differences in Vividness of Imagery and the Function of Imagery in Incidental Learning. Australian Journal of Psychology, 1971, 23, 279-288.
- SHEEHAN, P.W. A Functional Analysis of The Role of Visual Imagery in Unexpected Recall. In SHEEHAN, P.W. (Ed.) The Function and Nature of Imagery, 1972, Academic Press.
- SEYMOUR, P.H.K. A Model for Reading, Naming and Comparison. British Journal of Psychology, 1973, 64, 35-49.
- SEYMOUR, P.H.K. Contemporary Models of the Cognitive Processes, II. Retrieval and Comparison Operations in Permanent Memory. In HAMILTON, V., and VERNON, M.D. (Eds.) The Development of Cognitive Processes, 1976, Academic Press.
- SHEPARD, R.N., and FENG, C.A. Chronometric Study of Mental Paper Folding. Cognitive Psychology, 1972, 3, 228-243.
- SHEEHAN, P.W. A Shortened Form of Betts' Questionnaire Upon Mental Imagery. Journal of Clinical Psychology, 1967, 23, 386-389.
- SHEPARD, R.N., and METZLER, J. Mental Rotation of Three-Dimensional Objects. Science, 1971, 171, 701-703.
- SHIMRON, J. Imagery and the Comprehension of Prose by Elementary School Children. Unpub. PhD Thesis, 1974, Pittsburgh University.
- SHORR, J.E. The Use of Task Imagery as Therapy. Psychotherapy: Theory, Research and Practice, 1975, 12, 207-210.

- TUINMAN, J.J., and FARR, R. Passage Availability when Answering Passage - Based Questions for Long and Short Reading Selections. Measurement and Evaluation in Guidance, 1972, 4, 234-237.
- TURVEY, M.T. Constructive Theory, Perceptual Systems and Tacit Knowledge. In WEIMER, W.B., and PALERMO, D.S. Cognition and the Symbolic Processes, 1974, Lawrence Erlbaum.
- UNDERWOOD, B.J. Attributes of Memory. Psychological Review, 1969, 76, 559-573.
- UNDERWOOD, B.J. Are we Overloading Memory? In MELTON, A.E., and MARTIN, E. (Eds.) Coding Processes in Human Memory, 1972, V.H. Winston.
- VAN DER VEUR, R.W. Imagery Ratings for 1000 Frequently Used Words. Journal of Educational Psychology, 1975, 67, 43-56.
- VIELAU, A. Cognitive Vocabulary Teaching. Language Teaching and Linguistic Abstracts, 1976, 9, 51-52.
- WALL, S. Readability - a Neglected Criterion in Secondary Textbook Selection. Journal of the Reading Specialist, 1969, 9, 12-16.
- WAMBOLD, C.L., and HAYDEN, C. Training Cognitive Strategies in the Mildly Retarded: an Applied Approach. Education and Training of the Mentally Retarded, 1975, 10, 132-137.
- WASHBURN, J.N. The Use of Questions in Social Science Material. Journal of Educational Psychology, 1929, 5, 321-359.
- WEINER, M., and CROMER, W. Reading and Reading Difficulty: a Conceptual Analysis. Harvard Educational Review, 1967, 37, 620-643.
- WELLS, G. Comprehension: What it Means to Understand. English in Education, 1976, 10, 24-36.
- WHITCOMBE, V. Every Teacher a Teacher of English. English in Education, Spring 1973, 7, 43-56.

- SHORT, P.L. The Measurement of Mental Images. Science News, 1952, 24, 7-21.
- SIEGAL, A.I., LAMBERT, J.V., and BURKETT, J.R. Techniques for Making Written Material more Readable/Comprehensible USAF H.R.L. Tech. Report, August 1974, No. 74-47.
- SIEGEL, S. Nonparametric Statistics for the Behavioral Sciences, 1956, McGraw-Hill.
- SIIPOLA, E.M., and HAYDEN, S.D. Exploring Eidetic Imagery Among the Retarded. Perceptual and Motor Skills, 1965, 21, 275-286.
- SIMMONS, H.W. The School In Schools Council Working Paper No. 12, 1967, H.M.S.O.
- SIMMON, H.A. What is Visual Imagery? An Information - Processing Interpretation. In GREGG, L.W. Cognition in Learning and Memory, 1972, John Wiley and Sons.
- SIMON, J.L. Basic Research Methods in the Social Sciences, 1969, Random House.
- SLOBIN, D.I. Grammatical Transformations and Sentence Comprehension in Childhood and Adulthood. JVLVB, 1966, 5, 219-227.
- SLOBIN, D.I. Psycholinguistics, 1971, Scott, Foresman and Co.
- SMITH, F. Comprehension and Learning, 1975.
- SMITH, F. Making Sense of Reading - and of Reading Instruction. Harvard Educational Review, 1977, 47, 386-395.
- SMITH, N.B. The good Reader Thinks Critically. In DAWSON, M.A. (Ed.) Developing Comprehension, Including Critical Reading, 1971, International Reading Association.
- SNODGRASS, J.G., WASSER, B., and FINKESTEIN, M. On the Fate of Visual and Verbal Memory Codes for Pictures and Words: Evidence for a Dual Coding Mechanism in Recognition Memory. J.V.L.V.B., 1974, 13, 27-37.
- SQUIRE, J.R. Freedom and Constraint. English in Education, 1972, 6.
- STARKER, S. Two Modes of Visual Imagery. Perceptual and Motor Skills, 1974, 38, 649-650.

- STEINFUHRER, F. Some Problems of Vocabulary Work in Teaching Russian. Language Teaching Abstracts, 1973, 6, 198.
- STRANG, R. The Nature of Reading. In MELNIK, A., and MERRITT, J. (Eds.) Reading: Today and Tomorrow, 1972, University of London Press.
- STRATTA, L. Language and Experience: Thoughts on a Rationale for the Teaching of English. English in Education. Winter, 1972, 6, 98-112.
- STROMNES, F.J., and NYMAN, J. Immediate and Long-Term Retention of Connected Concrete Discourse as a Function of Mnemonic Picture - Type Sequence and Context. Scandinavian Journal of Psychology, 1974, 15, 197-202.
- SUTHERLAND, M.B. Everyday Imagining and Education, 1971, RKP.
- SWALM, J.E. Comparison of Oral Reading, Silent Reading and Listening Comprehension Assessed by Cloze. Unpub. EdD Thesis, Rutgers University, 1971.
- SWENSON, I., and KULHAVY, R.W. Adjunct Questions and the Comprehension of Prose by Children. Journal of Educational Psychology, 1974, 66, 212-215.
- TAYLOR, A.M., JOSBERGER, M., and WHITELEY, S.E. Elaboration Instructions and Verbalization as Factors Facilitating Retarded Children's Recall. Journal of Educational Psychology, 1973, 64, 341-346.
- TAYLOR, C.C. A Study of the Nature of Spatial Ability and Its Relationship to Attainment in Geography. British Journal of Educational Psychology, 1960, 30, 266-270.
- THARP, VANK. Mnemonic Techniques, Visual Imagery, and Brain Function. Biological Psychological Bulletin, 1972, 1, 10-20.
- THOMAS, G. The Process of Writing. English in Education, Winter 1972, 6, 74-81.
- THORNDYKE, P.W. Cognitive Structures in Human Story Comprehension and Memory. PhD Thesis, Stanford University, 1975, Rand Paper Series.
- TUINMAN, J.J. Determining the Passage - Dependency of Comprehension Questions in Five Major Tests. Reading Research Quarterly, 1974, 9, 206-233.

- WHITE, K.D., ASHTON, R., and BROWN, R.M.D. The Measurement of Imagery Vividness: Normative Data and Their Relationship to Sex, Age and Modality Differences. British Journal of Psychology, 1977, 68, 203-211.
- WHITFIELD, R.C. Modern Languages - Their Distinctive Contribution in the Context of the Whole Curriculum Modern Languages in Scotland (1975) No. 6, 22-30.
- WILBY, P. University Challenge. Sunday Times 14/8/1977, 46.
- WILBY, P. O level French is Useless for Survival. Sunday Times 29/1/1978.
- WILDER, L., and LEVIN, J.R.A. Developmental Study of Pronouncing Responses in the Discrimination Learning of Words and Pictures. Journal of Experimental Child Psychology, 1973, 15, 278-286.
- WILLIAMS, J.D. Teaching Problem-Solving. Educational Research, 1960, 3.
- WILLIAMS, R.C. Problems in Language Learning. Trends in Education (1976) no. 1, 13-18.
- WILSON, R.M. Comprehension Diagnosis. Reading World, 1974, 14, 49-50.
- WINITZ, H., and REEDS, J.A. Rapid Acquisition of a Foreign Language (German) by the Avoidance of Speaking. I.R.A.L., 1973, 11, 295-317.
- WITTROCK, M.C. The Generative Processes of Memory. UCLA Educator, 1975, 17, 33-43.
- WITTROCK, M.C. and GOLDBERG, S.I. Imagery and Meaningfulness in Free Recall: Word Attributes and Instructional Sets. Journal of General Psychology, 1975, 92, 137-153.
- WOLFF, P., and LEVIN, J.R. The Role of Overt Imagery in Children's Imagery Production. Child Development, 1972, 43, 533-547.
- WOLLEN, K.A., WEBER, A., and LOWRY, D.H. Bizarreness Versus Interaction of Mental Images as Determinants of Learning. Cognitive Psychology, 1972, 3, 518-523.
- WOLPERT, E.M. Length, Imagery Values and Word Recognition. Reading Teacher, 1972, 26, 180-186.

- WOOD, G. Mnemonic Systems in Recall. Journal of Educational Psychology Monographs, 1967, 58, (no. 6 Part 2).
- WRIGHT, P. Reading to Learn. In MELNIK, A., and MERRITT, J. (Eds.) Reading: Today and Tomorrow, 1972, University of London Press.
- YANOFF, J.M. The Functions of the Mind in the Learning Process. In SILBERMAN, M.L., ALLENDER, J.S., and YANOFF, J.M. (Eds.) The Psychology of Open Teaching and Learning, 1972, Little, Brown and Co.
- YARMEY, A.D., and BOWEN, N.V. The Role of Imagery in Incidental Learning of Educable Retarded and Normal Children. Journal of Experimental Child Psychology, 1972, 14, 303-312.
- YATES, F.A. The Art of Memory, 1966, Peregrine.
- YOUNG, R.K., OVERBY, G., and POWELL, G.D. Is There Sequential Information in a Mental Image? Journal of Experimental Psychology, 1976, 2, 663-670.
- YUILLE, J.C. A Detailed Examination of Mediation in Paired-Associate Learning. Memory and Cognition, 1973, 1, 333-342.
- YUILLE, J.C., and CATCHPOLE, M.J. Associative Learning and Imagery Training in Children. Journal of Experimental Child Psychology, 1973, 16, 403-412.
- YUILLE, J.C., and CATCHPOLE, M.J. The Effects of Delay and Imagery Training on the Recall and Recognition of Object Pairs. Journal of Experimental Child Psychology, 1974, 17, 474-481.
- YUILLE, J.C., and PAIVIO, A. Abstractness and Recall of Connected Discourse. Journal of Experimental Psychology, 1969, 82, 467-471.
- ZUPNICK, J.J., and MEYER, P.A. Long-Term Effectiveness of Imagery Instruction with Retarded Persons. American Journal of Mental Deficiency, 1975, 79, 519-525.

APPENDICES

APPENDIX A

(MISSING)

THE QUESTIONNAIRE

APPENDIX B

CONNECTING IMAGES SUPPLIED BY E (PILOT VOCABULARY STUDY),
IN ORDER OF PRESENTATION.

1. "Lunettes.....spectacles.....net."
"Imagine a football net with a pair of spectacles caught up in it, left there by the goalkeeper".
2. "Tapisserie.....tapestry.....tap".
"Imagine a tapestry hanging on a wall in a stately home, with taps embroidered all over it".
3. "Imprévu.....unexpected.....pray".
"Imagine you are praying for something unexpected to happen - for instance, for it to snow in the summer".
4. "Matelot.....sailor.....mat".
"Imagine a sailor floating on the sea on a mat".
5. "Loupe.....magnifying-glass.....loop".
"Imagine a tiny loop in a piece of string, so small you need a magnifying glass to see it".
6. "Canapé.....settee.....can".
"Imagine a huge can of baked beans sitting on a settee".
7. "Lapin.....rabbit.....lap".
"Imagine a giant rabbit, and you sitting on its lap".
8. "Côteau.....hillside.....cot".
"Imagine a cot sliding down a hillside".
9. "Baguette.....rod.....bag".
"Imagine a burglar carrying his bag. Sticking out of it is the iron rod he uses for breaking into houses".
10. "Batelier.....ferryman.....bat".
"Imagine a ferryman using a cricket bat as an oar to ferry his passengers across a river".
11. "Pample mousse.....grapefruit.....moose".
"Imagine a moose charging round trying to shake a grapefruit off his antlers".
12. "Pataud.....clumsy person.....pat".
"Imagine a clumsy person about to pat a friend on the back while standing on the edge of a cliff, missing and falling over".

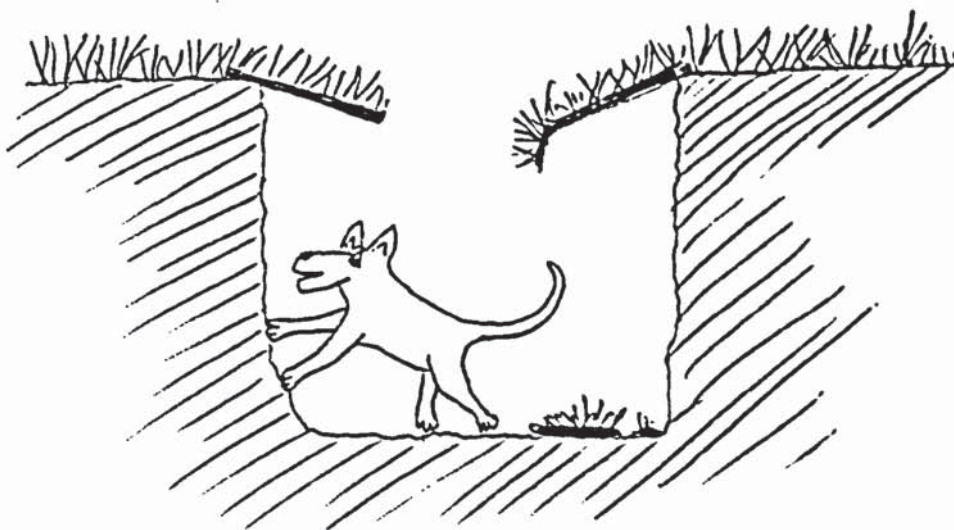
APPENDIX C

CONNECTING IMAGES AND PICTURES SUPPLIED BY E (MAIN VOCABULARY STUDY), IN ORDER OF PRESENTATION.

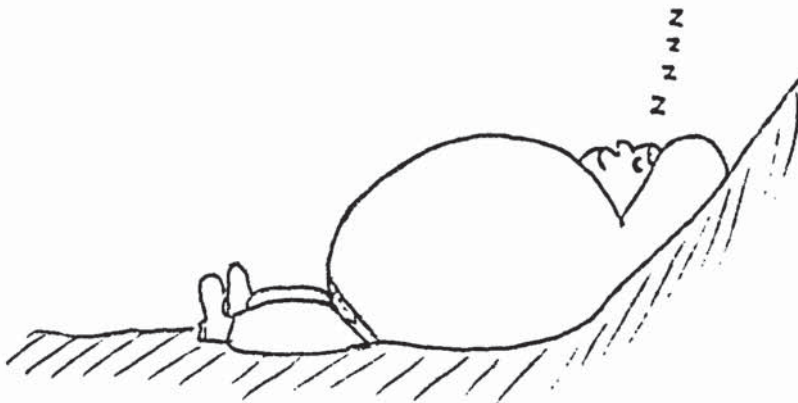
1. "Tard.....late.....tar".
"Imagine some black sticky tar in the road. You're going to school, but you get stuck in the tar, and arrive very late".



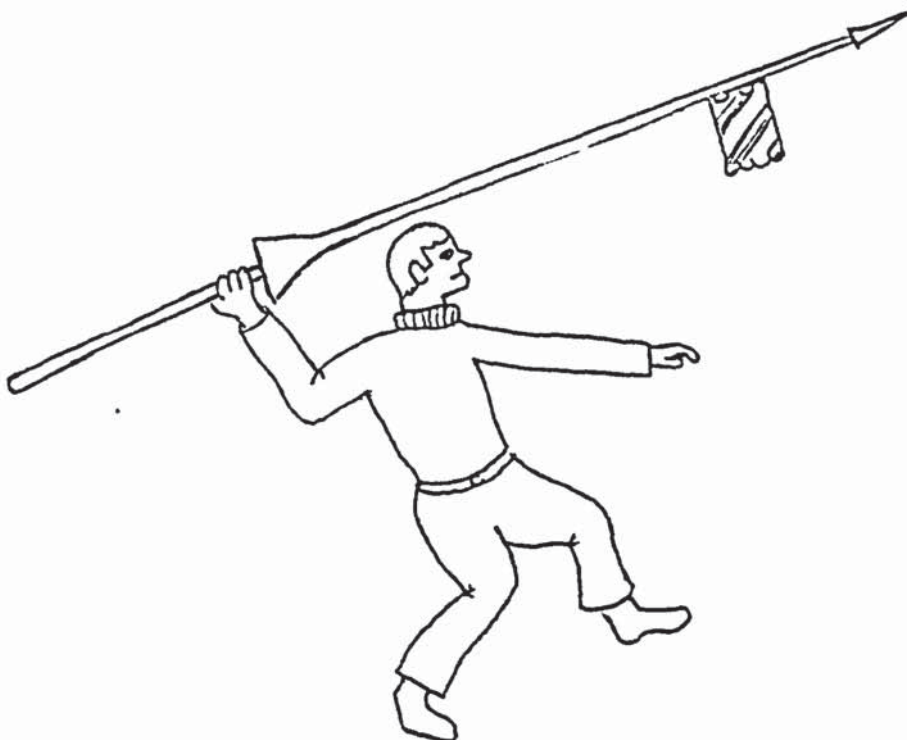
2. "Attraper.....to catch.....a trap".
"Imagine a trap hidden in the jungle to catch some wild animal".



3. "Fatigué.....tired.....fat".
"Imagine a very fat person, lying asleep by the road because he's so huge that he's tired out from walking a few yards".



4. "Lancer.....to throw.....lance".
"Imagine a big lance or spear like knights in armour used to carry, and you're trying to throw it"



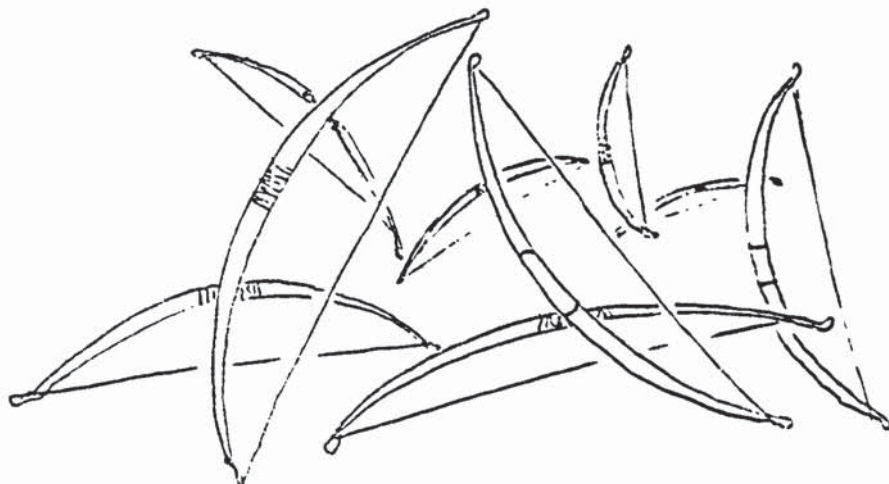
5. "Sage.....good.....sergeant".
"Imagine a police sergeant - everybody calls him sarge.
He's really kind and good - so good that he has a halo."



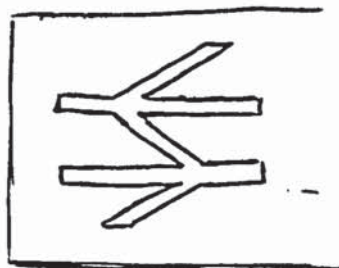
6. "Manger.....to eat.....manager".
"Imagine a bank manager with bowler hat and striped suit,
sitting at his desk. He's about to eat a bag of fish and
chips".



7. "Beaucoup.....many.....bow".
"Imagine a bow, then another one, and another and another until there are lots of them - many of them all over the place".



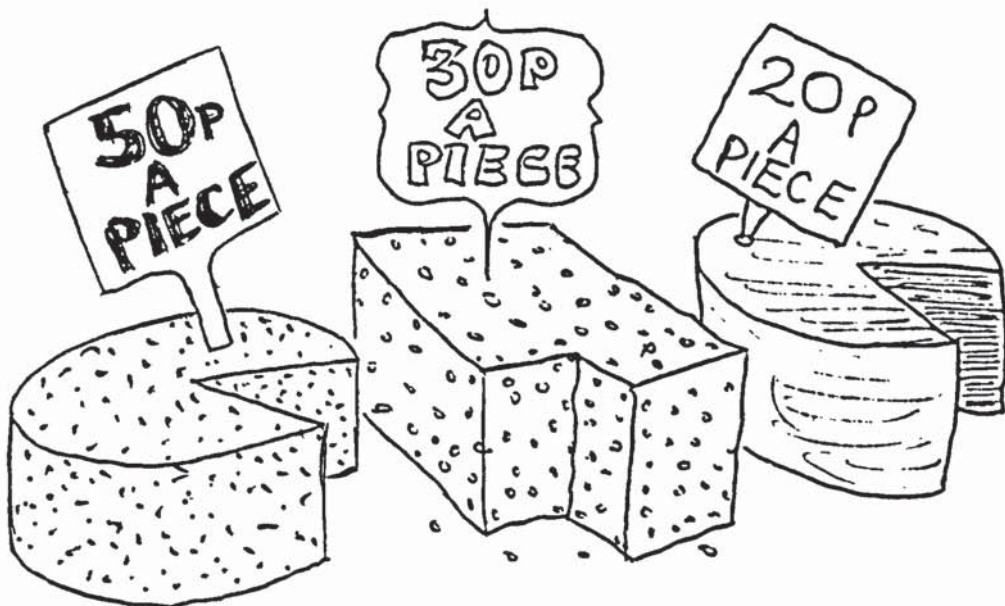
8. "Porter.....to carry.....porter".
"Imagine a porter at a railway station. He has to carry all sorts of boxes, cases etc. Imagine him staggering along, carrying them all".



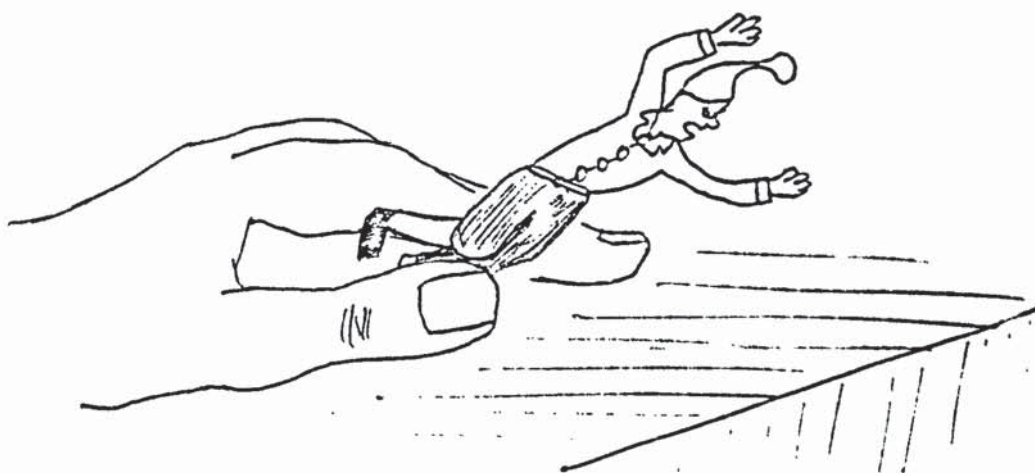
9. "Tasse.....cup.....ass".
"Imagine an ass or donkey. He's sitting up politely drinking from a large tea cup".



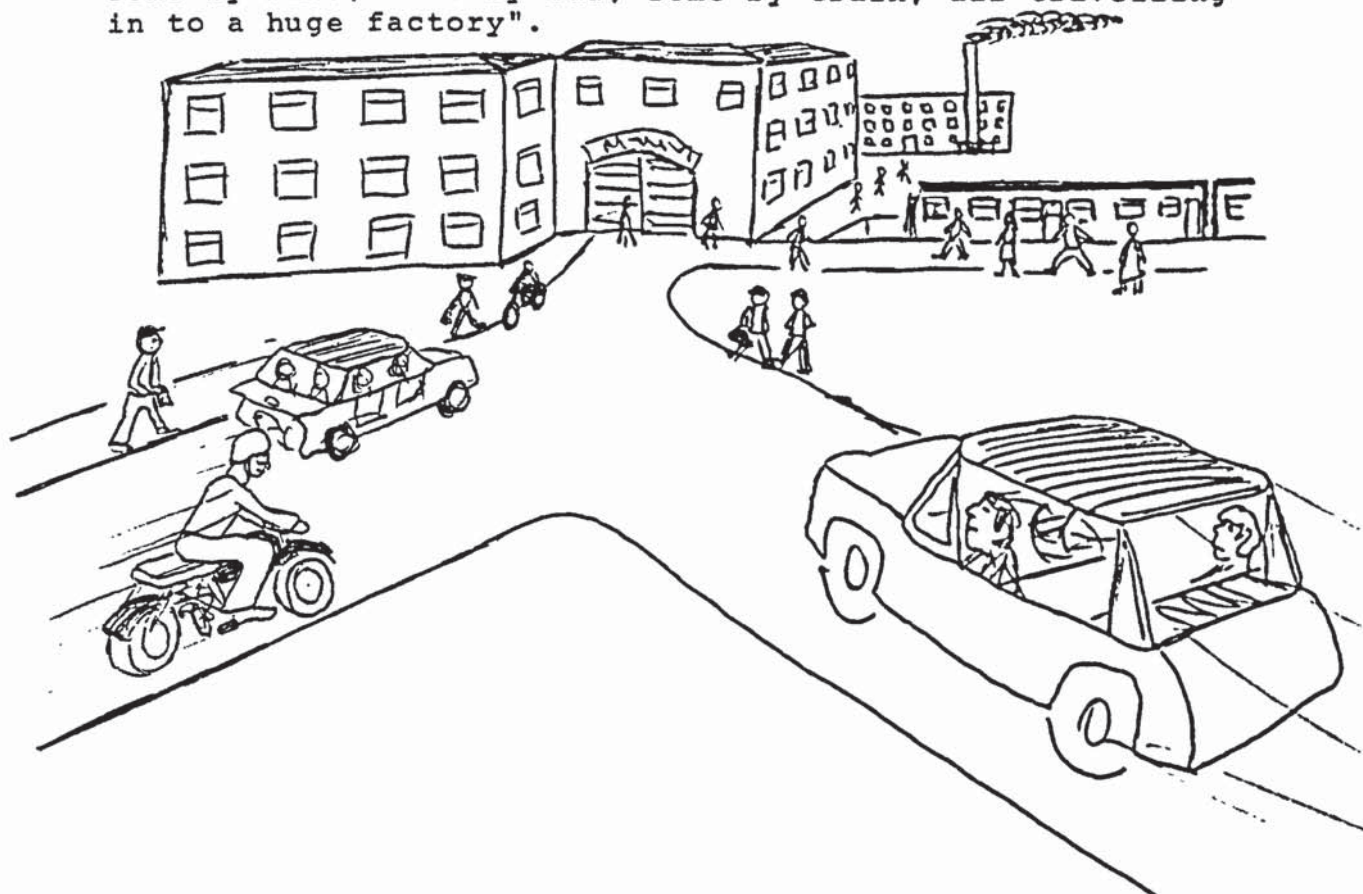
10. "Epicerie.....grocer's.....a piece".
"Imagine the window of a grocer's shop. In it you can see something with a big label on saying 50p a piece, an other label 30p a piece, another 20p a piece".



11. "Tomber.....to fall.....Tom".
"Imagine tiny Tom Thumb. He's so small that he walks across your desk and falls over your thumb flat on his face".



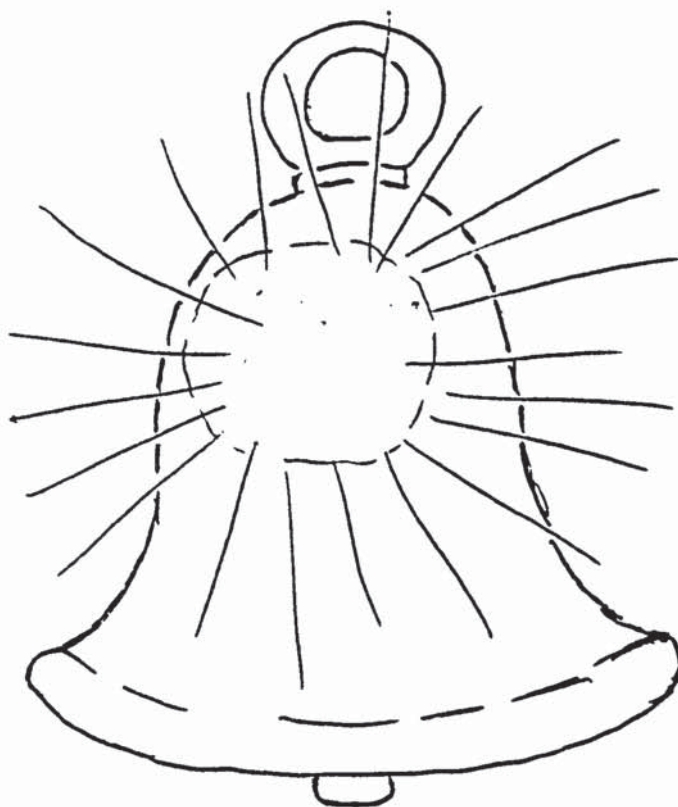
12. "Travailler.....to work.....travel".
"Imagine lots of people travelling to work, some by bus, some by bike, some by car, some by train, all travelling in to a huge factory".



13. "Trouver.....to find.....trove".
"Imagine you're digging for treasure trove and find a chest full of treasure. So you find a treasure trove".



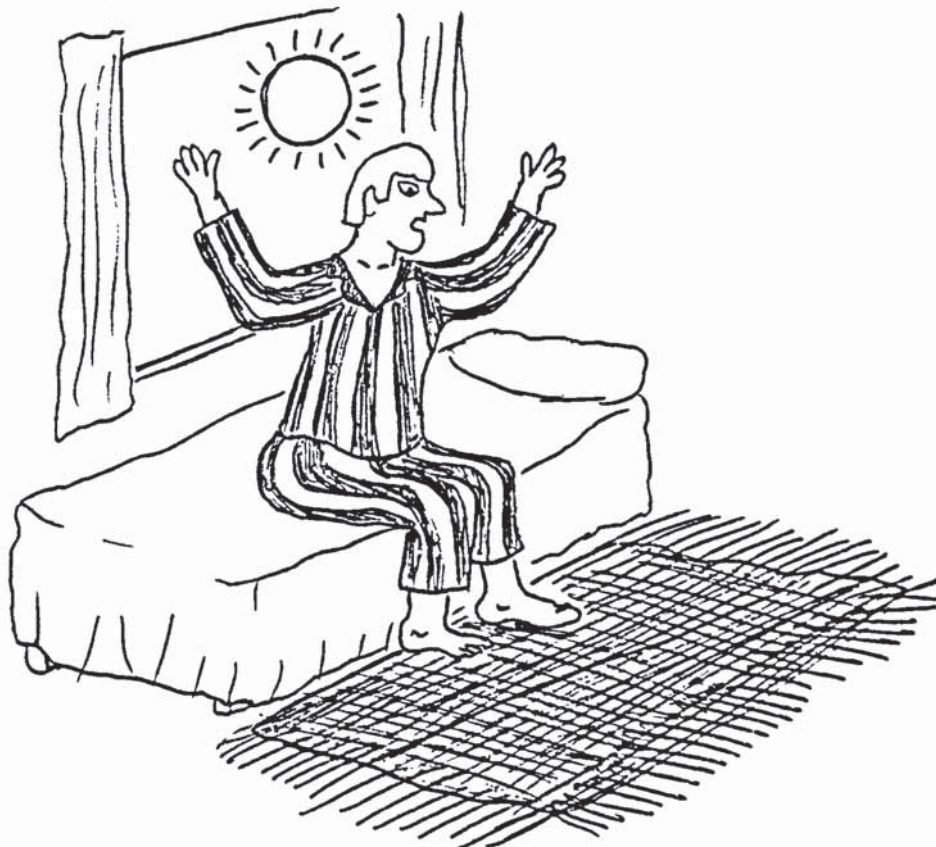
14. "Sonnerie.....bell.....sun".
"Imagine a big shiny bell, with the sun reflected in it, dazzling you".



15. "Ami.....friend.....me".
"Imagine you yourself - in other words "me" with your best friend at breaktime. So think of me and my friend".



16. "Matin.....morning.....mat".
"Imagine getting up in the morning. By your bed is a great big hairy mat, and you put your feet on it when you get up".



17. "Vilain.....naughty.....villain".
"Imagine a villain, with a top hat and a black cloak,
playing for Aston Villa. But he's really naughty and
keeps fouling the other side, so the ref. sends him off".



APPENDIX D

THE COMPREHENSION PASSAGE.

(Adjunct questions printed below each Section. Each question appeared on a separate page in the actual booklet).

The light April breeze rippled the prairie grass under the plodding hoofs of the bay horse. It touched the sober, wind-browned face of the boy in the saddle, and riffled the shaggy coat of the black-and-white dog trotting near the left stirrup.

Joe Beckett felt the warmth of the spring sun fading into cool dusk and, with an experienced eye, measured the distance to the timbered land ahead, which fanned out from the village of Council Bluffs on the Missouri River.

'Not more than a mile away, Charley. We'll make it before dark'.

1. On what river is the village of Council Bluffs?

The dog looked up at the boy and seemed to nod without breaking his effortless stride. Even when a rabbit scampered across the trail, Charley ignored any foolish chase.

Joe offered no word of praise for such behaviour. After all, Charley was a well-trained, working sheep dog who had learned early that only the flock was his responsibility. To the grey-backed sheep, Charley gave his loyal allegiance. To his master, he gave his heart, with lavish affection.

The boy glanced at the compact, wiry Border collie with a new feeling of kinship, suddenly aware that he now shared Charley's responsibility.

2. What breed of dog is Charley?

He pulled up the bay and looked back at the rolling plains, where he had left four thousand sheep in the care of Ab Cramer and three sheep-herders he didn't trust. In the year of 1853, there was little reason to trust any stranger on the frontier. The gold rush in California was still at fever pitch. Men would lie, steal, and even kill to reach that promised land where they were certain that fortunes could be made overnight.

A few years ago Joe had wondered why Alex Macleod, who had taken him from an orphanage to learn sheep-herding, had not rushed off with the others from Boston, Philadelphia, and New York to strike it rich. But the canny old Scot had explained in his own gruff way.

3. What year was it?

'Aye, we will go to California, lad, but not now. Let those in a hurry for money get there first. Supplies will grow short. Prices will go up. Soon there will not be enough beef or mutton to feed a mob of greedy men. While they are fighting over claims and getting hungry, we will gather up sheep from Vermont to Illinois. From the Missouri River we will walk them to California and sell at a high price. We can make a fortune with no need of pick or shovel. Care for the flock well, lad. They walk on golden hoofs.'

4. Where would they gather up the sheep?

Remembering this wise plan, Joe jerked himself back to the grim reality of the present. Alex Macleod was dead. Two days out of Council Bluffs, the durable old Scot had seemed to know his summer cold was the quick pneumonia of the frontier. He had called Joe into the wagon and had given him an oilskin-wrapped letter.

'Go back to Council Bluffs and get my old friend, Jennings, to be your flockmaster,' he had ordered. 'Walk the flock west, and deliver this letter and the sheep to Mr James Forbes in Sacramento. Get over the mountain before snow flies. Dinna grieve. The flock needs tending.'

5. Where did Mr James Forbes live?

The rest had been only a blur of sadness, a new grave, the uncertainty of the herders, the restlessness of the flock held too long on the bedding ground, and the whimper of a dog who waited by the empty sheep wagon.

Absently Joe dismounted and poured water into his battered wool hat for Charley to lap. Then he finished the rest of the water in the bottle and stretched his legs, standing in the dusty tracks made by a dozen spring wagon trains.

Once more he looked back over the treeless land to the west. He thought of the four thousand sheep, so dependent on guidance. He thought of the fourteen hundred miles to Sacramento, the rivers to cross, the mountains to climb, and the deserts to endure. And he thought of the Indians who attacked wagon trains.

6. What was Joe's hat made of? .

Suddenly the grief and loneliness tore at him like a cold wind, until he felt Charley's rough tongue lick his hand. He looked down at the sad, brown eyes of the collie and knew why the dog had been made his to train, to command, and to love. He ran a calloused palm over the white spot between the cocked black ears.

'I miss him, too, Charley, but Alex wouldn't want us to grieve. Come on. Let's find Mr Jennings. Once I get a flockmaster, I won't feel like I've got the weight of the world on my shoulders.'

Charley, seeing his master's face lighten, wagged his tail in approval as Joe mounted his horse and rode on toward the town by the river.

7. What did Joe feel on his hand?

APPENDIX E
THE TEST QUESTION SHEET

TEST QUESTIONS

(DO NOT WRITE ANYTHING ON THIS SHEET)

1. What time of day was it?
2. How can we tell that Joe was used to travelling on the prairie?
3. What did Charley do when he saw the rabbit?
4. Why did he do this?
5. How many men had Joe left to look after the sheep?
6. Why didn't Joe trust any strangers?
7. What sort of people would get to California first?
8. Why would the sheep be worth more in California?
9. What did Alex give Joe?
10. Alex knew he was dying. How can we tell this?
11. Name two dangerous things that Joe would have to cross on the journey to California.
12. How did Joe know that many wagons had already passed that way?
13. What did Charley do when Joe got back on his horse?
14. How would Joe feel once he found Mr Jennings?

APPENDIX F

MARKING SCHEMES

LITERAL COMPREHENSION QUESTIONS

1. (What time of day was it?)
Dusk, evening, any actual time between 5-8p.m. - 2.
Late afternoon - 1.
Afternoon or any other answer - 0.
3. (What did Charley do when he saw the rabbit?)
Nothing, ignored it, carried on etc. - 2.
Any other answer - 0.
5. (How many men had Joe left to look after the sheep?)
Four, Ab Cramer and three others - 2.
Any other answer, including "3 or 4" - 0.
7. (What sort of people would get to California first?)
Greedy ones; "those in a hurry for money" etc. - 2.
Any other answer - 0.
9. (What did Alex give Joe?)
A letter or message - 1.
Mention of oilskin wrapping, or any details about the content of the letter - 1. (max. 2)
11. (Name two dangerous things that Joe would have to cross on Rivers, mountains, deserts - 1 each for any 2.
Indians - 0, but accept "Indian country" - 1.
13. (What did Charley do when Joe got back on his horse?)
Wagged tail - 2.
Any other answer - 0.

INFERENCE COMPREHENSION QUESTIONS

2. (How can we tell that Joe was used to travelling on the prairie?)
He knew how far it was to the village - 1.
He knew how long it would take - 1.
Idea (or quotation) about measuring distance "with an experienced eye" - 2.
Any other answer, including vague ones about knowing his way around or knowing the dangers - 0.

4. (Why did he do this?)
Idea of Charley being well - trained - 1.
Sheep as his responsibility - 1.
Any other answer, including Joe's disapproval - 0.
6. (Why didn't Joe trust strangers?)
Mention of California and gold rush - 1.
Men therefore greedy, would do anything to get there - 1.
Any other answer, including theft of sheep - 0.
8. (Why would the sheep be worth more in California?)
Idea of supply and demand, or of high prices because of shortage - 2.
Only part of above (e.g. prices higher there, but no real explanation of why) - 1.
Any other answer - 0.
10. (Alex knew he was dying. How can we tell this?)
Idea that Alex told him to get a flockmaster to take over - 2.
Reference (or quotation) about Alex knowing his summer cold was pneumonia - 1.
Telling Joe not to grieve - 1 (max. 2).
Any other, including vague references to his illness - 0.
12. (How did Joe know that many wagons had already passed that way?)
Any reference to their tracks in the dust - 2.
Any other answer, including seeing wrecks or hearing of Indian attacks - 0.
14. (How would Joe feel once he found Mr. Jennings?)
Idea of relief; reference or quotation about "the weight of the world on his shoulders" - 2.
More vague answers - happy, glad, better etc. - 1.
Possibly mention sadness about telling Mr. Jennings of the death of Alex - 1 (max. 2).
Any other answer - 0.

ADJUNCT QUESTIONS

1. (On what river is the village of Council Bluffs?)
Missouri (or near spelling with same sound) - 2.
Any other answer - 0.

2. (What breed of dog is Charley?)
Collie or Border Collie - 2.
Sheepdog - 1.
Any other answer - 0.
3. (What year was it?)
1853 - 2.
Nineteenth century or any date in it - 1.
Any other answer - 0.
4. Where would they gather up the sheep?
Mention Vermont (or very similar name) - 1.
Mention Illionois (or very similar name) - 1.
Any other answer - 0.
5. Where did Mr. James Forbes live?
Sacramento (or very similar name) - 2.
California - 1.
Any other answer - 0.
6. What was Joe's hat made of?
Wool; sheeps wool - 2.
Any other answer - 0.
7. What did Joe feel on his hand?
Idea of Charley licking him (e.g. "Charley's tongue") - 2.
Any other answer - 0.

APPENDIX G

EASIEST AND HARDEST WORDS, VOCABULARY EXPERIMENT, KEYWORD ("IMAGE") GROUP.

% Subject with correct Recall	English Word	T-L Value	Image Value	Key-Word	T-L Value	Image Value	Bizarreness Rating
94%	naughty	13	-	villain	11	-	2.5
92%	friend	AA	5.4	me	AA	4.8	0.5
92%	late	AA	3.3	tar	10	-	1.7
90%	fall	AA	5.0	Tom	A	4.8	2.2
86%	tired	AA	4.6	fat	AA	6.0	1.5
69%	grocer's	11	6.0	piece	AA	3.9	0.5
69%	good	AA	3.1	sergeant	18	-	1.8
63%	throw	A	4.6	lance	16	-	0.2
59%	many	AA	3.6	bow	A	5.4	1.2
49%	cup	AA	5.8	ass	14	-	2.8
41%	find	AA	2.6	trove	-	-	0.8

T-L Value - frequency according to E.L. Thorndike and I. Lorge's The teacher's word book of 30,000 words, 1944, Bureau of Publications, Teachers College.

Imagery Value - Van Der Veur (1975).

Bizarreness Rating by six independent judges rating the image descriptions (Appendix C) on a four point scale. 0 = "normal, ordinary", 1 = "rather unusual", 2 = "very unusual", 3 = "really ridiculous or bizarre".