STRUCTURE AND ACCESS:

THE ROLE OF STRUCTURAL FACTORS IN TEXT COMPREHENSION AND INFORMATION ACCESS

CLAIRE EDNA O'MALLEY
B.A.

Submitted in accordance with the
requirements for the degree of Ph.D.

Computer Based Learning Unit
School of Education
The University of Leeds

September 1985

ABSTRACT

In this thesis it is argued that structural factors play an important role in facilitating the access, comprehension, and recall of textual information, especially when the content of the material is unfamiliar to the reader. A study was made of the effects of manipulating text structure, familiarity of subjects with the text type, familiarity with the content, and instructions given to subjects, on comprehending and recalling information from scientific research reports. The results show that subjects familiar with the text type are able to make use of structure as an encoding strategy, and that the use of this structural strategy improves comprehension and recall when the content is unfamiliar. The study suggests that teaching readers to make use of structure in processing text can facilitate comprehension and recall. These results provide support for previous theories concerning the role of text structure, most of which has focused on narratives, to the neglect of research on expository prose. It is argued that some of the problems involved in the research using narratives, in particular, the problem of the lack of distinction between structural factors and more general knowledge of the content, may be obviated by research with other text types, such as the one used in this study.

It is also argued that users of computer-based documentation systems need similar kinds of structural cues for accessing online information as they do for offline text comprehension and retrieval, and that the efficacy and type of such structural cues depends on several factors, such as the task requirement, as well the level of experience of the user. A second study examined the patterns of use of an online documentation system, and showed that users need different forms of organisation of the information as "access structures", depending on different task requirements. Finally, proposals are made for improving the design of online documentation systems and for conducting future research into the needs of users of such systems.

ACKNOWLEDGEMENTS

I would like to thank Roger Hartley, my advisor, for his support and guidance in developing the research described here. Special thanks are due also to Pat Greenwood, for her help and encouragement while I was with the Computer Based Learning Unit, and to Ken Tait and Andrew Cole for assistance with computing facilities. I benefitted also from discussions with Peter Jackson and Rachel Rymaszewski.

I am indebted particularly to Don Norman for giving me the opportunity to work at the Institute for Cognitive Science, University of California, San Diego. His advice and guidance, and the experience I gained in being at ICS, and in working with the Human-Machine Interaction group there were invaluable. Several people from the HMI project gave me many useful ideas, advice, and help. Some of the studies reported in this thesis were conducted in co-operation with members of that project, namely: Liam Bannon, Eileen Conway, Jan Graham, Lissa Monty, Paul Smolensky, and Jeff Sokolov (cf. O'Malley, Smolensky, Bannon, Conway, Graham, Sokolov & Monty, 1983; Bannon & O'Malley, 1984); also Steve Draper and Mary Riley (cf. O'Malley, Draper & Riley, 1984; Riley & O'Malley, 1984). Thanks to all these people for their co-operation and support. I would also like to thank Jean and George Mandler for their advice and support; I benefitted a great deal from working with Jean Mandler and Tamar Murachver. I owe a special debt of gratitude to Sondra Buffett for her help, advice, and encouragement.

This research was supported partly by the Economic and Social Research Council of Great Britain (Award No. H00428022035) while I was at the Computer Based Learning Unit, University of Leeds, and partly sponsored by grants from the Personnel and Training Research Programs of the Office of Naval Research (Contract N00014-79-C-0323, NR 667-437) and the System Development Foundation, while I was at the Institute for Cognitive Science, University of California, San Diego, USA.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION AND OVERVIEW OF THE THESIS

There is an increasing need for developing efficient and effective means of accessing large stores of information, whether the medium is the printed page or the computer. The practical problem addressed in this thesis concerns how to improve existing approaches to and methods for the design of online computer documentation or help systems. It is argued that the design of effective help systems requires not only an adequate task analysis and the provision of task-based help, but also an analysis of the needs of the user and the provision of user-centred help. One of the major problems facing users of online help systems concerns getting quick and easy access to the information they need. Such systems can be improved by the design and provision of better "access structures" (Waller, 1979, 1982).

It has been noted that people tend not to use manuals, whether online or offline, but prefer to go to other people for help. This is not just because the documentation is hard to understand or use, but also because it is hard to get access to the information quickly and easily. Users often don't know how to ask for help — not simply because they don't know the commands to use, but because they don't know how to formulate the appropriate queries in the first place. Users' problems are usually not stated in forms that match the ways the information is organised and accessed. Documentation systems are needed that provide flexible access structures, organised according to users' tasks, rather than system facilities, and according to users' needs and queries. These vary not only with level of expertise, but also with the kinds of tasks that users perform.

It is maintained that the use of structures for accessing information online is analogous to the use of structural cues for accessing information in printed text. In fact, Kommers (1984) has argued that putting information online turns text comprehension into information retrieval.

The thesis approaches this problem, therefore, by asking:

a) What are the structural properties of texts that serve to organise the content and make it a coherent unit of discourse?
b) Can readers use their knowledge of these structural properties to help them comprehend and remember the content of the text, and can such structural strategies be taught to readers unfamiliar with the content and the form of the text?
c) Could designers use similar kinds of "access structures" to improve computer-based text retrieval and comprehension — specifically online documentation? What kind of access structures are needed for using such documentation?

Thus, there are two main themes in this thesis. The first concerns the structural factors influencing the comprehension of printed text: what they are, and at what level of processing they operate. The second theme concerns the kind of "access structures" that are needed in the design, organisation, and presentation of computer documentation — in particular, online documentation — in order to cater for a variety of users with different information requirements.

Under the first theme, this research focuses on the global structures emergent from lower-level units that constitute a text (i.e., sequences of sentences) and that serve to make the text a coherent unit of discourse. Such texts can have structure at several levels: At the highest or global level there are structures that distinguish different text types (e.g., the conventional form of stories, technical manuals, scientific reports, and newspaper articles). These "superstructures" act as frameworks that organise the global content of the text, by virtue of the author's communicative or rhetorical purpose. For example, the author's goal in writing a research report is to persuade the audience to accept the argument being presented. The text, therefore, consists of various claims and evidence relating to the major hypothesis. This becomes formalised in a scientific report, in the organisation of hypotheses, predictions, procedures, results, discussion, and conclusions.

At a "lower" level, there are structures connecting topics that emerge as themes or "threads of discourse". These "macrostructures" are somewhat similar to summaries of the main points of the text, for example, a sequence of sentences forming the author's major hypothesis, or a collection of acts and events forming a whole episode in a story. At yet another level there are structures that emerge from the "micropropositions" of the text to form local sentential topics, connected by referential overlap.

Readers are generally able to produce summaries or abstracts of a text that agree considerably even after one reading, despite differences in their background knowledge of the content. Researchers have attempted to account for this fact by proposing that there are certain specific features or structures inherent in the text (either implicitly or explicitly signalled) that influence comprehension, and that readers use these features, together with their knowledge or schema of the text type, to guide processing and so organise the content of the text in memory, and to retrieve information from their mental representations of the text.

This thesis begins, therefore, with a comprehensive review of theories of text structure (Chapter 2), and theories concerned with the nature of the relationship between form and content (i.e., schematic and semantic structure) in a coherent text (Chapters 2, 3, and 4). A particular text type is chosen for analysis — the scientific report. This analysis provides a framework for the design of experimental materials used in the first study (Chapter 5).

A review is also presented of the psychological evidence for the role of structural factors in text comprehension. The function of structure is discussed with respect both to the "text-driven" aspects of processing (Chapter 3), and the to role of prior knowledge ("conceptually-driven" aspects) and inferential processes involved in establishing a coherent representation of the text in memory (Chapter 4).

Psychological research on text comprehension has focused predominantly on one particular text type: narratives — more specifically, stories from the oral tradition. One advantage of studying this text type is that it has a considerable degree of generality, both within and across cultures. However, recent criticisms of the story grammar approach point out that the evidence supporting this theory might be equally well accounted for by the fact that readers share similar structures in their understanding of causal and conventional event sequences. Thus, it is argued that story grammar studies are primarily examining more general phenomena, such as the acquisition and use of event and action schemas, rather than referencing text-specific structures.

There are several problems in using stories to try and separate out the effects of text-specific structure (form) from more general world knowledge (content). One of the practical problems is that the text-specific knowledge cannot be easily manipulated, since it is difficult to find cultures unfamiliar with the conventional structure of such texts. Therefore, any manipulations have to involve the text itself, rather than the background knowledge of the reader. However, there are other text types familiar only to certain subcultures — one example is the scientific research report.

The structure of a scientific article (e.g., a psychology research report) is a subset of the genre of argumentative texts. That is, the content is organised according to various claims and evidence, and becomes specialised in the case of a particular subset of these texts — the research paper. It is proposed that the canonical form of this type of text plays a similar role in comprehension as does the "story schema" in comprehending stories.

A study is described that was carried out in order to investigate the effects of manipulating the superstructure of the text, the familiarity of subjects with the text structure and content, and the instructions given to subjects, on the comprehension and recall of the text. The text type chosen for this study was the scientific research report. Specifically, it was predicted that knowledge of the conventional structure of the text would influence the ease of comprehension and memory for the content. The results of this study showed that readers familiar with the text structure use this knowledge to guide their processing of the text, in the form of an encoding strategy. The effect of manipulating structure was much greater for subjects familiar with this type of text (and who were therefore presumably using this structural knowledge to encode the text) than for those who were unfamiliar with such materials.

This research on printed text is extended in the second part of the thesis to online information — specifically, computer documentation associated with the UNIX[1] operating system (Chapter 6). As people come to rely more on information technology, issues concerning the access and presentation of online text become very important. One immediate research concern is the presentation and organisation of online computer documentation, especially as the types of user, and their levels of experience become more diverse. Issues concerning organisation and structure of text become crucial in online systems, since most users at the present time have only a 24-line by 80-character "window" onto the text. They cannot use the same strategies with this medium as they would with printed text — such as scanning, flicking pages, and so on — in order to get an overview of the organisation of the information. Therefore, such "access structures" have to be designed and provided in a form that is appropriate to this particular medium. Text comprehension thus becomes information retrieval, and the same sorts of structural and organisational issues raised in research on text comprehension become applicable in this domain.

Accordingly, the thesis presents a discussion of the applications of research on text comprehension to the design, organisation, and presentation of online computer documentation. It is proposed that lack of familiarity with the content of the material can be made up for to some extent by the provision of "access structures", which may be explicit or implicit. A discussion is presented concerning the different kinds of structure that users need in order to access information online. A study is described that was conducted in order to investigate the use made of an online documentation facility. This revealed that users need different forms of organisation of the information, depending not only on their level of experience, but also on the task for which the information is required. Another study is described that was conducted in order to investigate how users explain system concepts to each other when solving joint problems. The study indicated that different kinds of information are needed because users have different models with which they "view" their problems.

In summary, this research provides an extension of structural analyses of texts that have previously focused on narratives, to the analysis of scientific and technical texts. The research also provides some suggestions, based on empirical studies, for

1. UNIX is a trademark of Bell Laboratories.

the design of "access structures" in computer documentation, that help the user to access, comprehend, and use the information. It is argued that these proposals should be used in the design of intelligent help systems, which are likely to be an important development in the future.

CHAPTER 2

THEORIES OF TEXT STRUCTURE

This chapter presents a brief outline of the background to current theories of text structure. There were two major threads of influence: one from linguistics, the other from literary theory and anthropology.

Text Structure and Textlinguistics

It was not until the 1960's that the text became an object of linguistic analysis, rather than simply a source of data for it. However, the early text theories were simply extensions of sentential analysis. In fact, for some of the early theorists, texts were treated as little more than long sentences, analysable via adequate sentence grammars (e.g., Katz & Fodor, 1963). Other theorists, particularly in Europe, believed that the existing transformational and logical linguistic theories of the sentence should be altered in order to apply to texts (e.g., Petöfi, 1971; Van Dijk, 1972). It was realised not only that syntactic and semantic structure, and the interpretation of sentences, should be studied relative to other sentences in the discourse, but also that discourse has a specific overall structure that cannot be formalised in terms of sentence structure alone.

These early European text theories developed along two main lines: one characterised by "descriptive structural linguistics", based upon principles of "functional sentence perspective" and "communicative dynamism" (cf. Morgan & Sellner, 1980), and the other characterised by "generative text grammars". The *structuralist* view was of the text as a unit above the sentence. A distinction that was recognised between rhematic ("new") and thematic ("old") information in a sentence led to the development of suprasentential units. This, together with the realisation that there are segments that are dependent on previous and subsequent discourse, meant the broadening of sentence context and the breaking of the sentence barrier (cf. Langleben, 1981). This line of analysis, focusing on intersentential structures, led to the development of modern studies of cohesion and topic-comment functions (e.g., Halliday & Hasan, 1976).

One of the first large-scale inquiries into text structure and organisation in general was that of Harweg (1968). He postulated that texts are held together by mechanisms of "substitution" — one expression following another one of the same sense or reference, forming a cohesive or coherent relationship. Harweg defines text as a "succession of linguistic units ... that is built up by an uninterrupted chain of syntagmatic substitutions" (Harweg, 1977, p. 148). The process of "syntagmatic substitution" is one whereby "expressions are substituted in the sequence for previous expressions that denote the same or related content" (Harweg, 1980, p. 313). He describes concatenation through syntagmatic substitution as a left-to-right process, determining the textuality of a sequence of sentences, complemented by a top-down process, determining the particular text type (that which he called the "macrostructure").

The development of *generative text grammars* characterises the approaches of Janos Petöfi and Teun Van Dijk (Petöfi, 1971; Van Dijk, 1972; 1977c), which were based on the view that the goal of linguistic theory is to provide a mechanism for the generation of a well-formed text (as opposed to the generation of well-formed sentences). Petöfi's theory of texts is often called the "text structure / world structure theory". His idea was to integrate factors relating to the *users* of texts rather than to the text as an isolated artifact. Petöfi's theory was based on the postulate that there are correspondences between the structure of a text and the structure of the "world" that the text evokes. Van Dijk's theory involved the notion of "macrostructure" (though not in the same sense as Harweg's term). He argued that, since the generation of a text begins with a main idea that evolves into the detailed meanings comprising individual sentences, when a text is read there must be operations that work in the opposite direction to extract the main idea back out again (via "macrorules", such as deletion and generalisation).

Although European textlinguistics had been underway for at least a decade, it wasn't until the 1970's that the study of texts emerged as a major consideration in American research. Zellig Harris (Harris, 1952; 1970) was one of the first American textlinguists. He tried to apply methodology from American structural linguistics concerned with sentential analysis to the analysis of connected prose. Another major influence in America came from researchers in tagmemic theory: Kenneth Pike and Robert Longacre (cf. Pike, 1967; Longacre, 1972). Tagmemics is a method that involves analysis in terms of slots and fillers — that is, the positions "open" in a piece of text and the units occupying those positions. This method of syntactic analysis was applied to the discourse at the level of sentence, paragraph and text producing a hierarchy in terms of form and function. The method also involved an abstract "deep structure" of the discourse, and rules relating this structure to surface syntax in terms of intersentential relations and grammatical devices.

Influences from other disciplines also led to the development of current theories of text structure. In poetics and anthropology it was shown that the analysis of different discourse types should be given in terms of units, categories, and rules based on the semantic "macrostructures", or global meanings, of the discourse (Van Dijk, 1972). In sociology, sociolinguistics, and pragmatics, it was discovered that everyday conversations are bound by fairly strict conventional rules (e.g., Cicourel, 1975) and that structures in discourse run parallel with structures in speech act sequences (Van Dijk, 1977a).

An important influence on psychological interest in texts came from the field of artificial intelligence, when, in the 1970's, language understanding became a major research topic. The early language understanding programs focussed initially on words and sentences, but soon changed to include whole texts and discourse (e.g., Schank & Abelson, 1977). This kind of research showed that structure was not enough for theories of text understanding — they also needed process and world knowledge. However, it was soon realised that inferences generated and constrained only by low-level knowledge quickly get out of hand, so higher-level structures were needed to channel and constrain processing.

The Study of Narratives

BACKGROUND FROM LITERARY THEORY AND ANTHROPOLOGY

Artificial intelligence research on texts has tended to focus mainly on story understanding, which brings us to the second major influence on the development of modern theories of text structure. The study of narrative structure comes from two backgrounds: literary theory (including poetics) and anthropology.

Traditional narrative studies were actually theories of "the novel", although there were also studies of more specific narrative genres, including short stories and fables. Traditional narrative theory may be seen as a contemporary version of Aristotle's ideas about the epic, as set out in the *Poetics*. These ideas influenced much of the Anglo-Saxon and post-war German literary work on narratives. Another major influence came from the Russian Formalists. These theorists focused on the formal analysis of discourse and introduced notions, for example, about the thematic structure of stories, and the difference between the canonical structure of the story and its transformations in terms of the plot (e.g., Todorov, 1965).

However, it was mainly Vladimir Propp's (1968) work on *The Morphology of the Folktale* which became the background for the later work on structural analysis of the narrative, developed in France at first by Claude Levi-Strauss, and later taken up in America by, for example, Colby (1973). Propp classified folktales by identifying functions defined as acts performed by certain characters (e.g., the villain, the hero, etc.). He regarded all fairytales as having one structural type, in which the functions of characters were stable and constant elements, independent of how and by whom they were fulfilled, and that the sequence of functions performed is always identical.

Both structural and generative textlinguistics also became influential in studies of narratives. Structural semantics was applied to stories by Greimas (1966) in his account of Propp's narrative structures. He used a case grammar semantics, analysing events and actions into predicates and actor types, at the level of the sentence and the text as a whole. However, there was no real link between linguistics and narrative studies until the text grammars developed by Van Dijk and others (e.g., Petöfi & Reiser, 1973; Van Dijk, 1972, 1977c; Dressler, 1977; Van Dijk & Petöfi, 1977). Van Dijk's work involved a mapping of the narrative structures in text, which applied a number of semantic constraints. His system was based on findings from sociolinguistic studies, such as those of Labov and Waletsky (1967), concerning "natural narratives" (i.e., stories of personal experience), as well as on structural analyses. Labov and Waletsky (1967) found that everyday stories often consist of canonical structures such as setting, complication, resolution, and evaluation. In Van Dijk's system a superstructural schema was mapped onto the macrostructural level (i.e., the level of global meanings) of the story as a whole, by assigning narrative categorical functions, and thereby organisation, to the sequence of macropropositions derived from the propositional "textbase" of the discourse, with the aid of macrorules. Thus, the narrative categories were postulated to have constraints on the global semantic content.

CONTEMPORARY THEORIES IN PSYCHOLOGY AND ARTIFICIAL INTELLIGENCE

One of the earliest psychological studies of narrative structure was provided by Bartlett (1932). The importance of his use of the notion of a "story schema" was to indicate that it could account for some of the reconstructions he observed in recall. He suggested that when people cannot remember parts of a story, they can use this schema to reconstruct what might have occurred. This general notion also accounts for the increasing regularity of irregular stories over time, that is, the finding that recall tends to approximate the prototypical schema more than the actual input.

The kind of analysis performed by Propp led some researchers to see strong parallels between linguistic structure and story structure, and to propose "story grammars" that draw on the conceptual apparatus of syntactic theory. Lakoff (1972) proposed what has been regarded as the forerunner to modern story grammars. He reformulated Propp's (1968) theory of Russian folktales using rewrite rules. Colby (1973) also used rewrite rules in designing a specialised grammar of Eskimo folktales. Prince (1973) borrowed explicitly from Chomskyan syntactic theory, drawing parallels between readers' tacit knowledge of syntax and their knowledge of what makes a story. He also postulated event structures which he called episodes —

a major feature in contemporary story grammars. In his scheme, complex stories were derived by applying generalised transformations that could embed one story within another, just as the generalised embedding transformations of early transformational grammars embedded one clause within another. The first general grammar designed to apply to more than a small and restricted set of stories was proposed by Rumelhart (1975). Other grammars have been proposed by Van Dijk (1975), Thorndyke (1977), Mandler and Johnson (1977), and Stein and Glenn (1979). These more recent analyses will be dealt with in some detail in Chapter 4.

CHAPTER 3

TYPES OF TEXT STRUCTURE

As the preceding brief historical overview shows, text theorists gradually came to realise that a coherent text has some inherent structure over and above the structure of the sentences comprising it — that is, a text is more than a string of sentences. It is constructed according to specific purposes of the author and, in order for it to act as an effective communication medium, it must be structured in accord with the author's purpose, whether it is to state an argument, to persuade the reader, or simply to entertain. Of course, even if the text is well structured and coherent, it will not succeed in being an effective communicative medium unless the author has taken into account the structures in the reader's head — that is, the reader's prior knowledge, expectations, and intentions in reading the text. I shall focus on the reader's knowledge structures and the more constructive aspects of text comprehension in Chapter 4. In this chapter I examine the structures inherent in the text itself that serve to make it a coherent piece of discourse.

De Beaugrande (1980a) proposes that the text be regarded as a cybernetic system, or a set of elements functioning together. Iser (1980) has also noted that the relationship between text and reader can be regarded as a self-regulating system. The stability of the text as a system is maintained, for example, by the system of syntax which imposes organisational patterns upon the surface text. These organisational patterns or structures provide the text with coherence between its individual elements. They are also used by readers in processing, and function in organising the information in the text in a way that makes processing easier, given the limited capacity of human memory.

The proposal that higher order structure is required by the nature of the human information processing system is substantiated with respect to sentence comprehension by, for example, Fodor, Bever and Garrett (1974). They cite evidence for a clausal processing strategy, whereby words are held in working memory in relatively unchanged form until they can be grouped into a clause. At this point the representation becomes more abstract, and individual words are less likely to be retrieved. There is also evidence that the ability to monitor extrasentential information is reduced prior to clausal boundaries, suggesting that an increased load is being placed on the system, due to a recoding process.

The syntactic devices that serve to maintain the stability of the text occur not only at the lower level of clauses and sentences. In longer stretches of text there are cohesive devices such as recurrence, parallelism, pro-forms, ellipses, and so on. However, a text is comprehensible not only due to a continuity of *referents* (cohesion), but also due to a continuity of *senses* among the knowledge activated by the expressions of the text — this requires devices for coherence, such as schemas, frames, scripts, etc.

In constructing and comprehending a text it is important to maintain continuity, and although texts are linear on the surface, their underlying structures may be hierarchical. Figure 3.1 illustrates how a surface linear form can give rise to a

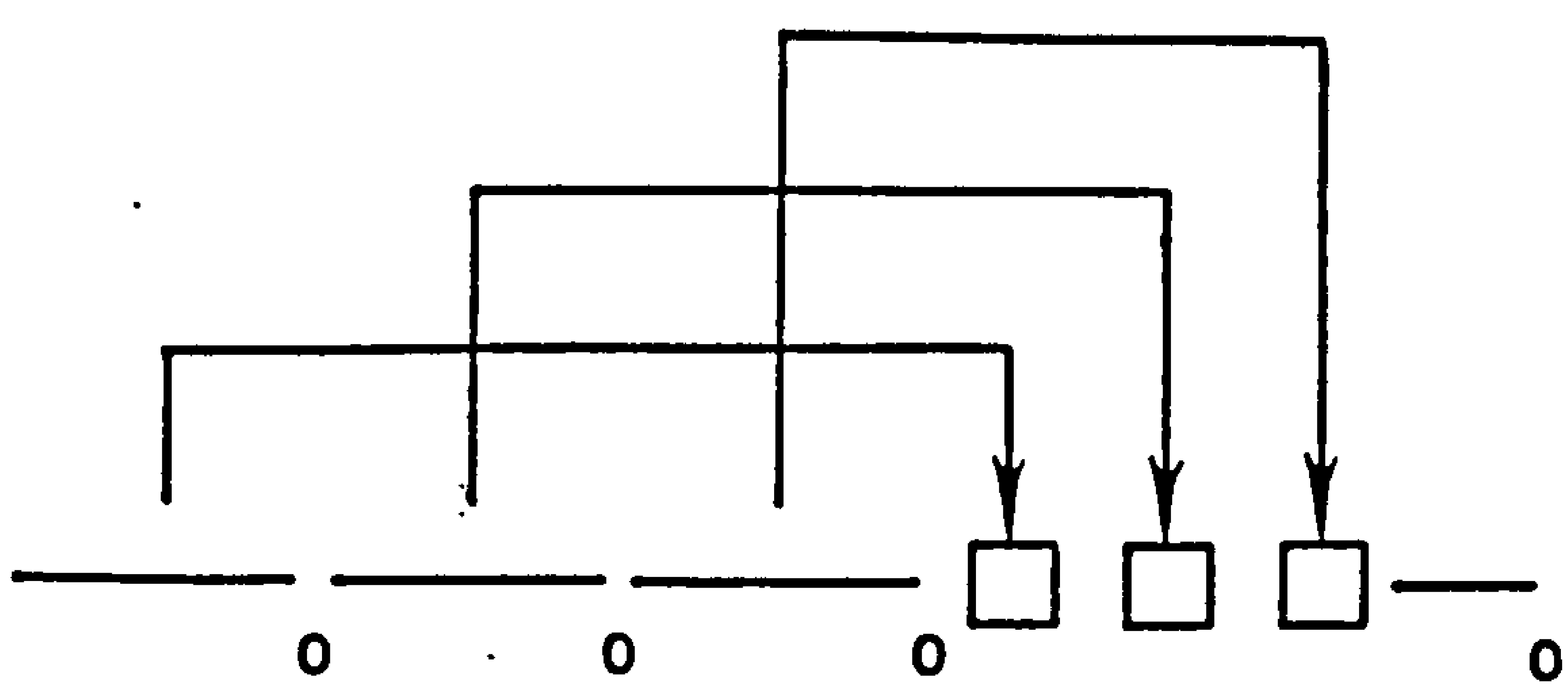

Figure 3.1. An example of text structure, showing how a surface linear organisation can give rise to a hierarchy. — stands for sentences, O for punctuation signs, and □ for resumptive linguistic entities (cf. Ballmer, 1981).

hierarchy, through such mechanisms as the "introduction-transport-use" principle.

It shows that a text has the shape of a linear sequence of sentences, including a series of sentences that serve to introduce and fix certain contexts, as well as devices to summarise and resume the introduced contexts. Texts may show many such units in a nested way so as to give rise to a hierarchy (cf. Ballmer, 1981).

Several researchers have criticised so-called "structuralist" theories of text as being too descriptive (e.g., Reiser & Black, 1982; Johnson-Laird, 1983), arguing that the emphasis should be placed on cognitive processes and functional accounts that are independent of the type of text. This argument will be dealt with in further detail in Chapter 4. For the moment it is worth noting a comment from De Beaugrande and Dressler (1981), who advise against "allowing the text to vanish away behind mental processes": "If that notion were accurate, textual communication would be quite unreliable, perhaps even solipsistic. There must be definitive, although not absolute, controls on the variations among modes of utilising a text by different receivers" (De Beaugrande & Dressler, 1981).

The Function of Form: Schematic Structures

Recent psychological research concerning structural factors in text comprehension has tended to suffer from confusions over the meaning and use of the notion of structure. Two particular controversies stand out: the criticism of story grammar theories for confounding form with content (the subject of the discussion in Chapter 4), and the variety of findings pertaining to the so-called "levels effect" (i.e., the notion that the "higher" in the structure, or the more "important" the information, the better it will be recalled — Chapter 3 contains a discussion of this topic).

There are several ways of distinguishing "levels" of structure in text (cf. Van Dijk, 1980a; Mandler, 1983, 1984). One can distinguish the whole from its parts (e.g., sets and members of sets); there may be several points of view or perspectives involved; one can distinguish between the general and the abstract (i.e., the gist or "upshot" of the discourse) and the particular details or instances; yet another notion involved is that of importance, relevance, or centrality to some topic or theme. What unites all these notions (i.e., of theme, topic, or gist), at least with respect to text structure, is that they all pertain to meanings — that is, they are *semantic* global structures. Van Dijk (1980a) distinguishes between these semantic or "macro-" structures and what he calls "superstructures". The latter are *schematic* global structures that give order or organisation to the semantic global structures. Schematic structures are categorical in nature, and pertain to the form of the text, rather than the content. The categories of superstructures are functional ones — i.e., superstructures organise macrostructures by assigning functions to them. They can be regarded as the "macrosyntax" for the global meaning of the text (Van Dijk & Kintsch, 1983). In other words, the distinction between macrostructure and superstructure is similar to that between the meaning of sentences and their functional syntax, in that superstructural forms put certain constraints on the content of the text.

These functional categories may become more or less conventionalised in a given culture and lead to the establishment of fixed schemas for the global content of the text (viz., the canonical structure of stories, newspaper articles, scientific reports, etc.). These functional categories also contain rules that specify what category can follow from, or be combined with which others. Thus: "the macrorules essentially define what is important, relevant, or more abstract information in a given text and ... this importance may change for the various communicative functions of the discourse, which again is related to discourse type" (Van Dijk, 1980a). In other words, the notion of "level" or importance in the text structure becomes defined relative to the functional categories organising the global meaning of the text, and these functional categories are determined by pragmatic and social factors as well as semantic ones.

There are two kinds of approach to analysing text structure: that which sees structure as a superimposed functional form, organising the content, and that which sees structure as emergent from the content relations in the text. For Kintsch and Van Dijk (1978; Van Dijk & Kintsch, 1983), the schematic categories of superstructures are functional ones, superimposed upon and organising the text content. Other researchers focus more on "content structures". For example, in Meyer's system (Meyer, 1975, 1985) the top-level structure is based on the content relations that can subsume the greatest amount of text. So, whereas for Kintsch and Van Dijk the top-level structure is an independent organisation overlapping the propositional analysis, in Meyer's system it is an emergent organisation or structure.

Van Dijk and Kintsch (1983) stress the functional role of schematic structures in arguing that, although knowledge of the world is necessary for understanding and making predictions about the course of a text, many discourse types are not predictable in this way — particularly when they have to introduce new information (e.g., scientific reports, instruction manuals, and so on). Thus, understanding the content of the text not only requires knowledge of the world, but also knowledge of the discourse type. Moreover, for specialised types of discourse, comprehension strategies may have to be taught explicitly to readers.

TEXT TYPES

Traditionally, rhetoricians have classified texts according to two properties: the purpose of the text, and what the text is about (cf. Faigley & Meyer, 1983). This is similar to Van Dijk's distinction between superstructure and macrostructure (Van Dijk, 1980a). In the *Rhetoric*, Aristotle (trans. 1960) classifies two kinds of argument: one relying on external evidence, the other on persuasion. Morris (1946) distinguished between discourse "uses" (purpose or function), and the ways ("modes of signifying") in which these uses are realised in discourse. Several other researchers have emphasised the *functional* nature of text structure (e.g., Halliday, 1982). For example, De Beaugrande and Dressler (1981) note that narrative texts are used to arrange actions and events in a particular sequential order, so there will be a certain frequency of relations such as cause, reason, purpose, enablement, and time proximity, whereas argumentative texts are used to promote the acceptance or evaluation of certain beliefs or hypotheses as true or false, so relations such as reason, significance, value, and so on, should be frequent in the text. Brewer (1980) classifies text types along the dimensions of "cognitive structure" and "force". He identifies three types of structure — descriptive, narrative, and expository — and four types of force — informative, entertaining, persuasive, and literary-aesthetic.

In keeping with this functional approach to text structure, readers are hypothesised to use the structure in the text to guide their comprehension. When readers begin a text they try to establish as soon as possible what is the initial theme or macrostructure (Van Dijk & Kintsch, 1983). They also need to establish the global function of the part they are reading, and its theme within the discourse as a whole. That is, a hypothesis is made about the schematic or superstructural category involved. Strategies may involve bottom-up processes — i.e., guessing the function of a particular proposition — as well as top-down processes — i.e., knowledge of the discourse type, and therefore the schema for the text. Thus, the role of the superstructure is to facilitate comprehension, storage, and retrieval of the text. In accord with basic memory organisation principles involved in human information processing, texts are processed in chunks, units, or categories (cf. Miller, 1956). More structure is often correlated with additional or more complex processing, which results in better structural representations, and therefore in better retrieval.

Assumptions about the plausible schema derived from contextual information or from previous experience with the text are combined with more local semantic information in order to form and confirm hypotheses about the actual schematic category:

> Since the general principles, rules or strategies, categories, and units, for each language and culture are learned by the language users, these organizational patterns become preprogrammed, so to speak, which also facilitates understanding ... Thus, the schematic categories of a story may function as the conventionalized discourse functions for semantic macrostructures, namely, as a possible form for the global content of the story. (Van Dijk & Kintsch, 1983, p. 237.)

Superstructures must be not only in the text, but also "in the head" — that is, the reader has to know about the conventional schema before he/she can use it. Studies have shown that the type of schema used to view a topic affects the type of information recalled (Kozminsky, 1977; Pichert & Anderson, 1977; Anderson & Pichert, 1978; Anderson, Spiro & Anderson, 1978; Meyer & Rice, 1982). Research has also demonstrated the operation of superstructures in several types of text: stories (Mandler & Johnson, 1977; Rumelhart, 1977), newspaper articles (Van Dijk & Kintsch, 1983), and expository prose (Meyer, 1977; Meyer & Freedle, 1984). Other studies have shown that readers can be taught to identify and use the overall structure of different discourse types with resultant changes in organisational strategies and recall (Meyer & Rice, 1982; Meyer, 1984).

Meyer (1984) performed a study that looked at comparative memory facilitation of different types of expository text. She postulated five basic ways of organising discourse (largely based on Grimes' (1975) analyses): collection, description, causation, problem/solution, and comparison structures (Meyer, 1975, 1985). Briefly, descriptive texts are organised according to grouping by association of topic with attribute. Collection is the same as description, but also involves sequencing in time. Causation involves both these kinds of organisation, and, in addition, causal or quasi-causal relations. In the problem/solution text, at least one aspect of the solution matches the problem in content and blocks an antecedent of the problem. Comparison texts are different from the others, in that they are organised on the basis of similarity or difference relations.

Meyer's hypothesis was that, since causation, problem/solution, and comparison schemas have more organisational components than collection/description schemas, they should facilitate encoding, storage, and retrieval. She found that performance was better for comparison, causation, and problem/solution types of organisation, for both immediate and delayed recall. Subjects who read the comparison, causation, and collection/description texts tended to use the same corresponding organisation in their recall protocols. However, those who read the problem/solution texts used a variety of the different possible organisational strategies. (This was attributed to the fact that these subjects disagreed with the author's message in the way it was presented in the problem/solution texts. Thus, there are certain cases where the author's organisational scheme will not be adhered to.)

Olson, Mack and Duffy (1981) compared the processing of stories and essays, and concluded from their findings that readers of stories have essentially a "prospective orientation", generating predictions and looking ahead to what is coming up. In contrast, readers of essays tend to process the text more retrospectively, fitting the current sentence in with earlier information that had been explicitly presented in the text.

Although several studies have found differential effects of prose type on comprehension and recall, others have only found weak effects (e.g., Harris, 1981). There are also different effects depending on the type of task. The results of a study by Graesser, Higginbotham, Robertson and Smith (1978) suggest that it is not the type of text alone that determines what sort of processing occurs. They found that when subjects process text without any task demands, they tend to show a "narrative bias", whereas when task demands are introduced, this bias disappears.

Kintsch and Young (1984) found that structure may actually hinder, depending on the task. Previous studies (e.g., Anderson & Pichert, 1978) showed that instructions to adopt a particular perspective can resulting the failure to retrieve material that is irrelevant to that perspective. Anderson, Spiro and Anderson (1978) found that stories arranged to emphasise different scripts led to different types of recall. Kintsch and Young investigated the effect of text type (narrative, descriptive, and expository text) on the likelihood of remembering decision-relevant materials when the reading goals are decision-irrelevant. They argued that if a text is recalled well because an efficient structure for it has been formed, then the parts of the text that are relevant to the structure should be recalled well, and the parts of the text that are irrelevant should be recalled less well. On the other hand, in texts that do not facilitate the construction of a well-defined structure, information not relevant to the structure might be recalled better, even if recall is low overall, since such information is not dismissed by the reader as irrelevant. Kintsch and Young reasoned further that since narratives are in general recalled better than expository texts, due to the operation of a story schema, and since incidental decision-relevant material would not be relevant to the structure, it would be recalled less well in stories than in expository texts. They found that overall recall was better in the narrative condition than for expository texts, but the decision-relevant information in target sentences in the expository condition was recalled better than in the narrative condition.

The Structure of Content: Semantic Structures

The type of structure discussed in the previous section concerned the form (conventional, pragmatic, and logical) that serves to organise the global content of the text. This section concerns a different type of structure — different not simply with respect to "level" (i.e., along the local/global dimension): The kinds of structure to be examined are those emergent from the content of the text — those structures that serve to make the text "coherent" or "meaningful" in general, rather than with respect to the purpose or communicative function of the text.

LOCAL AND GLOBAL COHERENCE

This thesis concerns chiefly the role of schematic structures in text comprehension. However, this is only one aspect of what makes a text coherent, and thereby understandable, to a reader. It is almost as difficult to say what the term "coherence" means, however, as it is to say what "meaning" is. The notion of coherence has at least an intuitive meaning, and it appears that readers themselves can agree to a large extent on what makes a text coherent. Kucer (1983), for example, found a high degree of agreement between readers concerning which propositions in the text contributed to their own summaries when the text was highly coherent (but there was little agreement when the text was low in coherence). The term has also been adopted in a technical sense by researchers in linguistics and text comprehension. However, for all that, it has become little more precise.

One confusion that has arisen concerns the fact that the terms "coherent" and "cohesive" are synonymous in standard usage, but not always in technical usage. Some researchers have tended to regard the syntactic devices contributing to cohesion (e.g., anaphora, ellipsis, etc.) as being responsible also for the text's coherence. As Charolles (1983) notes, the problem concerning the notion of coherence boils down to what in a text shows intrinsically that it is a connected whole. In other words, to what extent is the text itself (or rather, the author) responsible for its coherence, and how much is induced by the reader? However, this still does not clarify the distinction.

One might, for example, take the distinction to be similar to that between syntax and semantics (although the latter distinction also becomes blurred — thus, one hears about the "syntax of semantics" and the "semantics of syntax"). Halliday and Hasan (1976), amongst others, have made a major study of the syntactic devices occurring both within and between sentences that serve to make the sentences connected or cohesive. They state explicitly the distinction between coherence and cohesion: "A text is a passage of discourse which is coherent in these two regards: it is *coherent* with respect to the context of situation and therefore consistent in register, and it is coherent with respect to itself and therefore *cohesive*" (Halliday & Hasan, 1976, p. 23, emphasis original).

Thus, it appears that, for these researchers, the distinction is similar to that between syntax and semantics. In fact, elsewhere Halliday and Hasan state even more categorically that cohesion is not a matter of content: "Cohesion does not concern what a text means; it concerns how the text is constructed as a semantic edifice" (Halliday & Hasan, 1976, p. 26).

At other times one gets the impression that the controversy is about the distinction between the explicit cues in the text and those implied by the text, which the reader has to infer (cf. for example, Langleben, 1981). Tierney and Mosenthal (1981) investigated whether or not one could actually measure the coherence of a text (as rated by subjects) via a statistical analysis of its cohesive connections (i.e., reference, conjunction and lexical ties). They found no relationship between cohesive patterns and coherence ratings within a given text topic, and concluded that a cohesion index is causally unrelated to the text's coherence. (See also Morgan and Sellner (1980), who criticise the implied causality of linguistic forms such as cohesive devices on textual coherence.)

However, the distinction between explicit connectives and implied connections is not the whole story. For example, Halliday and Hasan (1976) use the term "cohesive" to refer to devices for endophoric (i.e., intra-textual) reference. According to them, exophoric reference links language with context, but they state that it does not contribute to the integration of units within the same text. In other words, the distinction appears now to be one contrasting connections between referents in the text with connections between what's in the text and what is not. However, although Halliday and Hasan argue that cohesion does not involve meaning, in anaphoric reference resolution the referents of the respective expressions have to be identified and compared, and therefore understood with respect to the reader's knowledge and the context, in order to be resolved. Furthermore, there is a good deal of psychological evidence that such processes as anaphoric and exophoric reference resolution are highly interdependent and interactive (e.g., Greenspan & Segal, 1984).

Charolles (1983) argues that the distinction between coherence intrinsic to the text, and that dependent on the reader or the context, is relative. He points out, for example, that an account that states that a particular series of two sentences is cohesive because the first one is taken up by a coreferring pronoun in the second, leaves open the question of whether the sequence of two sentences is cohesive in its own right, or coherent by virtue of certain operations performed by the reader.

Linguistic theories of coherence and cohesion have been criticised not only for neglecting contextual and semantic factors, but also for ignoring the pragmatic aspect. This aspect is emphasised by Hobbs (1979), who argues that coherence is deeper than the notion of a discourse just being "about" some set of entities (whether intra- or extra-textual). He views coherence from the point of view of the author, where the relations correspond to coherent "communication moves" — that is, the means of continuing a discourse in a relevant way. The goal of the author is to ease the processing load on the reader, by structuring the text in a way that will enable making the right inferences quickly. He suggests (Hobbs, 1982) viewing an utterance as being coherent insofar as it can be seen as an action in the implementation of some plan. (This notion is intended to apply to all discourse, not just dialogue, as the term "utterance" implies.)

Part of the problem concerning what makes a text coherent may be the levels chosen for analysis. Some researchers focus on the syntactic aspects, others on the semantic or psychological, and still others on the pragmatic, social, and contextual aspects of coherence. Confusions concerning what is coherence may have arisen because the devices for coherence operating at these different levels do not depend in any straightforward way on the "subordinate" levels, or that they are quite different in nature. Langleben (1981), for example, argues that the conditions for coherence will not be the same at every level: "From both the stratificational and the functional points of view, ascension via a hierarchy means that various aspects of the coherency come into action, gradually accumulating toward the pinnacle of the whole text set in the pragmatic framework of reality" (Langleben, 1981, p. 287). Halliday (1982) also recognises the existence of levels of coherence, and identifies several "higher-level" devices, for example, implicit conjunctions due to events following one another; conjunctions between larger scale entities such as "because of this...", referring back to earlier parts of the text; relations between functional categories such as premises and conclusions; thematic movement between "given-new", "background-foreground", and so on.

Giora (1983) also argues that devices such as the topic-comment function apply beyond the level of the sentence, and typify connectedness at various levels of discourse. An elementary organisation of a text is a linear progression from the rhematic ("new") part of the sentence to the thematic ("old") part of the following sentence. This view of intersentential thematic relations follows the principle of proceeding from old/given to new/foreground information, which requires that a sentence should start with given information (theme) and end with new information (rheme). In the subsequent sentence the theme is the rheme of the previous sentence.

At a higher or textual level there are similar kinds of segmentation and progression. However, these serve not only a cohesive function, but determine the informational structure of the text. Information that is introduced in a final position gains in foregrounding; introduced in the following segment initial position, it becomes informationally downgraded. Sequences of discourse topics produce such text segmentation in a way that creates "informational hierarchies" (Giora, 1983). Van Dijk and Kintsch (1983) also distinguish between local and global coherence, identifying several levels and types of coherence, including syntactic, semantic, stylistic, and pragmatic levels. They also greatly broaden the notion of coherence, adding such types as establishment of the same perspective, and level of description or degree of specificity.

In summary, as the preceding discussion suggests, the notion of coherence is problematic. For some theorists (usually linguists) it has a very specific definition; for others (usually psychologists) it is as all-embracing a concept as "comprehension", "understanding", and "meaning".

TOPIC, THEME, AND GIST

Concepts such as coherence and cohesion refer to the ways in which a text "hangs together" or is connected into a whole. However, without such related concepts as topic, theme, relevance, and so on, the structure of the text is "flat". With the introduction of these additional concepts the structure takes on a more elaborated form, usually conceived of as a hierarchy. Furthermore, a well-formed text is not simply coherent by virtue of its sequential connectivity, but also by virtue of its "aboutness". We have an intuitive notion of a text being "about" something. It is a common finding, ever since Bartlett's studies of memory for stories (Bartlett, 1932), that readers remember the "gist" of texts and can reproduce fairly accurate summaries, even if they cannot remember the details of what they have read. Furthermore, the summaries that readers produce tend to agree with each other to a large extent (Kintsch & Kozminsky, 1977).

As with the notion of coherence, it is hard to find definitions of concepts such as topic, theme, focus, relevance, importance, perspective, and so on, that are any more precise than our intuitive notions. Some researchers choose one term apparently to cover all the others; other researchers use some or all of them interchangeably.

The notion of sentence topic has the clearest definition. At this level, the term "topic" is used to refer to those parts of a sentence that are "presupposed", "old", or "given" (cf. Haviland & Clark, 1974); the term "comment" is used to refer to those parts of a sentence that are "focal" or "new". Presupposed and focal markings or signals generally occur at predictable sentence positions: Information early in the sentence is usually marked as presupposed (e.g., the grammatical subject), whereas focal (new) information normally appears near the end of the sentence (e.g., the direct object in a transitive sentence).

There is a good deal of psychological evidence for the topic-comment distinction at the sentential level. It has been found that, in general, sentences whose referents are given or presupposed are processed faster than when they are new (e.g., Haviland & Clark, 1974). According to the "integration theory" of Haviland and Clark, sentences are integrated in a three-stage process. First, the current sentence is broken down into its respective components, given or presupposed, and new or focal. Then memory is searched for an antecedent that matches the current given information. Finally, the new information is integrated into memory by adding it to the representation that contained the antecedent. If no antecedent is found, some form of bridging inference is necessary for integrating the new information.

Yekovich, Walker and Blackman (1979) found that linguistic characteristics of both context (presupposed) and target (focal) information affect the integration of sentences. When common information was marked appropriately in each of the two sentences, comprehension was facilitated, whereas inappropriate markings in one or both sentences led to slower comprehension. Chang (1980) gave subjects probe words after each word in two-clause sentences and found that decision times for words in the final clause were faster than for words in the previous clause, and also that pronouns in the last clause "activated" the meaning of the antecedent in the first clause — suggesting that pronouns are interpreted online and not after input of the whole clause in which they occur.

Van Dijk (1979; Van Dijk & Kintsch, 1983) makes a distinction between this kind of sentential topic (i.e., what is presupposed, known, or given) and the discourse topic, which includes both a central referent and the major predications involving this referent. In other words, in the view of Van Dijk and Kintsch, the notion of topic is only appropriately defined in terms of the relations between sentences and the rest of the text and/or the context, which is a similar assertion to that made concerning the notion of coherence, as discussed in the previous section. Van Dijk (1979) argues that at the level of sequences of sentences, relevance to a topic is due to referential coherence — that is, it is usual in a sequence that expressions denote identical referents. However, these referents achieve *discourse* topicality, importance, or relevance, by the fact that respective predications are "applied" to them: "Since we may conclude that, apparently, *both* the (identical or central) referent and the major predications of this referent may be relevant in the sequence as a whole, we must assume that in fact it must be a full *proposition* which should be taken as the 'topic' of a sequence" (Van Dijk, 1979, p. 117, emphasis original). In other words, the "topic" of a sentence is a discourse function that serves to produce partial coherence with the contextual representation of the previous part of the text (or context).

The effect of context on sentential topic has been demonstrated in several studies. For example, Caramazza and others (Caramazza, Grober, Garvey & Yates, 1977) have demonstrated that pronoun understanding involves a top-down strategy. They showed that the meaning of verbs in antecedent clauses influences the interpretation of pronouns in following subordinate clauses introduced with the connective "because", which they explain in terms of the implicit causality of the verbs. (However, Ehrlich (1980) has pointed out that this cannot be due to the meaning of the verb, since it doesn't happen with connectives such as "but".)

Van Dijk and Kintsch (1983) report an experiment in which subjects were asked to complete a sentence beginning with a pronoun, in the context of one or more preceding sentences. They found that most subjects chose a continuation in which the pronoun coreferred with the topical noun in the previous sentence, and that for sequences of sentences, subjects also chose as an antecedent the topical noun in the first sentence, even if several sentences intervened.

The status of the topic of the discourse is similar to that of the topic of the sentence, in that both capture the notion of "aboutness" (i.e., presupposed or given). The difference is in scope. At least in Van Dijk and Kintsch's view, the topic of a sequence of sentences or a whole discourse, includes both the central referent and what is predicated about that referent, whereas at the sentence level that which is predicated about the topic is distinguished as comment or focus. The notion of discourse topic thus captures the notion of the "main idea" of a sequence of sentences or a whole discourse.

Kieras has made extensive studies of the role of topic information in text processing. He hypothesised that comprehending an incoherent passage requires the maintenance of "topic pointers" in memory, one pointer for each unintegrated portion of input (Kieras, 1979). He also hypothesised that these pointers had to be maintained by rehearsal in short-term memory. However, in a simulation model of reading times, the hypothesised processing load failed to account for reading times that showed memory loss in a recall task, although it did account for it in other tasks (e.g., topic choice). Kieras found that in the recall task subjects appeared to deal with the incoherent parts of the passage in a way that differed from other tasks. He suggested that the hypothesised topic pointers do not undergo rehearsal unless the material is being encoded for future recall. Another suggestion was that rehearsal of unintegrated material and maintenance of topics are separate processes.

Van Dijk (1980a) uses the term "macrostructure" to account for the various notions of global meaning, such as topic, theme, or gist, as well as the notion of global coherence. Macrostructures consist of macropropositions, derived from the micropropositions expressed in the sentences of the text. Kintsch and Van Dijk

(1978; Van Dijk & Kintsch, 1983) postulate semantic mapping rules (macrorules) that relate proposition sequences with sequences at a higher level, and produce or derive the global meaning from the local sentential meaning of the text. Macrorules operate recursively, thus producing a hierarchical structure.

The Kintsch and Van Dijk Model of Text Processing

The model proposed by Kintsch and Van Dijk (1978; Van Dijk & Kintsch, 1983) assumes that the "meaning" of a text can be specified at two levels: at the level of the microstructure, where an ordered list of micropropositions represents the individual ideas in the text, and at the level of macrostructure, where the main ideas are represented by an ordered list of macropropositions. A text is said to be coherent at the microlevel to the extent that there is argument overlap, or referential coherence, among propositions. An important claim of the model is that the processes taking place at the level of propositions, or microstructure, are similar to those taking place at a more global macrostructural level.

The reader is hypothesised to establish a memory representation of the text in cycles — at the microlevel, a cycle is assumed to be a phrase or sentence in length. Studies have shown that sentences and major clauses serve as units of comprehension (Jarvella, 1971; Aaronson & Scarborough, 1977). Kintsch and Van Dijk (1978), and Miller and Kintsch (1980) have used these findings in constructing the cycles in their model of text processing. A major component of the processing cycle is the assumption of a short-term buffer, that assists in carrying over propositions to cohere with the input of the next cycle. According to the model, the reader attempts to relate (micro)propositions by argument overlap to propositions that are already in short-term memory, retained from the previous cycles. The resulting structure is represented as a memory network, in which each node is a single proposition and each link between nodes signifies argument overlap (referential coherence) between two propositions. Argument overlap is also used to determine a proposition's importance, or "level", in the network. Thematic propositions are assigned the highest level; propositions that share an argument with a thematic proposition are assigned the next highest level, and so on. It is assumed, therefore, that the reader attempts to represent a text by a single network, hierarchically organised.

If a single network cannot be formed, then a resource-consuming search of long-term memory is made. If a proposition in long-term memory is found that shares an argument with the input set, this "overlap" proposition is reinstated in the short-term memory set. Processing operates by the transformation of the input into a micropropositional textbase, and the application of "macrorules" to transform this textbase into a macrostructure.

In the first stage of processing, micropropositions are linked by referential coherence, and a hierarchical processing cycle is set up that results not only in multiple processing of some elements, due to referential coherence, but also in differential retention, since the number of times a particular element is processed determines its strength or activation value. Referential coherence in this model (i.e., the version described in Kintsch & Van Dijk, 1978) is taken as simple argument overlap: that is, two propositions are said to be referentially coherent if they share the same argument, or if one proposition is embedded as an argument of another. If a textbase is referentially coherent, that is, if there is some argument overlap among all of its propositions, it is accepted for further processing. However, if "gaps" are found, inference processes are initiated to close them, i.e., one or more propositions will be added to the textbase to make it coherent. The text is processed sequentially in chunks of several propositions at a time. When a chunk of propositions is processed, some of them are selected and stored in the buffer (the number selected being constrained by the capacity limit). If there is an argument overlap between the input set and the contents of the short-term memory (STM) buffer, the input is accepted as coherent. If there is no overlap, a resource-consuming search is made of previously processed propositions, now held in long-term memory (LTM). If there is still no overlap after this, an inference process is initiated. A network of coherent propositions is then constructed, the nodes of which are propositions, and the links of which are shared referents. The network is arranged hierarchically by selecting as the topmost proposition the topic, or if there is no topic, the one that results in the "simplest" structure, i.e., the one that is a presupposition of its subordinate propositions on the next levels.

In the 1978 model, the selection criteria for inclusion in the STM buffer, which determines the amount of processing a proposition receives, are arbitrarily based on principles of recency and frequency. If all the propositions of the input cycle are successfully added to the network, then a subset of these propositions is selected to remain in the buffer for processing on the next cycle. If some propositions cannot be added to the buffer because of the lack of argument overlap, then either LTM is searched for an interconnecting proposition, or a bridging inference is made. The STM buffer has a capacity limit of *s* propositions, selected on the basis of recency and referential overlap.

In the second stage of processing, macro-operators (cf. Van Dijk, 1980a) transform the propositions of the textbase into a set of macropropositions that represent the gist of the text. These operations work by deletion and generalisation of irrelevant or redundant propositions, and by construction of inferred propositions. Macrostructures are hierarchical, and macro-operations are applied in cycles, with increasingly stringent criteria of relevance.

The function of macrostructures, therefore, is the organisation of micro-information into a globally coherent structure, the reduction of complex information, and the definition of higher-level or global meanings that have been derived from lower-level meanings. This process may also involve the construction of new meaning (i.e., meaning that is not already a property of the individual constituent parts). Since texts represent courses of events, their meaningfulness depends on the correct conditional relationships being maintained in such sequences of events. However, textual proposition sequences are typically incomplete, and to varying degrees will require inferences in order to maintain coherence. The notion of sentential topic is different to that of text topic, in that a text topic is a macroproposition derived from propositions expressed by a sequence of sentences by applying various macrorules.

The full cognitive relevance of macrostructures appears in the representation of the text in long-term memory. Whereas in short-term memory they merely function as tentative global coherence links between propositions, their organisational power in long-term memory is demonstrated by the assumption that long sequences of propositions at the microlevel may be subsumed under macropropositions, and that in turn macroproposition sequences may be dominated by higher level macropropositions. The representation of the discourse in long-term memory thereby acquires a hierarchical structure, depending on the operations of short-term memory.

Complex information, such as that involved in the processing of text, must be organised and thereby reduced through the construction of higher level, global structures. Macrorules act as inference mechanisms that derive the macrostructure from the microstructure of a textbase (Van Dijk, 1980a). There are three main types of macrorule postulated:

[1] DELETION: A proposition is deleted if it is irrelevant for (i.e., not a condition for) the interpretation of subsequent propositions. A proposition is also deleted (or integrated) if it can be inferred from other propositions.

[2] GENERALISATION: Category names are substituted for category members, both with respect to arguments of propositions, and to the relational terms of propositions.

[3] CONSTRUCTION: A sequence of actions or events are summarised by introducing a name that refers to the sequence as a whole. Construction organises micropropositions by combining those sequences that function as one unit at some microlevel. It introduces information at the macrolevel which is "new" in the sense of not being part of the textbase nor entailed by the individual propositions of the textbase.

The early version of the model emphasised the processing and establishment of coherence at the microlevel. The later versions of the model emphasise much more the establishment of coherence among macropropositions, through the use of superstructures and world knowledge structures (cf. Van Dijk & Kintsch, 1983; Miller, 1985).

There is some evidence in support of the assumptions about the role of short-term memory and the processing cycles hypothesised by Kintsch and Van Dijk (1978). Evidence for the relevance of sentence and clause boundaries in short-term memory for discourse has been provided by Jarvella (1971, 1979). A sequence of sentences was interrupted at various points and subjects were asked to recall verbatim as much as they could. Jarvella found that subjects usually recalled the last clause, with sharp breaks in performance at clause boundaries, indicating that the syntactic structure of sentences was being used by subjects. On average, about two simple sentences were retained verbatim. Glanzer, Dorfman and Kaplan (1981) obtained similar findings, and also found that procedures that are traditionally used to interfere with short-term retention in list-learning studies have similar effects on discourse.

Fletcher (1981) used the Kintsch and Van Dijk model to predict what propositions would be selected for inclusion in the short term buffer. Using a cued recall task, he showed that propositions predicted by the model to be retained in the short-term memory buffer were in fact more available to subjects, even when they came from a previous clause, than those not predicted to be in short-term memory. Fletcher's study also suggested that short-term memory includes structurally important text propositions that are needed to establish a coherent text representation at the conceptual level.

The heavy reliance on referential coherence, or argument overlap, in the Kintsch and Van Dijk (1978) model has been criticised by Johnson-Laird (1983) on the grounds that it is possible for statements to be coreferential and not coherent, and that coreference is not necessary to produce coherence. The inadequacy of the principle of referential coherence has also been demonstrated empirically. Garrod and Sanford (1982) measured reading times in which the first sentence either explicitly mentioned the referent of the second sentence or merely implied it. According to Kintsch and Van Dijk, if there is no argument overlap, inferences would be required to "fill the gap", which would take up time. However, Garrod and Sanford found that it took no longer to comprehend the sentence when it was not referentially coherent than when it was. Keenan, Baillet and Brown (1984) showed that the integration of text units involves more than establishing referential coherence. They found that even when texts are referentially coherent subjects try to establish knowledge-based coherence. The causal relatedness of items in the text also affected comprehension times — the stronger the link the faster comprehension was.

Van Dijk and Kintsch (1983) point out that macrostructures are cognitive structures, and not purely linguistic ones: "It is more appropriate to account for meanings, and hence also for global meanings, as being *assigned* to a discourse by language users in the processes of understanding or interpretation" (Van Dijk &

Kintsch, 1983, p. 193, emphasis original).

Macropropositions have been treated in different ways in various versions of the early model, and the evidence concerning macroprocessing has used different versions. For example, in Kintsch and Van Dijk (1978) the cycles in macroprocessing represented successively higher levels of macrostructure hierarchy — thus a macroproposition at the top of the hierarchy was said to have participated in more cycles than a macroproposition near the bottom. Therefore, the reproduction probability for macropropositions was based mainly on structural factors. However, Kintsch and Vipond (1979) treat macroprocessing in the same way as microprocessing. Cycles are approximately paragraph length and macropropositions have their own "macrobuffer". The reproduction probability for a macroproposition would presumably be the number of times it was selected for this buffer.

Van Dijk (1980a) proposes that the reader makes a hypothesis about the macropoposition that is currently relevant as soon as one or more sentences provided enough information. He does not consider each processing stage as discrete, but rather suggests that the partial products of any one stage are available for further (higher level) processing. (This is the notion of "continuously available output" postulated by Norman and Bobrow, 1975.) The assignment of macropropositions is held to be determined not only by information in the text, but world knowledge and context also generate assumptions about the possible topics of the discourse. Hypotheses about topics can also be established by specific signals or cues in the text.

The Kintsch and Van Dijk (1978) model predicts that if the reader maintains coherence by means of memory searches, reinstatements and inferences, this will be reflected in increased reading time. On the other hand, if these processes are not undertaken, recall performance should suffer, since the reader's representation of the text will itself be incoherent. Vipond (1980) used texts with varying levels of difficulty and constructed lists of micro- and macropropositions according to the model. He found that the number of reinstatements, the number of propositions reinstated, and the number of levels in the micropropositional network were highly correlated with both micro- and macrocomprehension. Microinferences and the number of levels in the macropropositional network were correlated with microcomprehension, but macroinferences and reorganisations were not correlated with comprehension. Vipond suggests that macrostructure inference may be made automatically.

Spilich, Vesonder, Chiesi and Voss (1979) studied text processing in two groups who differed in prior knowledge of a particular domain (baseball). "High-knowledge" individuals were found to have a greater ability to relate the actions of the game to the goal structure, and to maintain the most important information in working memory. In general, high-knowledge subjects recalled more information that was goal-related than the low-knowledge subjects; they tended to integrate sequences of goal-related actions more than low-knowledge subjects; and they recalled more information in the appropriate order than low-knowledge individuals. Application of the Kintsch and Van Dijk model supported this interpretation, and suggested the operation of processes involving both microstructure and macrostructure information.

Spilich et al. (1979) also considered two variations of this model: The "independence model" assumed macrostructure propositions to be in a macrostructure buffer, while micropropositions were assumed to be in a micropropositional buffer. The second variation of the model assumed a separate macrostructure buffer, although macropropositions were permitted to have microproposition status; thus, when appropriate, macropropositions could be carried over in the microstructure buffer as well as in the macrostructure buffer. This second version was found to have a better fit.

Vipond (1980) also tested the relative independence of micro- and macroprocesses. Several types of dependence are possible. The Spilich et al. (1979) study investigated two types of partial independence. Vipond suggested another sense of dependence, in which micro- and macroprocesses can be viewed as independent or non-independent according to how they are allocated cognitive resources. One possibility is that each process has its own separate resource pool. This implies that extra difficulty in processing on one level would not affect the other level. (This was the assumption of both the Kintsch and Van Dijk (1978) and the Kintsch and Vipond (1979) versions.) Alternatively, micro- and macroprocessing may share resources at some stage in processing, resulting in an extra load at one level influencing the other. Vipond's study suggested that micro and macroprocessing do not share the same resources, and are therefore independent.

Cirilo (1981) investigated reading times and recall for sentences in a text, as a function both of their height in the textbase hierarchy, and of the amount of coreference they had with other sentences in the text. He reasoned that an input sentence can be connected to the rest of the text via immediate matches, based on the contents of the short-term store; via reinstatement matches, based on retrieval from long-term memory; or via inferencing, when the contents of memory are insufficient to establish coreference.

A sentence's reading time will vary depending on the type of connection it requires. This is the "distance effect". These three connections were examined by varying the context of single target sentences in stories. The targets either coreferred with a precursor, or they did not, thereby testing for the inferencing component. In the coreferential case, the distance between the target and the precursor was also varied, manipulating the probability of reintstatement versus immediate matches.

The Kintsch and Van Dijk (1978) model suggests that macropropositions should be available longer in short-term memory for making immediate matches. Therefore, Cirilo reasoned that a sentence should take less time to process if it is connected to a high-level macroproposition rather than to a low-level microproposition. This is the "height effect". It was tested in this study by varying the global relevance of precursors.

In a task emphasising microprocessing, it is expected that effects reflecting only microprocesses should be obtained, while effects reflecting macroprocesses should not. Thus the distance effect should appear, but the height effect should not. An attempt was made to realise such an emphasis by having subjects read coherent fragments from stories and then recall them immediately afterward. In a task emphasising both micro- and macro-processing, the distance effect should be present, and so should the height effect, with high-level sentences being recalled more frequently than low-level sentences. This was realised in the study by having subjects read the whole story and recall it immediately. Under a macroprocessing emphasis, distance effects should disappear, but a height effect should be found. This emphasis was created by having subjects read stories as casually as possible, by using stories as filler material in a larger experiment, and then administering a surprise recall test.

Cirilo found that greater distances (between target and precursor) and lack of coreference increased reading times, but only when microprocesses were emphasised. Reading times were faster when the precursor was structurally important, but only when macroprocesses were emphasised. The results showed that in conditions emphasising both micro- and macroprocessing, high-level precursors were recalled more often than low-level precursors, whereas in conditions emphasising microprocessing there was no difference in recall.

As well as evidence for the processing aspects of the model, there has also been some experimental support for the macrostructural aspects. For example, Shebilske and Reid (1979) found that eye movements in reading are influenced by comprehension processes that link sequential conceptual units together to form macropropositions. They predicted that when lower level conceptual units are directly related to the underlying structure, fewer processing resources will be required, and reading will progress faster. They investigated whether or not readers can predict macrostructure at any given moment while reading, on the basis of what they had already read. They found that higher importance ratings were given to sentences that were salient in the macrostructure of various passages, according to recall data obtained by Van Dijk (1975) and Thorndyke (1977). They also found that sentences that were indirectly related to macropropositions were read slower than sentences that directly expressed a macroproposition, due, they argued, to the extra processing resources required to maintain and integrate sentences into macropropositions. Kieras and Bovair (1981) also found that if a summary sentence (a macroproposition) is expressed in a paragraph, reading times, importance ratings, and online protocols indicate that people can recognise it as such.

Guindon and Kintsch (1984) performed a study to examine the formation of macrostructures. They reasoned that since macropropositions are formed and stored in memory with a greater probability than micropropositions (based on the processing model of Kintsch and Van Dijk, 1978) they should show greater priming effects. (Ratcliff and McKoon (1978) had shown that two words belonging to the same sentence produce greater priming effects than if they do not.) Guindon and Kintsch also reasoned that if the macrostatement from the previous paragraph were deleted, subjects should infer something like it, and therefore that test words from the macrostatement would be harder to reject on a recognition test than distractor items, leading to higher false-alarm rates as well as slower reading times. They found strong macrostructure priming effects, whether or not subjects were required to write a summary before being tested.

Evidence has also been found for the operation of macrorules. Kieras (1980) found that subjects were faster and more consistent at abstracting main ideas from generalisation-based passages than for construction-based ones. Brown and Day (1983) found a developmental trend in the use of macrorules, with deletion rules emerging first, followed by generalisation. Construction rules were the most difficult, and latest in developing.

Although Kintsch and Van Dijk (1978) gave an evaluation of the microstructural components of their model, they did not give an account of the knowledge-based aspects of comprehension. (Some studies have included knowledge-based macropropositions in their analyses — e.g., Spilich et al. (1979) — but these were derived from the experimenters' intuitions about the text being analysed.)

Miller and Kintsch (1981) attempted to address problems with the first model by making two additions to the coherence graph system: Two conceptual levels were added to the model — one more complex and one less complex than propositions, and all of these levels have access to an explicitly defined knowledge base. They conducted an experiment in which subjects were given segments of a text that were divided at phrase boundaries, and were asked to write down what they "thought would happen next". After a short time, the next segment was revealed and subjects again wrote a continuation for this segment.

The scoring of the segments was grouped according to whether the segment being predicted was at the beginning or the end of a paragraph, the beginning or end of a sentence, or a segment in the middle of a sentence. The results suggest that, as the reader progresses through a paragraph, the possible topics that might be discussed, and the ways that the text might develop, become increasingly constrained, so that accurate sentence continuation becomes increasingly likely. When continuation accuracy in both easy and difficult texts was considered, subjects were found to be more correct in the easy text than in the difficult text, especially at the end of a paragraph. This suggests that texts can be difficult for two reasons: Difficult texts are less constrained than easy texts, so that there are more ways in which a difficult text can develop while remaining acceptably coherent; the constraints established in difficult texts will not always reflect the true course of the text.

Miller and Kintsch (1981) also found that subjects read almost as fast at the beginning of a paragraph as at the end, suggesting that, at the beginning of a paragraph, a reader knows that the text is not yet constrained in any way, and so reads the beginning of a paragraph quickly, in order to get an idea of the text's content which might be useful in deriving future constraints. In addition, once readers are past the beginning of a hard paragraph, the complications of the text seem to slow them down, at least until the paragraph's end.

Miller and Kintsch suggest that the absence of appropriate knowledge did not keep the reader from using constraints in comprehension, but simply denies the reader access to a particularly powerful set of constraints. In the absence of top-down constraints, the comprehension of the text should be affected by those constraints at the level of individual sentences or propositions (e.g., linguistic signals). The appearance of a sentence in different contexts affects these constraints in the following way:

a) *Sentence in isolation*: When no context is present, the only possible constraints are those imposed by the sentence.

b) *Sentence embedded in a meaningful context*: Here the significant constraints should be those characterised by the context. The point is not that the content of the target sentence is irrelevant, but that the constraints derived from the sentence are ignored in deference to the more powerful constraints of the currently relevant knowledge structure.

c) *Sentence embedded in a nonspecific context*: Here the target sentence can be embedded in the text without yet specifying an explicit interpretation. Such a text need not be incoherent, but merely structured so that no single topic is identifiable as the central idea of the text. In this case, no global constraints are present, and any expectations that are generated should reflect the constraints specified by the local properties of the text.

Miller and Kintsch conducted an experiment to test these hypotheses. Subjects read target sentences either embedded in meaningful contexts, or in nonspecific contexts, and were asked to perform the continuation task, just as in the previous experiment. The results indicate that global constraints specified the interpretation of the meaningful context condition, but that local constraints specified the interpretation of both isolated sentences and texts in the nonspecific conditions.

Signals of Topicality

Although macropropositions are usually derived from or assigned to the text, they may be expressed in the text itself as signals of topic or theme, usually occurring at the beginning or end of the text, or its episodes or paragraphs. Van Dijk and Kintsch (1983) argue that the initial or final position of topical expressions follows from their semantic functions. At the beginning, they help the reader to form a hypothesis about the topic, so that subsequent sentences can be interpreted relative to that macroproposition. In the final position, topical expressions serve to check, remind, or correct the already established macropropositions, as well as to repeat what is already known, so that the reader has not only hierarchically dominant information, but also recent information to search for when the topic is needed in subsequent tasks.

Various cues at several levels can serve to explicitly signal the relevance, importance, or topicality of the text: graphical signals (e.g., type size, indentation, italics — these kind of explicit signals are discussed in more detail in Chapter 6), syntactic signals (e.g., word order, cleft constructions), lexical signals (words such as "important", "in other words", "the subject is", "the conclusion"; connectives such as "so", "thus"; superstructural signals, such as "the hypothesis is", "suddenly"), semantic signals (e.g., topic-comment functions, contrastive structures, summaries, level of description), pragmatic signals, rhetorical devices, and so on. (See Van Dijk (1979), for a more complete list.)

Evidence for the role of superstructural signals such as type of connection has been found, for example, by Smith and Frawley (1983), who analysed different text types according to the frequency of kinds of conjunctions occurring in them. However, Meyer (1985) notes that, although signals in the text can help readers hypothesise about which schema to assign to a text, the presence or absence of signalling in a well-organised text has minimal effects on readers who specifically use a structural strategy in comprehending the text (cf. Meyer, Brandt & Bluth, 1980; Meyer, 1984). Meyer and Rice (1982) found that the effects of explicit signals also depended on the type of text, as well as reader strategy.

Givón (1983) has suggested that various syntactic constructions are organise in a hierarchical continuum, which indicates the degree of topic continuity in a discourse. Constructions at one end of the continuum are used to signal that the current topic is the same as the previous topic (e.g., unstressed pronouns) while constructions at the other end are used to indicate changes in topic (e.g., cleft constructions). What determines the ordering of the various syntactic constructions is their markedness or explicitness. Fletcher (1984) investigated this notion in an experiment where subjects were asked to rewrite two short sentences as one single sentence. The form of the referent in the second sentence was found to depend on its continuity with the topic of the other sentence. Subjects tended to use the unmarked linguistic form in cases of high topic continuity, and marked forms when there was a topic shift. In another experiment, subjects were asked to interpret text fragments that contained ambiguous referents. The less marked the form of the referent was, the more likely subjects were to interpret it as coreferential with the preceding topic.

It has been found that the surface subject of a sentence usually marks the sentence topic. Perfetti and Goldman (1974, 1975) found that readers assign the topical referent of a passage to the surface subject position of a sentence. Van Dijk (1979) has noted that the assignment of items to either topic or comment positions in a sentence will be determined by the global topic of the passage. However, Kieras (1981a) showed that topic-comment assignment could also influence passage topic assignment. He notes that signals of topic or theme, such as titles, etc., are only weak ones. Schallert (1976) showed effects on recognition, and not on recall, in using passages with two possible topics and manipulating the title (cf. also

Kozminsky, 1977). Meyer (1977) also found only weak effects of marking phrases. Kieras (1985) reports an experiment in which he used passages with or without titles or marking phrases, marking either a thematically strong or weak item. Although he found an effect of initial mention, there was no other marking effect. Kieras argues that when the reader already knows the thematic content, marking is considered redundant, and that position effects are important markers rather than explicit signals.

Miller (1981) used a continuation task and manipulated semantic and syntactic focus, and found that subjects tended to focus their continuations on the agents of the final sentence of the text, the subject of the final sentence, and the part of the final sentence receiving semantic focus from the second sentence. The reading time for the final sentence was also found to depend on the agreement of that sentence's subject with the focus created by the semantic and syntactic focus manipulations.

Kieras (1981a) found that organising the passage around a single referent has a strong effect on the difficulty of identifying the topic, and that which referent appears as the surface subject of individual passage sentences is also a strong determinant of the perceived passage topic. According to Kieras, global coherence is required because readers can only process and store a limited amount of information while reading a passage; having the global topic in mind allows them to restrict their processing to selecting or inferring macropropositions about that single topic, while storing only the most important of them. Surface-level signals can be used because they require minimal processing resources, whereas if readers had to infer the macrostructural content strictly on the basis of the "deep" content of the passage, they would suffer the heavy processing load required at that level.

Kieras' study (1981a) was intended to show that violation of the global coherence rule would result in processing difficulties. He found that when a passage contained more than one major referent, readers could not pick the most frequently mentioned referent. They had to perform an extensive memory search and execute inference processes in order to arrive at a single-topic response.

According to Kieras, when there is only a single major referent, the passage macrostructure is built around this referent, and so supplying a statement of the main referent is simply a matter of selecting this central component of the macrostructure. However, when there are several major referents, the macrostructure for these passages consists of various thinly-connected parts, each built around its own central referent. In order to supply a single referent as the topic, the reader must engage in further macro-level processing to construct a higher level set of macropropositions that are organised around a single referent and tie together the separate parts of the original macrostructure. Kieras found that if one of the most frequent referents appeared in the first sentence as the sentence topic and reappeared as the subject thereafter, it was very strongly perceived as the passage topic.

One implication of macrostructure theory is that the structure of a passage determines not only how things are stored, but also how they are retrieved later. When given a cue for recall, the reader looks first under the corresponding "address" in memory for the passage information. If the desired information is found, recall proceeds smoothly. If not, the retrieval attempt will be disrupted, and a different retrieval strategy must be used, resulting in poorer recall. Hence the originally perceived topic of a passage will influence the effectiveness of a recall cue; a cue that matches the topic should be superior to one that mismatches.

Several studies (e.g., Meyer, 1977) have found that even though the most important content normally appears first in a passage, initial mention influences recall because it has topic-signalling value in itself. Kieras (1981b) investigated whether marking an item as the passage topic by initial mention and cueing recall with either the topic item, or a non-topic item, affects recall for propositions about passage items. He found that if the cue is the passage topic, recall favours the topic, but if the cue is not the topic, then recall of topic and non-topic information is equal. The implication of this finding is that marking an item as the passage topic does not govern what the reader stores. Rather, the original passage topic seems to determine what the effect of the later recall cue will be.

In a follow-up experiment a different passage topic marker was used — sentence surface subject assignment, which allowed different candidate topic items to be marked as topical by surface changes in the sentences, without changing the order of information in the passage. He used two types of passage: "balanced passages" were chosen to show little bias in favour of one of the two items; "biased passages" were used to show a strong topic bias. Propositions about topic items were recalled better than nontopic items; for the topic cue, recall of the topic item was superior to the nontopic item, and for the nontopic cue, the recall of the two items was similar. The patterns were similar for the balanced and surface subject passages, however, examination of the results for the biased passages showed the superiority of more specific topic items. The order of recall was closely related to the order of information in the presented passage: For the balanced and biased passages, the effects were due to the item marked as the topic, whereas for the surface subjects passages, the effects of topic and cue were equal. This implies that recall order is determined primarily by presentation order in the original passage. The recall cue has little or no influence on recall order. Thus, subjects recalled a passage in roughly the same order as it originally appeared, regardless of what cue was used.

Kieras' results (1981b), where recall of an item was highest if it was the cued item, regardless of which item was topicalised by the passage, imply that topicality and memorability are distinct characteristics. He found that general topic items were favoured as passage topics in a topic choice task, but specific concrete items were better recalled. This argues that standing in the passage macrostructure is not the only determinant of recall.

He also found that although there was little surface-level retention of the passage, there was a strong similarity between input and output (recall) order. This might have been an effect of macrostructure. If the passage macrostructure had a tree-like form that was recalled starting with the general information at the top of the tree and proceeding to the detailed information at the bottom of the tree, then the order of output should resemble the order of the propositions in the original passage, even without retention of the actual surface form.

However, macrostructure theory — in which topicalisation governs storage properties — predicts that overall recall would be best for topic cues and worst for non-topic cues. In this study (Kieras, 1981b), there were no main effects of cue: What was affected was not how much was recalled, but actually what was recalled. It appeared that subjects had access to the topic item even when it was neither marked as a topic nor cued. Kieras suggests that topic marking may be an aspect of the passage that the subject stores independently of retained propositions. If the reader is asked to recall a passage about a certain item (the cue), and also remembers that this item was marked as the topic, then the task is interpreted as meaning that this topical item is more important to recall than the nontopic item, and so it is favoured in recall. If the cued item is not the topic, then there is no need to favour one item over another. Thus, both the passage topic and recall cue act as instructions to the subject about what portions of the passage should be recalled.

IMPORTANCE, RELEVANCE, AND PERSPECTIVE

As was noted at the beginning of the previous section, there has been considerable confusion over the use of terms like topic, theme, and gist. The concepts of sentential and discourse topic just discussed capture the notion of "aboutness". However, this is also an inadequate conceptualisation, since "about" may mean "refers to", or "denotes"; on the other hand, it may mean "perspective" or "point of view".

Terms that have been used more or less synonymously with the concept of "topic", but which also imply perspective, or point of view, are "theme" and "focus". Theme was used by Halliday (1967) to refer to the "point of departure" for the rest of the sentence (i.e., similar to the topic part of "topic-comment"). Others have treated it as a much higher level concept such as "superstructure" or "script". "Focus" has been used by Chomsky (1970) for the part of the sentence that is not presupposed. Van Dijk (1979) appears to equate it with "comment". Reichman (1978), Grosz (1979), and Sidner (1983) use "focus" in the intuitive sense of topic or theme (i.e., "aboutness" in the sense of "denotes"). Grimes (1982) uses the term "reference spaces" to denote ways in which speakers treat some things as more central than others. This term is based on Reichman's notion of "context spaces" (Reichman, 1978).

Reichman (1978) found that pronouns sometimes don't get used in places where they would be appropriate. She presents a conversation in which the ordinary rules of pronominalisation showed that a pronoun should be used, but where a proper noun is used instead, several times in the same conversational turn. She argued that the less central in the "context space", the more explicit the type of mention of a referent that needs to be made, even if this principle must override the ordinary principles of pronoun use.

Grimes (1982) postulates that a text has one kind of structure, amongst others, whose basic units are "reference spaces". Within each reference space one pattern of reference holds true. As the "core" of the text develops, various mechanisms preserve continuity, including the intersection of referents between successive clauses and the referential component of macrostructures. He proposes that a "cell" is the simple minimal form of a reference space. Complex reference spaces consist of more than one cell, all of which share the same "core", but are separated from one another by other embedded cells. Grimes sees two kinds of relations as holding between successive cells: Development relations are where the reference space changes and the new space is not an interruption of the preceding one, but replaces it without the option of resuming the former space. Support relations are those in which one cell is elaborated on or explained by another in such a way that the cell that follows the supporting cell could resume the one that preceded.

Van Dijk (1982) is critical of Grimes' theory, in that it confuses the notions of importance (or relevance) with perspective or point of view. (Grimes (1975) initially used the term "staging" to refer to perspective or point of view.) The difference between perspective and importance, relevance, or centrality appears to be at least psychologically real. It is a basic assumption of schema theory that information may vary in importance depending on the schema used in encoding or retrieval (cf. Rumelhart & Norman, 1978). Pichert and Anderson (1977) found that the type of important information recalled was dependent on the perspective used to encode it. Anderson and Pichert (1978) found that subjects encoded more information related to a nonencoded perspective than they did when they recalled it first from the encoded perspective. However, these experiments concern the role of the reader's perspective in processing the text (a subject dealt with in more detail in the next chapter) rather than the perspective or point of view in the text (i.e., that generated by the author).

There are various cues in the text itself, such as titles, that indicate the author's perspective or point of view (cf. Bransford & Johnson, 1972; Dooling & Mullet, 1973). These studies showed that improvement in comprehension was found only if the title was presented before the text; if it was presented afterwards, readers did no better than those who had received no title. This supports the encoding strategy explanation for the role of perspective. However, Anderson and Pichert's (1978) results suggested a retrieval strategy. Fass and Schumacher (1981) showed that if retention interval was lengthened the results supported an encoding rather than a retrieval strategy.

Other findings suggest that titles have qualitative rather than quantitative effects: for example, subjects may mistakenly recognise sentences not contained in the text, but suggested by the title (Schallert, 1976; Kozminsky, 1977). Schwarz and Flammer (1981) found that the presence of a thematic title facilitated comprehension of a text that was slightly disorganised. Bock (1980) reports an experiment in which the title affected the selection of content words judged important, that the title raised the hierarchical positions of some propositions, and lowered that of others (where position was derived from importance ratings); and that propositions made "important" by the title were better recalled.

The Levels Effect

It appears from the results of several studies concerning what is known as the "levels effect", that the reader's internal representation of the text (the "textbase") is hierarchically ordered in terms of "importance" of its constituent propositions — where "importance" is in general equivalent to "relevance" to topic or theme. Evidence for this includes the findings that subjects tend to recall the most important or relevant propositions; that subjects tend to verify "important" test statements faster and more accurately than they do subordinate propositions; and that subjects tend to display considerable agreement over which of the propositions are important in relation to the others.

One of the major problems in testing for recall of "important" or topical information is in deciding which propositions are the most important. Many researchers have settled on intuition as a basis for deciding this, and argue that this is justified, since there is a high degree of inter-judgmental agreement about the relative importance of information within a text (cf. McKoon, 1977). Van Dijk

(1979) defines relevance as the result of an operation by which a reader assigns some degree of importance to some property of the discourse. He argues that it is a relative notion, which must be construed with respect to a certain text and context; it is also a contrastive notion, in that if some properties of the text are assigned relevance, others are not. However, as Van Dijk notes, the fact that people can and do assign relevance or importance to certain parts of the text does not explain how that relevance is assigned.

Given that there is some sort of hierarchy being used in comprehending and recalling a text, what kind of processes might be operating? Meyer and McConkie (1973) asked subjects to outline a passage and then converted the outlines to tree structures. From these structures, they developed measures of the importance of an "idea unit" in the structure of the passage: a measure of how high up in the hierarchy the unit occurred; a measure of the number of units that were beneath a given unit in the hierarchy; and a measure that combined these two. Important units were better recalled than unimportant units on all three measures. Serial position and rated importance were found to be correlated with hierarchical importance.

In a study by Kintsch and Keenan (1973) subjects were presented sets of sentences that were analysed into hierarchical propositional representations. Both propositional rank (analogous to Meyer and McConkie's "hierarchy depth score") and number of subordinate propositions (analogous to the "units beneath score") affected recall. The higher up in the hierarchy a proposition occurred, or the more propositions directly beneath it, the more likely it was to be recalled.

It is possible that important material is more likely to be remembered better because it is more likely to be encoded during reading. This was tested by Meyer (1977), who gave subjects a free recall test immediately after reading a passage in which a target paragraph was high or low in the content structure. She found that the target paragraph was recalled better when it was important; cued recall aided the recall of both important and unimportant material, but did not reduce the advantage of the important versions. Meyer argued that this result supports an encoding hypothesis (cf. also Meyer & Rice, 1982), because if the poor recall of unimportant material was due to retrieval failure, the introduction of cues should decrease this deficit by reducing the difficulty of the retrieval task. Meyer's results suggest that it is not simply the text structure per se that is influential on the levels effect, but the context in which the text segment is presented.

Another possibility is that important material may be remembered better because it is less susceptible to loss through forgetting. Meyer (1975) had subjects read and immediately recall a passage. A week later they were given a second free recall test and a cued recall test. When scores from the second free recall test were subtracted from the first, passages in which the target paragraph was important showed less forgetting. The introduction of cues produced better recall than either free recall tests, but it did not reduce the advantage of important versions.

A third possibility is that important material may be easier to access or more retrievable. In other words, the reader may select out a core of important material for "deep encoding". Because important elements are more deeply encoded, they are less subject to forgetting, an advantage that increases with retention interval (cf. Meyer, 1975). Evidence suggesting that the levels effect may be a retrieval rather than a storage effect has been found by Britton, Meyer, Hodge and Glynn (1980).

These hypotheses imply that subordinate propositions will be verified less accurately and more slowly than superordinate propositions. McKoon (1977) found that the importance of information in a text affected both the speed and accuracy with which that information was verified. Topic information was verified faster and more accurately than detailed information. The effect was significant when testing was delayed, but not when testing was immediate. Furthermore, neither the length of the text from which the tested information came, nor the serial position of the tested information in the text affected the speed or accuracy of the verification.

It may be that superordinate retrieval is faster and more probable because superordinate propositions are connected to many more propositions, on average, than subordinate propositions. However, some doubt is cast on this explanation by the results of recall experiments which show that probability of recalling a proposition does not depend on the number of connections to that proposition (Kintsch & Keenan, 1973). On the other hand, it may be that subordinate propositions require more retrieval time than superordinate propositions because propositions are retrieved in order, beginning with the most superordinate, and ending with the most subordinate. Another possibility is that the encoding of a subordinate proposition is less elaborate, or less complete, than the encoding of a superordinate proposition. The choice of either an encoding or a retrieval explanation depends upon which theory of text structure is adopted. A hierarchical theory of structure is compatible with either an encoding explanation or a retrieval explanation, whereas a theory that proposes a non-hierarchical structure (e.g., Anderson & Bower, 1973) requires an encoding explanation.

Manelis (1983) suggested that one factor underlying the distinction between high and low content positions is amount of elaboration (cf. also Rothkopf, 1977). In other words, the more information is elaborated (hence, subordinated), the more important the superordinate. Manelis had subjects rate the amount of elaboration in Meyer's (1975) passages and found that subjects judged the target paragraphs in the high-content version as more elaborated than in the low-content one. He then asked subjects to sort paragraphs from Meyer's texts in terms of hierarchy, and found that the relative number of immediate subordinates was greater when the target paragraph was in the high-content than the low-content version, but there was no significant effect of level. However, Meyer and Rice (1982) specifically manipulated amount of detail independently of signalled level, and found that detail did not influence the levels effect, nor the organisation of protocols, for certain text types (those with problem/solution structures), although it did have an effect on other types of text (those with comparison structures).

The arguments in favour of elaborations hinge on the fact that the deeper the level of processing (cf. Craik & Lockhart, 1972; Cermak & Craik, 1979), or the more elaborative the processing (Anderson & Reder, 1979), the better the material is remembered. Reder (1982a) suggests that elaborations redirect activation away from interfering facts in memory and towards the to-be-remembered information, that readers can elaborate on cues or probes at test to generate additional concepts from which to spread activation, and that elaborations allow for inferential reconstruction of the to-be-remembered information. (These hypotheses are based on ACT theory — cf. Anderson, 1983.)

Reder and Anderson (1980) tried to demonstrate that the important facts of a text can be considered elaborations of the main points and can facilitate retention of these important facts. The main points may be remembered better due to their position in the logical structure within a text. Meyer & McConkie (1973) found that centrality in a hierarchy is a better predictor of recall than a subject's rating of importance of propositions. On the other hand, as Reder and Anderson suggest, it may be that these findings are due to the fact that the reader implicitly recognises the importance of the central points and assigns greater capacity to their processing. Elaborations may actually hinder the acquisition of important points, especially if there are too many facts to remember, in which case readers would have to devote some of their processing capacity to unimportant facts. It may also be harder for readers to appreciate or extract the important points if they are embedded in details. Hierarchical analyses of text structure suggest that access to details is through the higher level nodes, and thus dependent on recalling the main points, not the subordinate ones. However, elaborations may allow the reconstruction of the main points, if they imply them, and access to central points may be available sometimes only via details.

Reder and Anderson (1980) performed a pair of experiments that tested subjects' recall of a piece of text or a summary of the text, both immediately after study and after a delay interval of one week. Subjects were given questions that could be answered on the basis of a summary. They found that summaries yielded higher recall than texts. Furthermore, summaries maintained their advantages at retention intervals from twenty minutes up to twelve months; summaries were superior both for questions taken directly from the text, and for inference questions which required the subject to combine facts that had been studied. Moreover, summaries yielded better transfer in a task that examined ability to learn new, related material as a function of how previous material had been learned from a summary. Superiority was maintained even when the main points in the text were underlined. Reder and Anderson concluded that their findings of the advantages of summaries over texts present a problem for their elaboration hypothesis of memory. The data are however consistent with a hierarchical analysis of text.

The evidence for the facilitatory effect of elaboration comes from studies where subjects themselves made the elaborations (e.g., Schallert, 1976; see Reder, 1980, for a review). However, Reder (1982a) argues that elaborative effect can be found in author-generated texts where they are redundant, and where they allow inferential reconstruction of the to-be-remembered information — in other words, they have to be related to the facts. Bradshaw and Anderson (1982) had subjects learn a given fact either in isolation with an unrelated fact, or with a related fact. After a retention interval of one week, they found that important or central facts were better recalled in the isolation condition than in the unrelated condition, but that the related facts produced better recall than the other conditions.

In general then, there are two alternative accounts of the "levels effect": The "top-down" approach emphasises the saliency of textual cues, and suggests that the reader is guided by a representation of the text structure in order to process high-level sentences more thoroughly. An alternative "bottom-up" approach relies in part on a form of the "given-new" hypothesis, and emphasises the difficulty of processing a proposition due to its lack of connection to the earlier text. High-level propositions typically introduce new material, thus there are few bridges between the concepts involved in the high-level proposition and the concepts already introduced in the text. The content of the high-level propositions may be less predictable, and therefore require more processing than low-level propositions, which tend to elaborate on already established ideas.

Cirilo and Foss (1980) tested these two alternative accounts. They hypothesised that a sentence will require more processing if it is high in the hierarchy than if it is low, either because high-level sentences are cued as such, or because they are more difficult to integrate. Reading time was used as the measure of processing difference. Single sentences were taken from structured stories and placed in new stories at a different hierarchical level. It was predicted that a sentence at a high level in one story would take longer to read than the same sentence at a lower level in another story. The experiments were also designed to investigate the effect of serial position of a sentence on reading time. If the processing of most new information occurs relatively easily, then later information will be easier to integrate and hence faster to process.

Cirilo and Foss found that high-level units took longer to read than low level units. An interaction between height and position was also found: The early parts of the story took longer to read than the later parts, although this effect appeared to be restricted to the high-level units. Kintsch and Van Dijk (1978) accounted for these "levels effects" in memory by reference to the number of processing cycles in which a proposition occurred, or correspondingly, the amount of time that a proposition remained in the short-term memory buffer. However, this study indicates that readers do not spend equal amounts of time processing high and low-level propositions when they are first encountered.

Cirilo and Foss (1980) suggested that subjects might be able to determine which propositions of a story are important (and therefore should be processed more thoroughly) as they are encountered, because text processing is guided by story frameworks or schemas, and because the texts may contain cues concerning the role that the various sentences play. Alternatively, they suggested that high-level sentences might require more processing time because they are more difficult to integrate into the existing structure that the reader is building. (However, some propositions in a text may see important when first encountered, even though they are not central to the final representation.)

Some findings appear to contradict the levels effect. Britton, Meyer, Simpson, Holdredge and Curry (1979) found no differences in reading times of high and low-level paragraphs. Yekovich and Thorndyke (1981) found no levels effect in recognition tests. However, these findings are still compatible with the model of Kintsch and Van Dijk (1978), which attributes the levels effect to multiple processing of high-level information in the STM buffer, i.e., to more practice allotted to high-level propositions.

The finding of differential level effects for recall and recognition is supported by a retrieval model (e.g., Meyer, 1975; Britton et al., 1979; Thorndyke & Yekovich, 1980), where text information is stored hierarchically with an equal amount of encoding of all of the text propositions. The retrieval process is a top-down search of this hierarchy. High-level propositions are recalled, but not recognised better than low-level information, due to access to retrieval starting with the high-level information. Although this model explains the recognition data, it doesn't explain the cued recall data (Meyer, 1975) nor the different reading times for high and low-level propositions (Cirilo & Foss, 1980).

Criticisms of the Levels Effect

Stein (1982) is critical of most studies of "importance" for their reliance on recall data. She notes that several studies have found that some events in narratives that have been identified as "central" or "important" are often not recalled — for example, goals and emotional reactions are frequently deleted in recall even though they are rated as important on judgement tasks (Stein & Glenn, 1979).

The main criticisms of the "levels effect" concern the status of the concepts "important", "relevant", or "high-level", and (a related problem) the nature of the hypothesised structure. As Van Dijk (1982) notes, there are several *senses* in which a proposition is more important or relevant than others. There is a distinction between what is important at a particular point in the text (i.e., accounted for in terms of sentence structure and immediate sentence relations, such as the topic-comment relation) and what is relevant for a larger part of the text, or the text as a whole. One can distinguish between what is semantically important, i.e., in terms of content, and what is important for the author's main point (pragmatic relevance). Certain categories may be more important than others with respect to the particular text type (e.g., a complication in a story, a conclusion in an argument, and so on). Importance may depend on the particular rhetorical device being used — e.g., opposition and contrast. Mandler (1984) also notes confusion over different types of *levels* of importance. One sense of level is scope, which is a structural characteristic, referring to the number of nodes subsumed by a given node in the hierarchy. Within a branch, a node at a higher level has more scope than a lower level node on the same branch, but this doesn't hold for relationships between nodes at the same level across branches. Yet another sense of level is generality. Importance can also be a functional characteristic, depending on the type of hierarchy and what the text is used for.

Mandler (1984) also notes confusion over the types of *structure* proposed in studies purporting to show a "levels effect". One kind of hierarchy is a taxonomy or class inclusion structure, in which each unit is an example, or member, of the next higher unit. Another kind of hierarchy (e.g., the one that describes a story structure or an argument schema) is that of a collection, where each unit is a part of the next higher unit. The units are related directly to one another, and not merely associated by virtue of their membership in a superordinate class. Furthermore, there are different relationships between units: some are causal, some are temporal, and some atemporal.

There has also been confusion over the level of the structure hypothesised. Mandler (1984) argues that the hierarchical implications of microstructure, which describes the connectivity and semantic content of sentences, is not to be confused with the overall hierarchical structure, or macrostructure, which provides the gist of the text as a whole. Another confusion concerns the confounding of relative contributions of membership in a given unit, and the importance of an individual proposition to that unit. Mandler criticises the Cirilo and Foss (1980) study, where they used Thorndyke's (1977) grammar to place single sentences at either a high or low position in the hierarchy. In the example they used, the sentence occurring at a high level was the statement of the outcome of an episode, whereas the same statement at a low level formed part of a simple reaction — which, as Mandler points out, is typically faster to read and more poorly recalled than other statements in stories. Therefore, the effect could have been due to the height in the structure, or to the particular categories that were used, or to the relation of the individual sentences to their respective categories.

CHAPTER 4

FORM AND CONTENT

The Constructive Nature of Text Comprehension

A full account of text comprehension must also include the knowledge structures of the reader, the processes by which that knowledge is brought to bear in reading the text, and the strategies used to guide the processing of the text. Several studies have shown that the representation of the text in memory contains not only the information that was explicitly presented, but also implicit or inferred information. Moreover, subjects often appear to be unable to distinguish between what was actually presented and what was inferred. Bransford and Franks (1971) presented subjects with semantically related facts that were parts of a single complex idea. When tested, inferences (new combinations of the component ideas) were recognised as confidently as subjects recognised sentences that had actually been presented. Bransford and Franks concluded that learners spontaneously integrate the information expressed in related ideas, resulting in abstract semantic memory representations that contain more information than is actually presented.

Kintsch (1974) extended these findings by testing response latencies to explicit and implicit statements. He argued that learners make inferences during reading to fill the gaps in the text, and that these inferences are stored as part of the learner's memory representation for the meaning of the text. He predicted that accuracy in verifying test sentences, as well as response latencies, would be independent of whether the test information had been explicitly presented or only implied in the text. Results failed to confirm this, but Kintsch argued that explicit statements had the advantage of a short-term surface representation. When a delay was introduced the results confirmed the predictions. Kintsch concluded that readers may infer information during reading in order to integrate and comprehend explicitly presented propositions, and that inferred propositions are stored in memory as part of the textbase, which is the subject's propositional representation of the memory of a presented text. Thus, while Kintsch agreed with Bransford and Franks that both expressed and inferred information are present in propositional memory, he also argued that memory for surface features of a text may produce faster response times to explicit statements immediately after presentation.

Frederiksen (1975) found that repeated exposures of a text did not eliminate derived structures, indicating that inferences made at acquisition become an integral part of the learner's memory representation of the text. His data were interpreted as reflecting processes of adjustment to information overload during acquisition rather than as reconstructive processes occurring during recall. However, Moeser (1976) showed that semantically related sentences presented as discrete items are not necessarily incorporated into an integrated memory system. She found that varying specific aspects of the input situation and retrieval tasks appeared to have an effect upon whether or not integration would occur. Moeser concluded that independently presented verbal items can be stored as discrete memory units, and that when two items containing related information are stored independently, inferences based on the relationship between them cannot be retrieved.

In general, inferences can fulfill two functions. First, they make connections between surface structure fragments in the text (explicit propositions) and between propositions and knowledge already in memory, in order to establish coherence and continuity in the text. This is what Collins, Brown and Larkin (1980) refer to as the "text-based" view of inferences. The second approach is a "model-based" view that emphasises the constructive nature of comprehension, in that the central purpose of inferences is to synthesise an underlying model that organises and augments the explicit representation of the text. In other words, in this view the process of making inferences is guided by the knowledge structures possessed by the reader.

One text-based factor influencing inference making and integration is referential overlap. Garrod and Sanford (1977) found that reading time for a referring expression of the second sentence in a pair was faster if the content was more closely related to the referent. As was discussed, Kintsch and Van Dijk (1978) emphasise referential overlap as being necessary for coherence. Manelis and Yekovich (1976) addressed the issue of argument repetitions directly. In sentences where the same concepts repeatedly occurred across the underlying propositions, reading times were faster and recall was improved. Manelis and Yekovich explained that since repeated arguments necessarily share a common referent, integration is facilitated by argument repetition. Yekovich and Walker (1978) have also reported that referential overlap facilitates integration and comprehension. However, this doesn't explain how the process operates.

The "given-new strategy" (Haviland & Clark, 1974; Clark & Haviland, 1977) provides one model of how new information is inferred and integrated into memory. When a reader encounters the second of two related facts, the probability of connecting the second to the first (the referent) is increased if the referent is readily available in memory. According to the given-new hypothesis, the integration of new information into pre-existing knowledge structures is a three-step process. First the learner extracts the familiar (given) information from the sentence. The second step is to match the familiar information to an antecedent already present in memory. The third step is to revise memory by attaching the unfamiliar (new) information to the established antecedent.

It has been argued that certain inferences are activated by the first sentence in a sequence, and that subsequent sentences are processed faster if their content is consistent with the inferences triggered by the previous sentence (Haberlandt & Bingham, 1978). Lesgold and Perfetti (1978) argue that the given-new strategy suggests that in the ideal situation, one can start with the most recently understood facts from a discourse and search backwards until an antecedent is found. The further back the antecedent and the less straightforward the mapping, the longer it will take to complete. Hayes-Roth and Thorndyke (1979) found that temporal proximity and similarity of wording affected the integration process. However, Lesgold and Perfetti argue that staging procedures, such as foregrounding, enable the reader to be predisposed to process certain information.

In talking about the psychological process of integrating information, Carpenter and Just (1977) introduced the concept of a "discourse pointer". A discourse pointer is a symbol in the comprehender's mind that indicates the current topic or focus of the discourse. The discourse pointer is like the given-new strategy, in that both are based on the linguistic distinction between presupposition and assertion

(Hornby, 1974). According to Hornby, comprehenders are less likely to notice events inconsistent with a presupposition than events inconsistent with an assertion. This finding can be explained in terms of the discourse pointer. The first sentence sets the discourse pointer on the asserted proposition. The next sentence should exploit the preceding context by acknowledging that the asserted information from the first sentence is now presupposed or known, and the discourse pointer is moved to a new focal topic — that is, the asserted information in the new sentence.

Lesgold and Perfetti (1978) found that when intervening sentences between antecedent and target preserved the theme within which the antecedent had been presented, comprehension of the target was faster than when intervening sentences dealt with a different theme. Gruber, Beardsley and Caramazza (1978) found that a noun that is the subject of the first clause of a sentence is preferred over a noun occurring in the predicate as a referent for a pronoun occurring in a second "because" or "but" clause. Lesold and Perfetti interpreted their results in the context of Anderson's (1976) activation theory. They argue that when intervening sentences address a different theme, the antecedent information currently active deactivates, which makes it unavailable for immediate matching to the given portion of the target sentence. Lesgold, Roth and Curtis (1979) found that when a sentence refers to earlier material in a discourse, it takes longer to understand when the material referred to has been "backgrounded" than when the material has foreground status.

McKoon and Ratcliff (1980) used a priming or "activation" technique to investigate the inferential process. Their technique was based on what they called the "three-component process model" of inference: First, it is assumed that the concept to be inferred has to be accessed in long-term memory; this concept then has to be activated, or brought into short-term or working memory; finally, the information that caused the concept to be activated has to be connected to the concept. The connected structure — the result of these three component processes — is the structure that is stored in long-term memory as a representation of the text. McKoon and Ratcliff required subjects to verify whether a test word occurred in a previously presented text. It was assumed that a decrease in response time to the test word showed that the test word was activated by information that immediately preceded it. Response time for the test word was decreased both when the test word was a referent of an anaphor mentioned in the last sentence of the paragraph, and when the test word was in the same proposition as the referent of the anaphor in the last sentence. They concluded that an anaphor activates both its referent and the concepts in the same proposition as its referent. They also argued that if response times for sentences expressing inferences are equal to response time for sentences appearing explicitly in the text, then it can be inferred that subjects have stored the result in the long-term memory representation of the text.

Further evidence for staging or priming effects was found by Tanenhaus and Seidenberg (1981), who showed that inferences necessary to construct an antecedent for a definite noun phrase are drawn prior to the end of a clause or sentence. Frederiscksen (1981) also found staging effects in pronominal reference. He argues that these results support a "reinstatement theory" in which a set of prior potential referents are reconsidered at the time a pronoun is encountered. Finding the referential relation signalled by the pronoun begins immediately and doesn't stop until further disambiguating semantic constraints become available. This account is at variance with the Kintsch and Van Dijk (1978) model, where they evaluate referential relationships solely on the basis of a stored set of abstract propositions. These findings suggest rather that the internal representation of a sentence is sensitive to topical status as well as to propositional content.

THE ROLE OF PRIOR KNOWLEDGE

It seems, therefore, that macrolevel textual characteristics such as topicality and theme influence the generation of inferences, as well as more lower level and syntactic factors. The levels effect discussed in the previous chapter applies to implied as well as explicit information in the text. Goetz (1979) found that increasing the importance of an inference in a story will increase the probability that the inference will be drawn. Walker and Meyer (1980a) found that information high in the content structure (based on Meyer's analysis scheme — Meyer, 1975) is more likely to be integrated than information low in the content structure. Graesser, Robertson and Anderson (1981) found that "structural centrality" was the best predictor of verification ratings, indicating that the representational structure is an important determinant of the perceived truth of inferences.

A similar but independent distinction to that between "text-based" and "model-based" inferences is that between "necessary" and "elaborative" inferences. Necessary inferences are those that have to be drawn — logically — to preserve the continuity in the text. Elaborative inferences are possible and plausible inferences that are not necessary to establish coherence, but which may be drawn nevertheless. Although a good deal of evidence exists that necessary inferences are made at comprehension (e.g., Haviland & Clark, 1974; Garrod & Sanford, 1977), there is some disagreement as to when elaborative inferences are made. Evidence by Singer (1979, 1980) suggests that elaborative inferences are postponed until they are needed (i.e., at test). However, Garnham (1982) reports an experiment suggesting that such inferences are not deferred but made at comprehension. Granger and Holbrook (1983) provide some evidence that either account is plausible and that people use different inferencing strategies, but that despite these different strategies (e.g., immediate versus deferred inferencing) readers can come to the same interpretation of the text. Many of the studies just discussed have used memory tests to extrapolate to online processing — as Garnham (1982) points out, this relies heavily on the assumption of the encoding specificity principle (Tulving & Thomson, 1973).

In general it has been assumed that all presented information, including the inferences drawn in comprehension, is encoded enough to store in long-term memory. Spiro and Esposito (1981) report an experiment suggesting evidence for an alternative hypothesis: that predictable information, however "central" to the discourse, is taken for granted by readers, processed superficially, and receives an attenuated representation, or no enduring representation at all, and that "accommodative reconstruction" (Spiro, 1977, 1980) takes place at recall. (These suggestions are supported by findings that goal-related information from stories is recalled less well than others types of information, even though it is central to the representation of the text (Stein & Glenn, 1979). It may be that such information is "taken for granted" either at processing or recall — i.e., subjects may not mention it because it is "obvious".)

Several researchers are critical of the "text-based" approach to inference processes in comprehension, arguing that such processes cannot be characterised solely in terms of finding connections between elements in the text (Collins, Brown & Larkin, 1980; Goetz & Armbruster, 1980). Webber (1980), in discussing anaphora, argues that one of the important points to understand about the referent of a definite pronoun is that it is not an element *in* the text, but one suggested *by* it — that is, one of the concepts evoked into the reader's "discourse model".

There is a good deal of evidence for the role played by prior knowledge in generating inferences. For example, Anderson, Reynolds, Schallert and Goetz (1977) found that Jewish subjects drew more inferences reflecting the use of a Passover schema than non-Jewish subjects; Steffenson, Jogdeo and Anderson (1978) found that Indian subjects produced more elaborative inferences for an Indian wedding letter and more distortions for an American wedding letter, while the converse was true for American subjects.

Of course, prior knowledge affects more than just the inferences drawn in comprehension. As was discussed in Chapter 3, the "perspective" taken in reading a text influences the type of material recalled (e.g., Pichert & Anderson, 1977; Anderson & Pichert, 1978; Fass & Schumacher, 1981). Spilich, Vesonder, Chiesi and Voss (1979) had subjects with high or low knowledge of a baseball game recall a segment of a game, and found that high-knowledge subjects recalled more information related to the goal structure of the game than low-knowledge subjects. High-knowledge subjects are also better at recalling the sequence of important events (Voss, Vesonder & Spilich, 1980). Chiesi, Spilich and Voss (1979) showed that, as the number of context sentences increases at input, target sentence recall became greater for high-knowledge subjects than when target sentences were presented by themselves. However, for low-knowledge subjects, as the number of sentences preceding the target increased, recall of target sentences without the context sentences being presented at recall deteriorated.

Chiesi et al. argued that when high-knowledge subjects are presented with a meaningful sequence of domain-related sentences, they integrate the sentence, whereas low-knowledge subjects are not readily able to integrate, and so retrieval suffers. Voss (1984) reports experiments showing that high-knowledge subjects have a more organised structure in their representation of the text, since context sentences induced a "Haviland and Clark effect". The data appeared to support the hypothesis that high-knowledge individuals are more adept at constructing the macrostructure from the text than low-knowledge subjects. It was found that high-knowledge subjects verified macrostructure probes as rapidly as they verified probes more directly related to specific text information, whereas for low-knowledge subjects performance accuracy was lower for macrostructure probes, suggesting a weaker ability to construct the macrostructure.

Despite all the evidence showing the facilitatory effects of prior knowledge on comprehension and recall, other evidence has been somewhat equivocal. Anderson (1981) found that subjects given prior knowledge about individuals learned new information about the individuals faster, but retrieved the information more slowly. Graesser, Hoffman and Clark (1980) found a slight facilitative effect of prior knowledge on reading times, but Graesser, Hauft-Smith, Cohen and Pyles (1980) found that prior knowledge can actually have a detrimental effect on recall. Johnson and Kieras (1983) found that the facilitative effects of prior knowledge can be task dependent. Subjects lacking in familiarity with the content can compensate for it in self-paced reading tasks. Johnson and Kieras found weak effects in a forced-pace task, and the greatest effects in an incidental condition. They suggest that prior knowledge facilitates encoding, or the building of a text representation, rather than providing elaborations, as, for example, Anderson and Reder (1979) suggest.

Spiro (1980) notes that, although the constructive aspects of text comprehension have received a good deal of attention in research, the more reconstructive aspects (cf. Bartlett, 1932) have tended to be neglected. The reconstructive or accommodative hypothesis concerns the attempt to restructure the knowledge representation to accommodate new information. Potts and Peterson (1985) have also recently argued that many models of comprehension treat the representation of the text and the reader's world knowlege as separate bodies of information, as though the representation of the text were "compartmentalised" or isolated from the reader's existing body of world knowledge (presumably with some "leakage" to allow for knowledge effects in comprehension). Other models (e.g., Anderson, 1976) take the other extreme and assume that whenever a person encounters a new piece of information that relates to a concept already existing in their knowledge structures, the new information is directly linked to the existing concept node. Still others choose to take a middle ground (e.g., Collins, Brown & Larkin, 1980; Gentner, 1981). Potts and Peterson (1985) argue, on the basis of their studies, that integration is a matter of degree and can vary depending on the situation as well as individual differences. Potts (1977) found that subjects sometimes fail to use their world knowledge even when it is necessary for the task, and that there were large individual differences in this effect. Chaffin (1979) found evidence suggesting that, although both linguistic and world knowledge are part of the same initial representation of sentences, they are in fact distinct. He found that world knowledge is only used when it is needed, but that the knowledge-based aspects of the initial representation are not attentuated completely.

The effects of prior knowledge, as I have discussed, can be influenced not only by task conditions, but also by instructions given to subjects and by strategies employed by them. Walker and Meyer (1980a) instructed subjects either to learn or to simply read a text, and then tested their ability to relate pairs of facts from the passages studied. They found that while both the learn and read groups were equally good at recognising explicit statements from the text, the learn group was much better at verifying implicit inferences. Meyer and Rice (1982) also found that strategies used by readers influenced the effects of organisational structure and the use of signals for importance in the text. Graesser, Higginbotham, Robertson and Smith (1978) found that in incidental reading of a newspaper text, subjects tended to select material that covered topics they were familiar with, and that the active or narrative aspects were selected more than the static, descriptive ones, whereas in a "task-induced" condition selection of active and static aspects was more or less equal.

Form and Content in Narratives

This section draws together the themes concerning types of text structure, discussed in Chapter 3, with the topic of the previous section, concerning the role of world knowledge. As I have argued in the beginning of Chapter 3, recent psychological research into text comprehension has suffered from confusions over the meaning and use of the notion of structure. One such controversy — the levels effect — was discussed in Chapter 3. The other problem dealt with in this thesis concerns the nature of the relationship between structural aspects (form), and the content expressed in the text. This controversy is particularly apparent in theories of story structure.

The findings of Graesser, Higginbotham, Robertson and Smith (1978) suggest that there is a "narrative bias" in self-selective reading. There has also tended to be a "narrative bias" in research on text structure — that is, researchers have predominantly focussed on the role played by narrative structures in text comprehension.

There are several reasons for this preoccupation with narratives. For one thing, as noted in Chapter 2, there was already a substantial amount of psychological, linguistic, and anthropological research on structure in narrative text (e.g., Bartlett, 1932; Propp, 1968; Lakoff, 1972; Colby, 1973). Furthermore, stories provided readily available and naturally occurring research materials, familiar both to adults and children across many different cultures. They also provided materials that were simple enough to work with, but complex enough to identify important textual and psychological properties. An analysis of the structural regularities in these texts provided a way of investigating the characteristics of the corresponding schema for the text and the way in which it was used in processing. It also offered a link to previous linguistic and psycholinguistic work on sentences. Furthermore, research in knowledge representation provided systems for describing the world knowledge implicit in the text, and research in social psychology on plan and goal structures (e.g., Abelson, 1975) provided a vocabulary for describing the actions of characters in goal-directed situations.

Bartlett's (1932) use of the notion of a "story schema" suggested that it could account for some of the reconstructions found in recall. He suggested that when people cannot remember parts of a story, they can use this schema to reconstruct what might have occurred. The notion also accounts for the fact that over a period of time recall tends to approximate the canonical form or schema more than the actual input.

The forerunners of modern story grammars were provided by Lakoff (1972), who reformulated Propp's (1968) theory of Russian folktales using rewrite rules; Colby (1973), who used rewrite rules in designing a specialised grammar of Eskimo folktales; and Prince (1973). Rumelhart (1975) proposed the first general grammar, designed to apply to a larger set of stories, and other grammars have been proposed by Van Dijk (1975), Thorndyke (1977), Mandler and Johnson (1977), and Stein and

Glenn (1979). The grammars of Mandler and Johnson, and Stein and Glenn are the best developed and most extensively tested in psychological experiments, so they will be taken as the prototypical story grammars.

These grammars are both closely related to that of Rumelhart (1975, 1977). He suggested that stories reflect human problem-solving situations, where there is a focus on the way in which individuals formulate goals and achieve them. There are actually many different types of story, with various functions. Mandler and Johnson (1977), and Stein and Glenn (1979) focus on stories from the "oral tradition", which, it is argued, have certain implications for studying psychological processes involved in text comprehension.

Historically, the story was used to preserve the culture of a given civilisation. Without written records, they had to be passed on by word of mouth, sometimes over several generations. It is argued that the evolution of the structure of these stories must therefore reflect to some extent the organisational structure of the information processing systems through which they passed (Johnson & Mandler, 1980). There are practical reasons for the regularity of such stories: Since oral transmission via several people tends to produce gross distortions (cf. Bartlett, 1932), they must be structured in familiar ways if they are to be understood and remembered as they were presented.

Stories were used not only as historical records, but also as instructional devices, both to explain natural phenomena and to convey and preserve the society's social and moral codes. An additional important function of storytelling was in the reorganisation of personal experience (cf. Labov & Waletsky, 1967). Therefore, the regularities in these stories presumably reflect commonalities in human experience. This function of reorganising personal experience still exists today (cf. Stein & Policastro, 1984), although most people these days tend to regard stories as being primarily for entertainment or enjoyment — and this has been a basis for one of the criticisms of story grammar theories, as will be discussed later on.

According to the systems proposed by Mandler and Johnson (1977), and Stein and Glenn (1979), a simple story consists of a protagonist who is motivated to perform certain types of actions, with the intention of attaining a goal. (However, Johnson and Mandler (1980) also allow for non-goal-based stories in their definitions.) Figure 4.1 gives the outline of the underlying structure of a story, according to the Mandler

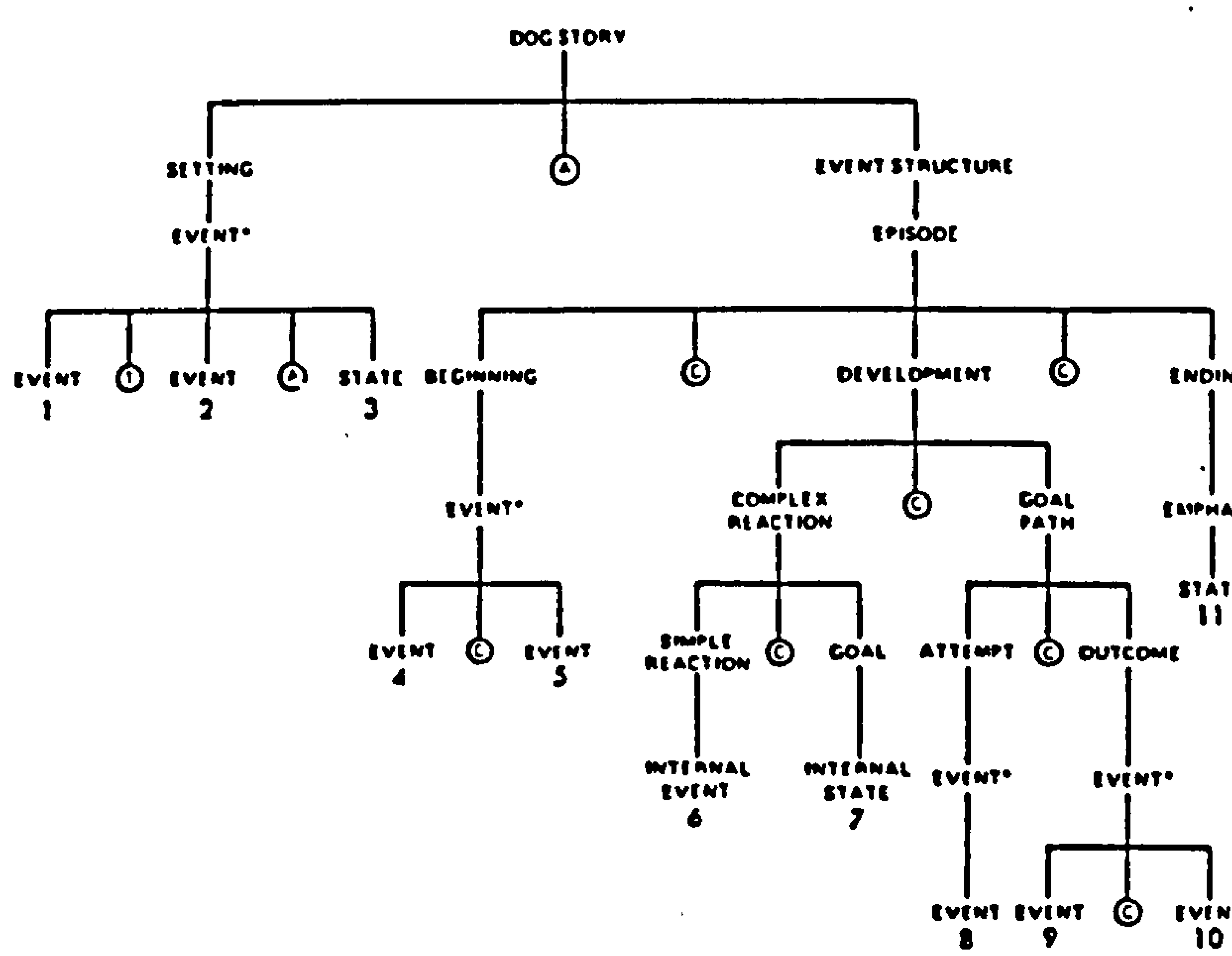

Figure 4.1. This structure represents the *Dog* story used by Mandler and Johnson (1977). The connections *and*, *then*, and *cause* have been abbreviated to A, T, and C, and encircled. The numbers under the terminal nodes refer to the surface statements of the story.

and Johnson (1977) grammar. The essential structure of a single episode story consists of a protagonist who is introduced in the setting, an episode in which something happens, causing the protagonist to respond to it, which in turn brings about some event or state of affairs that ends the episode. Although higher level nodes in such structures are never directly expressed in the text, all terminal nodes represent either a state or event, and typically correspond directly to some surface expression. Nodes are connected by three types of relationship: The *and* relation connects two nodes involving simultaneous activity or temporally overlapping states; the *then* relation connects temporally ordered nodes (both arbitrary and conventional ordering); the *cause* relation connects two nodes when the first is a reason or cause for the second.

Episodes are causally connected by embedding properties of the beginning, outcome, and ending nodes. In the more common ending-embedding episodes, the development of one episode causes an ending that makes up the beginning of a new episode. A beginning-embedded episode occurs when an entire episode forms the beginning of a second episode. That is, a series of events that themselves form an episode constitute the stimulus to the development of a further episode. The third type of causal connection is outcome-embedded. This involves a series of unsuccessful goal paths, where each outcome forces the protagonist to form a subgoal towards a larger goal represented in the higher level episode.

Although the original system proposed by Mandler and Johnson (1977) consisted of little more than rewrite rules, Johnson and Mandler (1980) added the concept of transformations. The beginning, complex reaction, and ending nodes may be omitted or deleted, but these deletions are governed by rules that are dependent on the reader or listener being able to recover the underlying structure. For example, complex reactions may only be omitted if the beginning of it is present. Johnson and Mandler (1980) justify this by arguing that readers or listeners may be familiar enough with the motivations for various types of actions that they can easily infer the nature of the complex reaction from the beginning and the subsequent attempt. Finally, a canonical story need not specify the causal connections between nodes in the surface structure, except that, when propositions are out of order, the causal relations between the moved nodes must be stated.

Although the term "story grammar" has been applied to the type of theory of story comprehension proposed by Mandler and Johnson (1977), and Stein and Glenn (1979), a strong distinction is made between a story *grammar* and a story *schema*. A story grammar, in these theories, is a rule system for describing the structural regularities found in a particular type of text (i.e., stories from the oral tradition). A story schema is a mental structure in the comprehender's head, that incorporates some or all of the regularities and is used in processing the story (Mandler, 1982a, 1984). The story schema guides comprehension so that a coherent representation can be constructed, by generating expectations about what will occur next, and providing a means for checking the plausibility of events as they occur (Stein, 1982). Furthermore, the schema enables the comprehender to reconstruct the story at retrieval.

A good deal of evidence has been provided for the existence of the story schema. Mandler (1982a; Mandler & Goodman, 1982) distinguishes between support for the psychological *reality* of story schemas and their psychological *validity*. Story structures may be said to be psychologically real if people are able to think and talk about them. They are psychologically valid to the extent to which they may be shown to influence processing, regardless of subjective awareness.

There is some evidence that people are able to identify and agree upon "good" and "bad" stories (Mandler, 1982b; Stein & Policastro, 1984), and that their judgements of the important units in the story correspond to the proposed underlying structure (Brown & Smiley, 1977; Pollard-Gott, McCloskey & Todres 1979; Stein & Glenn, 1979; Mandler, 1984). The psychological validity of story structure has even more support. There is evidence for its role in processing both at encoding and retrieval.

Mandler and Johnson (1977) suggested that during encoding the story schema functions by directing attention to the relevant aspects of the input, helping the reader or listener keep track of what has gone before, and indicating, by virtue of its framework, which parts of the story are now complete and can therefore be stored. It has been found that high-level sentences require more processing time, either because they are perceived to be more important, or because they are less redundant with the preceding context. Although there are difficulties with equating a rewrite-rule system with importance, the sense in which story grammar constituents represent a particular level of importance is that the sentences that express the central meaning of each local topic unit are considered to be important to the story. Hence, a story grammar has only two levels of importance: high and low. There is another sense in which topic change is important: It tells the reader that the story is moving forward and that the next unit or category has begun.

Haberlandt (1980), and Haberlandt, Berian and Sandson (1980) found that subjects took longer to read sentences at the boundaries of episodes (cf. also Mandler & Goodman, 1982). It has also been found that people take longer to read a story unit when it is moved out of its normal place, even when it is signalled (Haberlandt, 1980; Mandler & Goodman, 1982). Mandler and Goodman postulated that the most likely reason for slower reading time is that the reader now knows that the previous unit is now finished, and that at this point time may be required to form a macroproposition.

Recall is a function of the role the sentences play in the overall structure of the story, and the extent to which the story matches an idealised schema. However, the overall story may not be clear to the reader or listener until the whole story is processed, at which time some reorganisation may occur. At retrieval, the schema operates in three ways: First, it tells the subject what sort of general information is to be retrieved; second, it provides a temporal sequence to find the specific content; third, if the exact content of a category in the sequence cannot be retrieved, the schema allows the subject to generate an approximation, based on the structure of the schema itself.

Children as young as four years old have shown better recall for important versus unimportant story propositions (Poulsen, Kintsch, Kintsch & Premack, 1979). Subjects are also more likely to recall central material from story categories than elaborations on these units (e.g., Black & Bower, 1979; Mandler & Johnson, 1977; Omanson, 1982; Stein & Glenn, 1979). Additional new material (other than elaborations) in recall tends to conform to the unit being recalled (Mandler, 1978; Mandler & Johnson, 1977; Stein & Glenn, 1979). People also tend to preserve the schematic order in recall (Mandler, 1978; Mandler & DeForest, 1979). Mandler and Goodman (1982) found that it takes longer to recall a target sentence when its cue is from a different unit, and that people make errors in a cued recall test when the target sentence has been presented out of its canonical position. Finally, evidence exists that subjects can reorder a story in accord with the canonical form when it has been presented out of order (Mandler & Johnson, 1977; Rumelhart, 1977; Thorndyke, 1977; Mandler, 1978; Mandler & DeForest, 1979; Stein & Nezworski, 1978; Stein & Glenn, 1979), although if it is presented in a truly randomised order there is some difficulty in recovering the underlying structure (Stein & Nezworski, 1978).

CRITICISMS OF STORY GRAMMARS

Several researchers (especially those in the artificial intelligence tradition) have been critical of the story grammar approach on several grounds, including the adequacy of the grammar, the nature of the proposed structure and how it is used in processing, and the functional account of stories.

Story grammars have been criticised on the basis of their formal properties as grammars (Black & Wilensky, 1979; Black & Bower, 1980; Wilensky, 1983), however, I shall not dwell in any detail on this particular objection here, it is dealt with in Frisch and Perlis (1981). Briefly, the objection hinges on the assumption that the story grammar approach, as I have outlined it (i.e., in terms of the theories of Mandler and Johnson, 1977, and Stein and Glenn, 1979), is a linguistic enterprise, and that these grammars are intended to characterise stories in the same way that sentence grammars are purported to characterise sentences.

Black and Wilensky (1979), and Black and Bower (1980) have argued that story grammars are inadequate because they cannot distinguish between stories and non-stories, that there are instances of texts that have all the necessary features of stories but that would not be classified as a story, and that the grammar would generate some non-stories as stories. The example they choose to illustrate this last claim is the procedural text used in Graesser (1978). However, as Stein and Policastro (1984) note, the problem with the example they use is that it does not include all the parts and relationships contained in a prototypical story (i.e., no protagonist, overt attempt, consequence or reaction). Mandler and Johnson (1980) point out that in order to substantiate their claim that story grammars would incorrectly accept as a story a procedural passage, Black and Wilensky had to transform Graesser's procedures to make it sound more or less like a story. They also point out that Black and Wilensky fail to make the distinction between stories for which their grammar was intended and those for which it was not. They argue that their grammar was intended to apply to those stories from the oral tradition, and that therefore Black and Wilensky should have addressed their arguments to those kinds of stories.

Rumelhart (1980) points out that Black and Wilensky's definition of a grammar has little to do with why a grammar might be psychologically interesting. They seem to hold the view that a grammar is a device for generating all and only the sentences of a language, which presupposes the view that a language is properly defined as a set of sentences, and that a grammar is merely a recursive devise for enumerating them. Rumelhart argues, however, that the psychologically interesting thing about a grammar is that it proposes an analysis of the constituent structure of a linguistic unit.

The main counter to the objection is that it wrongly characterises the function and status of the grammar. Story grammar theories, at least those of the sort under consideration here (i.e., not the text grammars of earlier linguistic theory) are not intended to be linguistic endeavours, but descriptions of both story structure and the schematic knowledge possessed by readers of such stories (cf. Mandler, 1984).

However, even though one might dismiss the objections to the formal adequacy of story grammars, a more fundamental problem with these theories concerns the ways in which schematic form is confounded with semantic content. In fact, according to Morgan and Sellner (1980), story grammar theories have confused three distinct properties: linguistic form, content, and presentational structure. Of course, these properties are not altogether independent in any text. The content of a text will determine several matters of presentational structure (e.g., ordering) and linguistic form (e.g., signalling devices, devices for cohesion, etc.). Morgan and Sellner (1980) point out that the distinction between syntactic and semantic rules in standard grammars hinges on the fact that the categories of syntax are categories of form (although some grammarians believe the distinction between syntax and semantics to be somewhat more blurred — e.g., Fillmore, 1968). They criticise Rumelhart's system because it concerns relational terms rather than categories of form. Johnson and Mandler (1980) are quite explicit about the fact that their structures involve relationships between story categories, and in this sense they are more dependent on context for their assignment than are sentential categories. This underlies another objection of Black and Wilensky (1979): With such systems, in order to apply a syntactic rule to a pair of sentences, the reader must first determine the semantic relationship between them. In other words, they object that before a rule can be applied, the reader must have already accomplished the purpose for which the rule was designed.

The confounding of form and content is even more fundamental: Not only are the categories of story structure relational ones, but these categories are relational because they involve statements about events that are causally and/or temporally ordered. In other words, story grammars confuse the distinction between "text grammars" — i.e., representatons of configurations of textual elements, defined independently of content — and "content schemas" — i.e., representations of the organisation of possible content facts (Kieras, 1985).

The fact that story structure appears to be dependent on the causal/temporal chaining of events and actions has led some researchers to suggest that they are unnecessary constructs, and that story comprehension can be understood in terms of world knowledge structures concerning goals, events and actions, such as scripts (Schank & Abelson, 1977; Bellezza & Bower, 1982), knowledge of typical plots (Lehnert, 1981), typical actions of main characters (Omanson, 1982; Weaver & Dickinson, 1982), plans (Bruce & Newman, 1978; Collins, Brown & Larkin, 1980), and so on. Several of these alternative theories have proposed what has been called the "primary causal path" hypothesis (cf. Voss & Bisanz, 1985). This is the hypothesis that the crucial components of a story are the narrative units that make up the sequence of goal-directed, causally-related events that lead to the resolution of the plot (i.e., the achievement of the protagonist's goal). Various terms have been used to characterise this causal path: "critical path" (Black & Bower, 1980), "central content" (Omanson, 1982), "causal chain" (Schank & Abelson, 1977), although there are differences in the way the causal path is analysed in these different theories. Some of the theories involve fairly linear chains (e.g., Black & Bower, 1980; Omanson, 1982), while others assume causal networks (e.g., Trabasso, Secco & Van Den Broek, 1984; Trabasso & Van Den Broek, 1984). However, all of them assume that findings such as the "levels effect" are due to causal connectivity, rather than to level in a hypothesised story structure.

There is some evidence for the causal path theories. Causally connected passages tend to be read faster than those that are not causally connected (Haberlandt & Bingham, 1978). Black and Bower (1980) showed that causal connections, especially those on the main causal chain, produce better memory than other cohesive devices not involving causal connections (e.g., argument repetition). Black and Wilensky (1979) used Schank's (1975) causal chain theory to test the hypothesis that the closer a statement is to the main causal chain of a story, the better it will be remembered. They found a high correlation between rankings of story statements and distance off the main causal chain. They also found in a regressional analysis that the causal chain was improved in predicting recall by being combined with some hierarchy variable, therefore, they concluded, the best theory should be one with a hierarchical causal chain. Lehnert, Black and Reiser (1981) found that causal connectivity of an event or goal is a good predictor of its likelihood of being included in a summary of a story.

Several researchers have proposed a taxonomy of possible causal connections between the actions and states described in stories. Omanson (1982) presents an analysis that classifies the content of stories as central, supportive, or distracting. In Omanson's view the story can be divided into three types of function: characterisation of the world in which the story takes place; identification of the characters; and depiction of the events involving the characters that take place in the described world (focal events). In this analysis the text consists of content units (events or states) which are classified as identifying, characterising or focal. The relations comprising event sequences are identified as being either componential, purposeful, causal, disruptive, or enabling. Identifying units that introduce main characters are classified as central, as is the final focal content unit of the narrative. If the unit identified as central is the purpose of any other unit, these units are also classified as central, as are any units that enable, cause or are disrupted by the central units. Central content, therefore, is that which is part of the purposeful-causal sequence of events that lead to the end of the narrative. (Disruptive relations are viewed as purposeful relations that fail, and enabling relations as necessary preconditions for causal relations.)

Kemper (1982) found support for this kind of causal taxonomy — based on the analysis of Omanson, Warren and Trabasso (1978), and Warren, Nicholas and Trabasso (1979) — since, when given an opportunity to improve stories by supplying missing links, readers inserted new information into the gaps, repairing violations of the event chain taxonomy. She concluded that readers can recognise the connectivity of the intact story events since few unnecessary insertions were made. However, causal chains such as Omanson's are more or less linear. Kemper (1982) also found that hierarchy of superordinate and subordinate actions and events interacted with causal connectivity — that is, more superordinate actions were restored than were subordinate ones. The location of actions in the story was also an important determinant. Thus, it seems that a hierarchical structure is operating.

The analyses of Trabasso, Secco and Van Den Broek (1984), and Trabasso and Van Den Broek (1984) assume a causal network of events. The relations between events are analysed using logical criteria of necessity and sufficiency, then they are represented in a causal network with events as nodes, and inferences as arcs. A causal chain of the most important events in the story is then found, using criteria for opening, continuing, and closing the chain. Causal cohesion for the story is quantified in terms of the percentage of events in the story that are contained in the causal chain.

It is assumed that the causal chain is opened by the setting statements of the story since these are necessary to provide the background conditions in which subsequent events take place. The chain is ended by criteria concerning expectations with respect to the protagonists goals and inferred plans, such as whether the protagonist succeeded in achieving his goal. In the event of failure the chain is ended. Trabasso et al. (1984; Trabasso & Van Den Broek, 1984) compared the results of their analyses with data collected by Stein and Glenn (1979). Defining coherence as the proportion of events that lie on the causal chain, they found the coherence of the stories to be linearly related to the proportion of events recalled. They also found that events on the causal chain were recalled more often, and were not forgotten over a one-week delay; events that were not on the causal chain (dead-end events) were recalled half as well and underwent further loss over time. The number of connections of events was also correlated with recall for both causal chain and for dead-end events. The order in which categorised events were recalled was highly correlated with the proportion of those events that were in the causal chain for that category. They also found that the order of importance assigned to events was correlated with the proportion of those categorised events that lay on the causal chain.

Trabasso and Van Den Broek (1984) extended the validation of their analysis by comparing it with the data obtained by Omanson (1982). They predicted that if an event is integrated into a causal chain, it should be retrieved with a high probability, since it is linked in both forward and backward directions to other events. They found that the events that were in the causal chain were recalled higher, proportionately, in both immediate and delayed recall, and showed less of a reduction over time than the dead-end events. Causal chain events also had a higher probability of being summarised than dead-end events, and were rated as more important. The predictions of the causal chain were also very close to Omanson's divisions of central and peripheral units.

Mandler (1982a) suggests that Omanson's (1982) centrality analysis is compatible with and complementary to story grammars to a large extent, since both analyses assume that readers encode narrative content in terms of events and states, and since a major aspect of story understanding involves the use of knowledge about social actions to explain the actions of main characters. She suggests that Omanson's analysis might supplement story grammars in enabling one to determine which sentences express the gist (centrality) of the story category. Trabasso and Van Den Broek (1984) also regard their analysis as somewhat compatible with a story grammar analysis, but they see it more as a link between that approach and the approach of Schank, Black, Bower, Wilensky, and others, that views story comprehension as a process of understanding events, states, and actions of the characters in the story.

A corollary of the hypothesis that it is world knowledge that guides story comprehension is that similar findings concerning the structure of action and event descriptions should also be found in non-linguistic contexts. Lichtenstein and Brewer (1980) suggest that some components of the story schema are not linguistic or story specific, but rather rely on more generalised event understanding schemas that are also used in nonlinguistic contexts. They found that memory for prose descriptions of events were very similar to memory for videotapes of events. Superordinate goal-directed units were recalled at a higher rate than other units, and goal-directed preceding units were recalled with greater frequency than subsequent units. They also explored the effects of moving some event components out of canonical order, and found that when the critical event was described in its canonical position it was always correctly recalled in that position, but when it was in a noncanonical position, it was three times as likely to be recalled in a wrong location as in its original location. Furthermore, for the majority of the incorrectly placed units, the action units were recalled in their canonical position. Lichtenstein and Brewer admit that for such things as settings or morals in stories subjects may rely on story schemas in recall, but argue that for much of the content of the story, such as sections describing the behaviour of characters, it is not the knowledge of literary conventions that subjects use, but knowledge of the structure of naturally occurring sequences of events.

However, Mandler (1982a) argues that analysis of social actions by themselves does not account for knowledge of how these sequences are typically structured within stories. She points out that what is at issue is not that content knowledge is important, but whether there is an organisation to any particular kind of schema irrespective of the content involved (Mandler, 1984). There is an important distinction to be made between an action discourse in general and a story (or narrative discourse) in particular. Action discourses are coherent if they denote connected action and event sequences. Since, however, discourses need not be fully explicit, they only represent part of the actions and events of this sequence, and require inference in order to fill the missing links. Much of the work in artificial intelligence is about action discourses rather than stories. However, not all action discourses are stories, and stories are not just action discourses, but represent events and actions that may interfere with normal or expected courses of events. Stories have a number of semantic and pragmatic constraints distinguishing them from other action discourses. Furthermore, for a given culture these constraints may become conventionalised. Since stories are action discourses, they must also be organised on the level of adequate action descriptions, but where such factors as "completeness", style, ordering, perspective, etc., are also relevant. Both kinds of account are needed for a complete theory: that is, a system of narrative categories and rules, together with a specification of the structural action content of the narrative categories as represented by the story (cf. Van Dijk & Kintsch, 1983; Mandler, 1984).

Other criticisms of the story grammar approach have been made by Garnham, Oakhill and Johnson-Laird (1982), and Johnson-Laird (1983), who argue that any text can be understood through the joint operations of referential coherence and plausibility, with the implication that no purpose can be served by knowledge of an underlying structure. Johnson-Laird (1983) distinguishes between coherence and plausibility, since a discourse may be coherent, yet describe a nonsensical sequence of events. Coherence (or the construction of a "mental model" of the discourse) depends on both coreference and consistency. Plausibility depends on the ability to interpret the discourse in an appropriate temporal, spatial, causal, and intentional framework. However, Mandler (1984) points out that the factor of plausibility, which is content-based, is orthogonal to issues of structure. She also notes that although referential coherence is influenced by knowledge of story structure (since people can use a story schema to fill in the gaps in imperfectly connected sentences), it is not a structural characteristic of texts.

Yet another set of criticisms of story grammar theory involve the fact that these theories ignore the affective aspects of stories. It is argued that stories have interest or entertainment value, and that an affective response must be experienced by the comprehender in order for a text to be a story (e.g., Morgan & Sellner, 1980; Brewer & Lichtenstein, 1981; Wilensky, 1983). However, both Mandler and Johnson (1977), and Stein and Glenn (1979) state that, although the explicit emotional response of the protagonist doesn't have to be included in the surface structure of the text, the comprehender has to be able to infer the specific emotional response of the protagonist. Furthermore, the emotional response must be directly (causally) related to the prior initiating event as well as to the subsequent goal of the protagonist.

In summary, there is a good deal of evidence that content-specific knowledge of goals, events, and actions plays an important role in story comprehension. However, none of the theorists of the "story grammar" school have denied the importance of this type of knowledge. What they maintain is that in addition, knowledge of text-specific conventions — story structure — is also used to guide comprehension. Van Dijk and Kintsch (1983) argue that specific text representations are needed because discourses express content in a specific linguistic way, and it is often necessary to use surface structures in processing the text. Discourses also have a particular style and may possess certain rhetorical devices or conventional structures. Furthermore, the facts described in two different texts may be the same, but expressed from different points of view.

The problem with using stories in studying the role of text structure in comprehension is that they are so closely tied to actions and goals that it is very difficult to unconfound text and action structures. As Van Dijk and Kintsch (1983) point out, a convincing refutation of the cognitive relevance of schematic structures in texts must include other text types than narratives. For example, although expository texts may also contain causal chains, in narratives the reader already possesses the knowledge necessary for following the links in the chain, whereas instructive or argumentative texts are designed so as to provide the reader with *new* knowledge structures. In the next chapter, a study is described that looks at the role of schematic structure in a different text type — the scientific research report — in an effort to avoid the entanglement of form and content inherent in narratives.

CHAPTER 5

STRUCTURAL FACTORS IN THE COMPREHENSION OF SCIENTIFIC TEXT

As the previous chapter has shown, the problem with using narratives in exploring text structure is that it is difficult to separate factors to do with readers' knowledge of the conventional structure of the text from their more general world knowledge — that is, knowledge of events, actions and goals. It has been suggested that investigations of other text types are needed to demonstrate the role of the reader's knowledge of the text structure. This is the purpose of the study described in this chapter.

A Study of the Role of Structure in Comprehension and Recall of Scientific Research Reports

This study involves an investigation of the effects of knowledge of the text type, knowledge of the topic (or content), and instructions given to subjects, on their comprehension and recall of the text. The text type chosen for the study was a scientific text: a psychology research report. This particular text type was chosen because of its highly conventionalised structure. Although the structure may vary slightly according to the particular field of research, or the rules of a particular journal, the overall global structure has a certain conventional and invariant form that distinguishes it from other types of scientific and technical texts.

The psychology research report was the subject of an analysis presented by Kintsch and Van Dijk (1978) to illustrate the operation of their text processing model. In their model Kintsch and Van Dijk also included as an important feature the role played by what they call the "superstructure" of the text: the global organisation of the text topic(s) according to the particular text type. Whereas macrostructures organise and reduce the semantic structure of the text, the relative relevance of the information in the text is determined by the superstructure. The Kintsch and Van Dijk model is a start towards providing a more general account of structural factors than has been provided by the story grammar approach, however, although the original account of the model concerned research reports, no attempt was made at that stage to investigate the role of the superstructure of this particular text type. This study, therefore, is also an attempt to add to this particular model by providing evidence for the role of the superstructure of this particular text type. This model, and its recent revisions, were discussed in some detail in Chapter 3. However, little attention was paid there to the proposed operation of the textual schema in processing:

The superstructural schema is postulated to control the operations of macrorules in both comprehension and recall. The reconstruction or recall process is modeled with three reproduction operators consisting of the inverse application of the macro-operators, and resulting in the reconstruction of some of the information deleted from the macrostructure, under the control of the textual schema. Reconstruction includes the addition of plausible details and properties, particularisation or specifications of conditions, components, or consequences of events. Reconstruction takes place under the control of the textual schema. The micropropositional network is matched against the schema and slots are instantiated as in the following example, where the superstructure is a psychology research report:

The network is searched in a top-down depth-first manner, for propositions that match the *introduction* schema. Once *n* slots have been filled, this instantiated schema drives the search for the next pattern. The search continues through the network for propositions to fill slots in the method and results schemas. The propositional fillers for the *introduction* schema are then compared with the *results* schema, and inferences are made to fill the slots in the *conclusions* schema. These are then matched against the remaining portion of the network, and when a match is found, the process stops. Those propositions that have been successfully matched are then given higher probability weightings. The combined probabilities from the initial processing cycle and the schema match are used to predict the recall probabilities for each proposition. With probability weightings determining the structure of the hierarchy, the resultant network then drives the production part of the comprehension process, which is controlled by the inverse application of the reproduction operators.

Van Dijk (1977b) notes the importance of superstructure in argumentative texts: "Without the typological categories determining a macrostructure, we would perhaps have information concerning what the discourse is about, but not what (macro-)propositions count as premises, and which as conclusions" (Van Dijk, 1977b, p. 139). Several researchers have identified problem-solution structures in both narrative texts (e.g., Rumelhart, 1977) and, more commonly, in technical and scientific texts (e.g., Meyer, 1975, 1985; Jordan, 1980; Hoey, 1983). The structure of argumentation has been studied for a long time in more or less precise terms in the classical Aristotelian tradition. Aristotle distinguished between poetics as the rationale of imaginative discourse, and rhetoric as the rationale of practical argumentation. The schemas for admissible reasoning in syllogisms are well-known. However, everyday argumentation seldom follows the acceptable form of reasoning studied by logicians (see the studies by Wason and Johnson-Laird for psychological evidence of this — e.g., Johnson-Laird & Wason, 1977; Johnson-Laird, 1983).

The structure of practical or applied, as opposed to formal argumentation was emphasised especially by Toulmin (1958). Toulmin's argument structures may be seen as more general schematic structures upon which other more specialised structures (e.g., scientific reports) are built. He distinguishes between the *claim* or conclusions, whose merits the argument is meant to establish, and the facts appealed to as the foundation for the claim — the *data*. He also recognises many other components of arguments, including the means of strengthening a claim by showing that the step from claim to data is appropriate and legitimate — the *warrant*. This is the authority relied upon for taking the step from one to the other. Toulmin also notes that there may be a need for explicit reference to the degree of force which the data confer on the claim, by virtue of the warrant. This is a *qualifier*, or a condition for *rebuttal*. Warrants may also need to be supported by a *backing*, to explain why in general the warrant should be accepted with some authority. The schematic form of an argument in Toulmin's system is shown in Figure 5.1a. Figure 51b shows an example of a simple argument (taken from Toulmin, 1958). Toulmin's structure has been used by several researchers as the basis for analysing argumentative discourse (e.g., Cohen, 1980) and text (e.g., Van

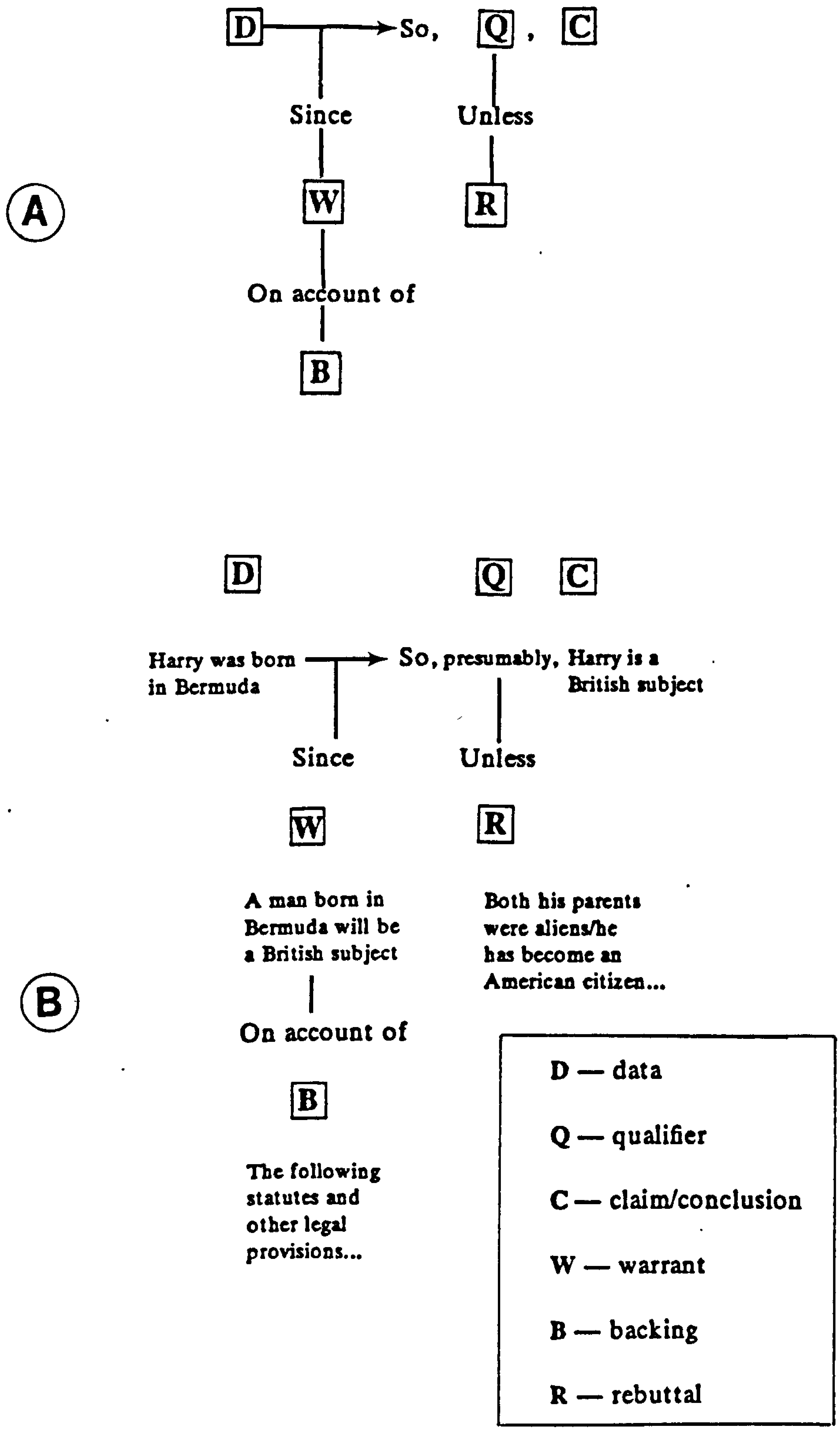

Figure 5.1. Toulmin's argument schema (Toulmin, 1958).

Dijk, 1980a).

As well as distinguishing between poetics and rhetoric, Aristotle also distinguished between rhetoric as everyday argumentation before an audience of ordinary people, and the specialised sciences, which argued before specially qualified audiences (cf. Halloran, 1978). The structure of a scientific article (for example, a psychology research report) is a subset of the genre of argumentative text types. That is, there is a general organisation of the material according to various premises and conclusions (or evidence and claims), which becomes specialised in the case of a particular subset of these texts (the research report). Scientific papers are conventionalised because of the constraints of methodology upon the reported experiments themselves. The argument schema is often embedded in several of the categories of the schema.

The general schema for a psychology research report is shown in Figure 5.2, taken from Van Dijk (1980a). The top-level categories in this schema consist of statements concerning the problem under investigation, together with an account of the solution to this problem. This can be seen as a more specific case of the general argument schema consisting of claims and evidence for these claims, or conversely, as sets of premises and conclusions.

The statement of the problem and its formulation into more specific premises and conclusions consists of some setting or background information, from which are drawn certain hypotheses, which in turn generate more specific predictions. The setting information may contain a general statement of the problem, and some information concerning previous research findings relevant to the problem. There then follows a more specific statement of the major hypotheses under investigation, and the specific predictions are then made. The solution category consists of two other major categories, the description of the experiment and its results, and an evaluation of the experiment in the light of the predictions and hypotheses set forth at the beginning. This section can be seen as providing facts, warrants and backing for the claims made in the problem section. The first major subnodes consist of a general description of the design, and an account of the actual procedure and the results. The design section includes information about subjects and materials used, together with a general description of the experimental design (i.e., a summary of

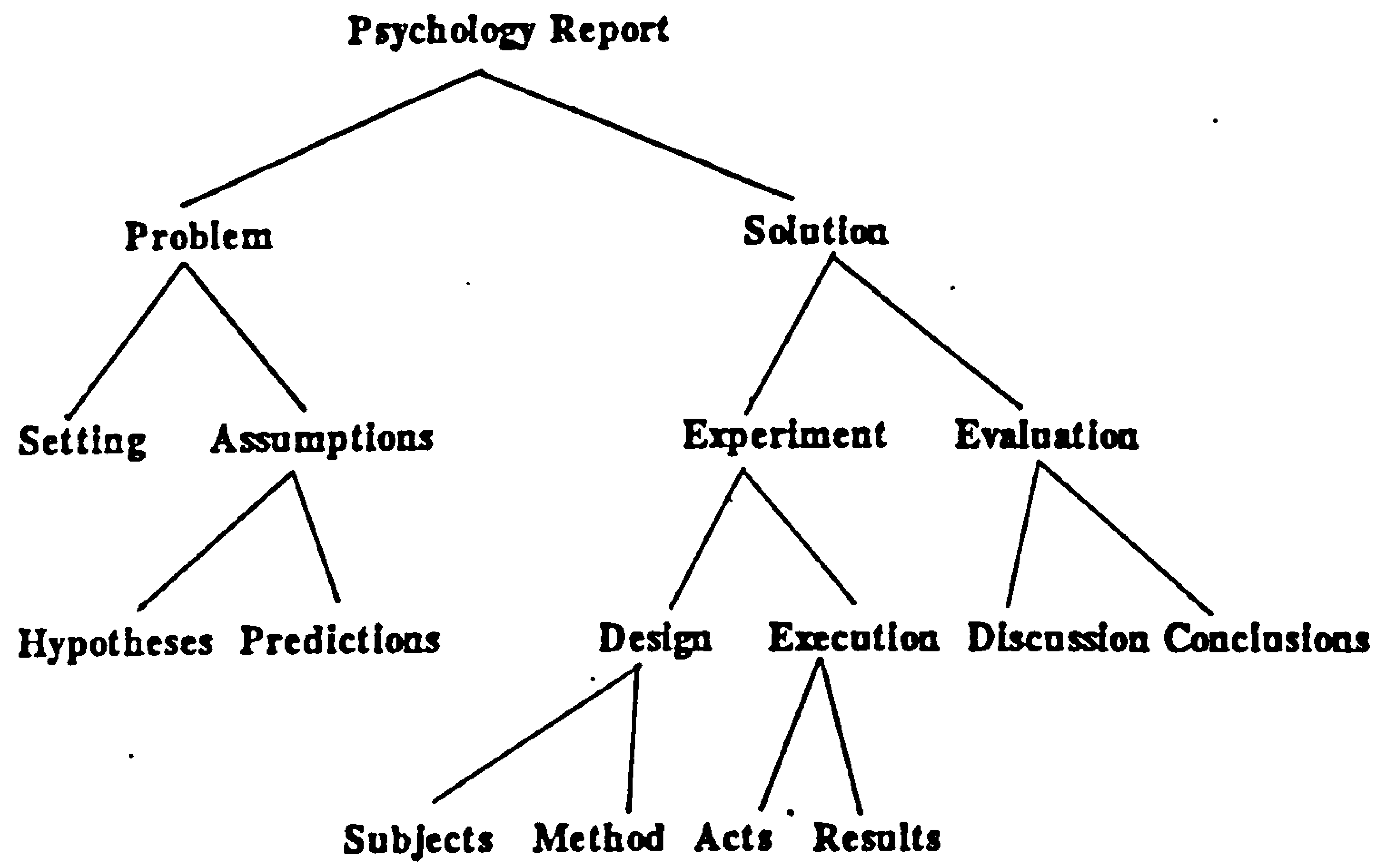

Figure 5.2. Superstructure of a psychology research report (from Van Dijk, 1980a).

the conditions and type of design). The description of the execution of the experiment consists of more detailed accounts of instructions given to subjects, and what they were required to do, followed by a description of the results of the experimental manipulation. The evaluation of the experiment involves drawing conclusions from the experiment and relating them to the hypotheses and predictions made in the first part of the report. This involves a more general discussion of the results, perhaps together with alternative accounts of the data, and any previous research to which the results are related. Finally, the major conclusions are stated, usually with specific reference to the original problem.

In summary, the reasons for choosing to study this particular text type are that firstly, it does not confound form with content to the same extent that stories do, as was seen in the previous chapter. Whereas readers need knowledge of events, goals and actions in order to use the story structure in comprehension, readers of this text type do not need to rely so much on general knowledge of the content in order to use the structure in processing. However, there is some relationship between the categories of the report schema. The problem and solution (or claims and evidence) are logically related. Therefore readers do need some knowledge of how to reason in order to use the structure, but, it is proposed, they do not need to know the specific content — in fact, it is the very purpose of the schematic structure to inform readers of the content by relating the various claims, warrants, backing and so forth. In other words, although the report structure is a schematic one, with functional and related categories, it is not bound to the specific content in the way that stories categories are — i.e., by virtue of the relationship between the events described.

Secondly, it was hoped to provide some validation of the claim made by Kintsch and Van Dijk (1978; Van Dijk & Kintsch, 1983) that readers are able to use the report schema to facilitate comprehension and retrieval. Thirdly, since scientific reports are familiar only to a subset of the population, and not as culturally pervasive as stories, it was hoped to subsantiate arguments for the role played by knowledge of the text structure by providing contrastive evidence between individuals familiar with the text type, and those unfamiliar with it — something that is difficult to demonstrate with stories, since even children as young as four years old have been shown to use a story schema (Poulsen, Kintsch, Kintsch & Premack, 1979).

This study was conducted in particular, therefore, in order to investigate the role that familiarity with the superstructure of the text plays in comprehension and recall. The superstructural formats of texts differ because texts have specific communicative purposes. It is hypothesised that the reader uses these conventional formats in comprehending the text, and in generating expectations about the content of the text, which then facilitates the construction of a representation of the text in memory.

Contrastive studies in text comprehension between high-knowledge and low-knowledge subjects have also been performed by, for example, Spilich, Vesonder, Chiesi and Voss (1979) and Voss, Vesonder and Spilich (1980), to demonstrate the role of prior knowledge of the content of the text. Vesonder (1979; cited in Voss & Bisanz, 1985) used a similar method to examine familiarity with the structure of a scientific report. He presented science passages to college students who were either science majors or non-science majors. His results suggested that science majors had a better knowledge of the prototypical structure of experimental reports, as indicated by their pattern of performance. However, Voss and Bisanz (1985) note that the passages used in this study also had scientific content — a possible confounding factor — i.e., confounding knowledge of prototypical text structure with domain-specific knowledge. Therefore, in this experiment two different texts were used, in order to avoid the possibility that a psychological research report might not only contain an unfamiliar structure to the low-knowledge subjects, but that it also might be confounded by unfamiliarity with the topic, in that it might contain technical terms. The two texts used were a report of a visual discrimination experiment, containing several technical terms, and a social psychology research report, which contained few technical terms, and was readable by people without a psychology background.

In order to establish further whether or not subjects could use the superstructural schema in comprehending and recalling the text, two versions of each text were used. One version preserved the original structure, while in the other version the structure was disturbed by reordering groups of sentences representing certain categories of the schema. Several studies have shown that subjects are able to recall scrambled stories in the correct canonical order (e.g., Mandler & Johnson, 1977; Thorndyke, 1977; Mandler, 1978; Stein & Nezworski, 1978; Mandler & DeForest, 1979; Stein & Glenn, 1979). It was hypothesised that readers familiar with the report structure would be able to use it to reorganise the text if it were presented in an unstructured way, but that readers unfamiliar with the structure would not be able to do so, and that their recall should more or less preserve whatever input order they were given.

In order to further substantiate this, it was hypothesised that if low-knowledge subjects were actually instructed to use a report schema, it should ameliorate the effects of restructuring the text. Therefore, subjects were either given instructions concerning the report schema before reading the text, or after reading but before recall, or not at all. Several researchers have suggested that the superstructure of the text is used as an encoding stratgey to guide comprehension of the text (Mandler & Johnson, 1977; Olson, Mack & Duffy, 1981; Van Dijk & Kintsch, 1983). In this sense, the knowledge or schema of the report structure acts as an "advance organiser" (cf. Ausubel, 1960). However, the schema may also be used at retrieval, to reorganise the text. Many of the previous studies have not been able to distinguish between these two possiblities, since they have relied simply on recall measures. Therefore, it is hoped that by manipulating when the instructions for structuring the text are given (i.e., before or after reading), these two possibilities may be distinguished.

Predictions for Reading Times

It was hypothesised that if readers are using the report schema to encode the text, there should be an increase in reading times at category boundaries. The rationale behind this is that an increased reading time at category boundaries indicates an extra cognitive processing load. This is purportedly due to the fact that subjects are, at that point, either initialising a new schema, or slot in their schema, as they form hypotheses about what is coming next (the prediction hypothesis); or that subjects are consolidating and reorganising the chunk that they have just read, and forming macropropositions from the micropropositions in the text (the postdiction hypothesis — cf. Kintsch, 1980). Specifically, it was predicted that reading times for the first sentence in a new category should be longer than the rest, since it may not be apparent when the end of the category is reached until the beginning of the next one.

Several researchers have demonstrated this effect with stories. Haberlandt (1980) found that reading time at boundaries of episodes (i.e., beginnings and endings) were higher than reading times of the intermediate consitituents of the episode (goals and attempts). The explanation was that subjects use episodic boundaries in encoding the propositions of a story. When the same sentences were presented in a nonnarrative form, these differences disappeared, indicating that the findings could not be attributed solely to the specific sentences (cf. also Mandler & Goodman, 1982).

If readers are familiar with the structure of the text, they should find it easier to comprehend. Therefore, overall reading times should be shorter than for readers unfamiliar with the structure of the text. Johnson and Kieras (1983) argue that familiarity with the material leads to faster processing, since there is some saving in the process of building a representation of the text if the necessary knowledge structures already exist. However, the texts used in their study were well-organised.

Kintsch, Mandel and Kozminsky (1977) found that even though subjects were equally able to produce well-organised summaries from reading scrambled texts, their reading times were longer than those given well-organised texts. It has also been found that people take longer to read a story unit when it has been moved out of its normal place, even when it is signalled (Haberlandt, 1980; Mandler & Goodman, 1982). Kintsch et al. (1977) suggested that texts that have a structure unfamiliar to readers would be even more sensitive to the effects of scrambling. This might be the case where scrambling leads to disruption of local coherence. However, if local coherence is maintained, and only structure is manipulated, one might predict that readers using a structure strategy to encode an unstructured text — that is, readers familiar with the normal structure, would show longer reading times, since the actual order violates their expectations, and since it requires more effort to reorder the text into the canonical form. Readers not familiar with the canonical structure would in any case be reading in a much more bottom-up fashion and, therefore, as long as local coherence is maintained, may be faster in processing the text. (However, this may only show up with the social psychology text, since the extra difficulty due to processing unfamiliar content in the perception text may swamp the effect.)

One might also argue that if high-knowledge subjects, or low-knowledge subjects given instructions for using the structure prior to reading, show longer reading times, it may not be due to using extra processing effort to reorganise the unstructured text, but simply because they find it harder to understand — even with local coherence maintained. If, however, the effect is at category boundaries (presumably where the reorganisation takes place) rather than within categories, then we might assume that the effect is due to the representation-building process (cf. Johnson & Kieras, 1983) rather than to overall comprehension difficulty. Finally, the effect should be less for high-knowledge subjects than for low-knowledge subjects with instructions, since the former already have a structure, and don't have to build it afresh.

Low-knowledge subjects should find the perception text harder to comprehend in general, since it contains several technical terms. Therefore, their reading times should be slower than with the social psychology text, which contains few technical terms. There should not be as much difference, if any, between the two texts for the high-knowledge subjects. However, if low-knowledge subjects are given instructions for using a schema to organise the text, it should make the perception text somewhat easier to understand, at least for the structured version. Therefore, the manipulation of instructions should ameliorate the effects of unfamiliar content.

Predictions for Recall

Any predictions concerning quantity of recall in a self-paced reading task have to be made with care, since there is a general trade-off between reading and recall (cf. Vipond, 1980). Johnson and Kieras (1983) found weak effects of familiarity on recall in a self-paced reading task, and suggested that subjects were compensating for lack of familiarity at reading, by giving the text extra processing. The effects on recall were greater when the task was forced-pace or incidental.

However, one can make some predictions about qualitative as opposed to quantitative effects on recall. Several studies have shown that subjects are able to reorganise stories at recall when they have been presented out of canonical order (Mandler & Johnson, 1977; Rumelhart, 1977; Thorndyke, 1977; Mandler, 1978; Stein & Nezworski, 1978; Mandler & DeForest, 1979; Stein & Glenn, 1979). Therefore, it was predicted that high-knowledge subjects should show a greater tendency to reorganise the unstructured text to conform to the canonical order than low-knowledge subjects, who should be more likely to stick to the original input order. However, if low-knowledge subjects are given instructions for structuring the text before recall (but after reading) then their recall order should follow more that of high-knowledge subjects. If it does not, then one might attribute this to the fact that the schema is being used as an encoding strategy, and, consequently, low-knowledge subjects given instructions before reading the text should show the expected reordering.

DESIGN

Method

In summary, the investigation involved the effects of manipulating the superstructure of the text (i.e., the ordering of superstructure categories), the familiarity of subjects with the text structure, the difficulty of text content, and whether or not instructions were given to subjects before or after reading, on their comprehension and recall.

Subjects

Seventy undergraduates took part in the experiment. They were each paid one pound for participating. Thirty-nine of the subjects were psychology undergraduates (the "high-knowledge" group), and thirty-one were non-science undergraduates (the "low-knowledge" group), none of whom had prior experience with this type of text, nor with any formal training in psychology. The high-knowledge group consisted of twenty-seven females and twelve males; the low-knowledge group consisted of sixteen females and fifteen males.[2]

Materials

In this experiment, naturally occurring texts were used. One was a social psychology research report, and the other was a report of a visual discrimination experiment. The social psychology text was chosen because it was felt to be readable by those without a psychology background. The text (herein referred to as text1) was taken from the *Journal of Social Psychology* (Goldman, Florez & Fuller, 1981). The original text was altered as little as possible, but such things as quantitative data were replaced by verbal accounts of the results, and the text was shortened to about 1100 words. (See Appendix A for original texts, and Appendix B for resulting materials that were used.) The visual discrimination text was taken from the *Journal of Experimental Psychology: Human Perception and Performance* (Taylor, 1982). This text was felt to be more difficult for low-knowledge subjects. Again the text was altered to eliminate quantitative material, and was shortened to about 1100 words.

---

2. Unequal cell sizes are a result of dropping subjects who did not complete all measures - see Table 5.3.

The differences in terms of difficulty between the two texts is confirmed by the

Table 5.1

Readability analysis			
Variable		Text1	Text2
Readability	Kincaid	12.1	15.5
	Automated	11.9	16.5
	Cole-Liau	13.4	15.8
	Flesch	14.6	14.9
Sentence info.	av. sent. length	17.2	23.2
	av. word length	5.25	5.59
	av. nonfun length	6.91	7.2
	short sentences	22%	25%
	long sentences	8%	12%
Sentence types	simple	41%	35%
	complex	55%	55%
	compound	2%	4%
	compound-comp	3%	6%
Verb types	ToBe	28%	33%
	Auxiliary	33%	15%
	Infinitives	16%	24%
	Passives	16%	20%
Word usage	Prepositions	8.5%	16%
	Conjunction	2%	2%
	Adverbs	5%	2%
	Nouns	23.8%	27.6%
	Adjectives	16.1%	24.5%
	Pronouns	3.8%	2.2%
	Nominalisations	1%	3%
Sentence openers	Prepositions	13%	16%
	Adverbs	11%	4%
	Subject	69%	63%
	Verbs	0%	0%
	Sub conj.	6%	12%
	Conjunctives	0%	0%
	Expletives	2%	6%

readability indices shown in Table 5.1. The social psychology text (text1) has a lower readability index on all four measures used in the STYLE program (cf. Cherry, 1982). The Flesch readability index (Flesch, 1948) may be taken as the more accurate of these measures (cf. Coke & Koether, 1983). It indicates that the social psychology text is readable by those of at least first-year undergraduate level, whereas the difficulty of the perception text (text2) requires a final-year undergraduate ability (not taking prior knowledge into account). However, readability indices such as these are not very reliable predictors of ease of comprehension. The point of the comparison is to help confirm the objective differences between the texts chosen for this study.

The first step in preparing the materials was to reduce the size of the text, and eliminate all the quantitative data. Appendix A shows the original texts, with the material actually used in the experiment highlighted in italics. Sentences were chosen on the basis of the experimenter's intuition, as being those that captured the important points of the text. This often tended to coincide with the first and last sentences of paragraphs. (These are the usual positions for topical sentences.) Thus, all elaborative details, irrelevant for understanding the main point, were discarded, and each text was thereby reduced to approximately 1100 words.

An example of this kind of elimination of elaborative material is shown below (taken from the social psychology text). The italicised portions are those actually selected for use in the experiment.[3]

> *Altruistic actions, generally referred to as prosocial behaviour, involve activities where one person aids or assists another, but receives no obvious reward for his help.* Studies in this area have investigated helping behavior, such as aiding a seizure victim, protecting property from theft, or offering assistance to a motorist with a flat tire. There are, however, behaviors which aid or help others but are less dramatic, where the cost of carrying them out is considerably less, which involve less time, and where not engaging in the behavior has less consequence for another individual. Examples include informing someone that they have forgotten to turn off the lights of their car. *These behaviours would be regarded as accommodating, courteous, or thoughtful and an individual doing these acts would be seen as well-mannered, polite, and considerate.* (Goldman, Florez & Fuller, 1981.)

This one is taken from the perception text:

> So far as the author is aware, however, *the present data are among the first supporting age differences in the near periphery using a task that explicitly required peripheral discrimination as opposed to mere detection. Since the abilities of each of the two age groups to discriminate peripheral letters were assessed relative to their abilities to discriminate letters presented at the fixation point*, these results cannot be readily attributed to variations with age in extraneous factors such as motivation, decision bias, attentiveness, or memory. *A developmental difference specific to peripheral visual processing seems indicated.* (Taylor, 1982.)

The resulting texts yielded 65 sentences for the social psychology text, and 51 sentences for the perception text. Each sentence was then assigned to a category in

3. Differences in spelling for the italicised version are intentional - the original text used American spelling.

the superstructure, and to a position within each category. (In order to maintain local coherence, some slight rewording was necessary for some sentences — usually the addition of signals such as connectives — as well as some minor reordering within categories. See Appendix B.)

The superstructural format for this experiment was taken from Van Dijk (1980a) — see Figure 5.3. As Figure 5.3 shows, in Van Dijk (1977b), there were only nine

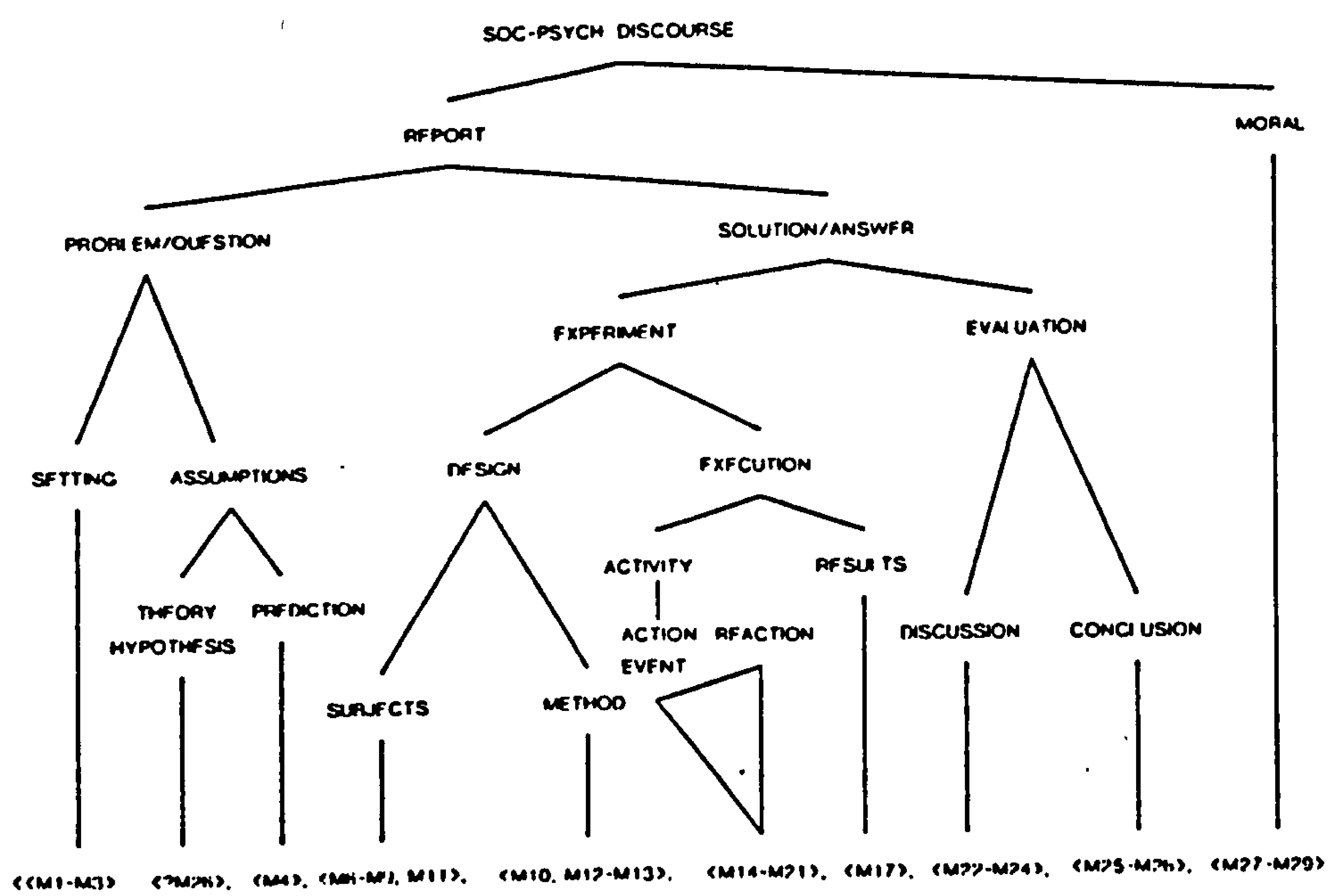

Figure 5.3. Superstructure for a psychology research report, taken from Van Dijk (1977b).

terminal nodes instead of sixteen, as there are in this analysis. This was felt to be justified, since the sentences that fall under the categories higher up in the hierarchy may be regarded as generalisations, and in some cases, metastatements of the information subsumed under them — that is, explicitly stated macropropositions. For example, although the node *problem* subsumes the nodes *setting*, *assumptions*, and so on, it seems reasonable for a text to contain some general summary sentences describing the problem at this level, and then going on to elaborate upon this in more detail. Furthermore, the materials used in this experiment were taken from naturally occurring texts in which there are sentences within the text corresponding to each of the nodes in the superstructural hierarchy — including those pertaining to non-terminal nodes. It is possible for each terminal node to contain sentences that differ in their generality with respect to the local topic. That is, it is possible, for example, to find in the node *acts*, a general statement of what happened, and then more detailed sentences describing the events or acts. However, the more general sentences might just as well correspond to the *execution* node, if one were to take that as a terminal node.

As an example of the resulting text, here are the sentences, according to superstructure category, for the social psychology text (see Appendix A for the context of the original sentences):

PROBLEM

1 The problems being investigated in this study are the factors influencing courteous behaviour.

SETTING

2 Researchers investigating prosocial behaviour have found that it is possible for a person to influence the likelihood of receiving help.

ASSUMPTIONS

3 Prosocial or altruistic behaviour involves activities where one person assists another but receives no obvious reward for his help.
4 These kind of behaviours are generally regarded as thoughtful or courteous.

HYPOTHESES

5 Social contact theory is related to helping behaviour.
6 This suggests that if A had previously had brief verbal contact with B, then B would be more likely to help A than if this initial contact had not occurred.
7 Norm-of-reciprocity theory states that people should help those who have helped them.
8 This theory suggests that if A has helped B, then B would be more likely to reciprocate by helping A.
9 Equity theory states that interacting individuals try to balance their rewards and costs.
10 This theory suggests that if A has helped B, then B has been rewarded at no cost to himself.
11 Therefore, B would be motivated to re-establish a balance by helping A at the first opportunity.

12 It has also been suggested that females are less prosocial than males.

PREDICTIONS

13 At first it may seem that the norm-of-reciprocity theory and the equity theory would make similar predictions.
14 Both theories would predict that if A helps B, then B would help A.
15 However, it is possible to set up a situation in which the two theories would make alternative predictions.
16 Equity theory implies that motivation to help will be induced regardless of whether or not the imbalance has been produced intentionally.
17 Therefore, if A helps B unintentionally, equity theory would predict that B would be likely to help A.
18 However, norm-of-reciprocity theory would predict that B would be likely to help A only if he thought that A had helped him intentionally.
19 If the norm-of-reciprocity theory is valid, then the intentional courtesy should produce more frequent courteous behaviour than no courtesy.
20 Equity theory would also predict this.
21 Equity theory would imply that courteous behaviour should occur to the same extent in both the unintentional and the intentional conditions.
22 The norm-of-reciprocity theory would predict that more courteous behaviour should occur in the intentional than in the unintentional condition.
23 Social contact theory would predict that if B has had previous contact with A, then B would be more likely to help than if no contact had occurred.
24 It would also be predicted that males would show more frequent courteous behaviour than females.

SOLUTION

25 An investigation of these hypotheses might reveal which factors influence courteous behaviour.

EXPERIMENT

26 The present study was designed to test the aforementioned predictions.

DESIGN

27 A two-by-four experimental design was used.

SUBJECTS

28 Eighty-four adult males and eighty-four adult females who were using a car-park next to a shopping centre served unwittingly as subjects.

METHOD

29 Sex, male and female, was compared with door-holding behaviour under conditions of control, unintentional, contact and intentional.

EXECUTION

30 Subjects were randomly assigned to each of the four conditions.
31 A male confederate, dressed similarly to the subject population, waited for a subject to park his car and walk towards the car-park exit.
32 A trial was begun when the confederate saw a subject walking alone towards the exit door.
33 In order to leave the car-park, subjects had to walk towards the exit door, open the door, proceed through a corridor, and finally open a second door at the end of the corridor.

34 It was noted whether or not the subject held the second door open for the confederate.

ACTS

35 In the control condition, the confederate preceded the subject to the exit door, opened the door and shut it behind him.
36 This required the subject to open the door for himself.
37 In the corridor, the confederate bent down to adjust his socks, which allowed the subject to arrive at the second door before the confederate.
38 In the unintentional condition, the confederate preceded the subject to the exit door and, whilst holding the door open, he bent down to tie his shoelaces.
39 He took no notice of the subject as the subject walked through the open door.
40 Then the confederate followed the subject to the second door.
41 In the intentional condition, the confederate preceded the subject to the exit door.
42 He held the door open, allowed the subject to walk through, and then followed the subject to the second door.
43 In the contact condition, the confederate preceded the subject to the exit door and held the door open, allowing the subject to pass through.
44 As soon as the subject had passed through the door, the confederate asked the subject for directions to the local theatre.
45 When the subject had responded the confederate thanked him and followed behind the subject to the second door.

RESULTS

46 The results of an analysis of variance showed that the mean courteous behaviour for males was significantly greater than for females.
47 The different door holding conditions yielded a significant difference in the rate of courteous behaviour.
48 The interaction of the main effects was not significant.
49 When compared with each other, the differences between means for control, unintentional and contact conditions were not significant.
50 The mean for the intentional condition was significantly greater than the means for each of the other conditions.

EVALUATION

51 The present study found that it is possible for an individual to influence the likelihood that another will behave in a courteous manner towards him.

DISCUSSION

52 The intentional condition induced significantly more courteous behaviour than did the control or unintentional conditions.
53 The unintentional condition did not produce significantly more courteous behaviour than the control condition.
54 However, it appears that intentional help must occur under specific conditions.
55 In the contact condition, the subject received intentional help.
56 This was followed by additional contact with the confederate asking for directions.

57 Courteous behaviour did not increase above that in the control condition.
58 The norm-of-reciprocity theory can account for this absence of increased door-holding behaviour.
59 The confederate acted courteously, holding the door open for the subject.
60 The subject then reciprocated by answering the confederate's questions.
61 Any obligation that the subject felt towards the confederate was removed by his responding to the confederate's request.
62 Another reason why the contact treatment did not produce increased courteous behaviour might have been that the subject was not certain that the confederate held the door open simply to be polite.
63 The subject might have understood that the confederate was holding the door open to reciprocate for the help for which the confederate was asking.

CONCLUSIONS
64 It appears, therefore, that the results support the norm-of-reciprocity theory and do not support the equity theory.
65 If courteous behaviour is to be reciprocated by an individual, then that individual needs to know that he was helped intentionally.

The next manipulation involved the structure of the text. Where the nodes of the superstructure matched the sections of the original text, the structure was preserved. However, subdivisions of these sections were made in order to include the remaining nodes in the Van Dijk hierarchy. In order to maintain local coherence, only the categories or units in the hierarchy were moved out of order. The sentences within each category are elaborations of the main category statement, and therefore, if within-category structure were disturbed, it would affect local coherence, possibly confounding the desired structuring effects (cf. Kintsch et al., 1977). Therefore, each sentence of the text was identified with a particular node in the hierarchy. Figure 5.4 shows the structure for the social psychology text (text1) and the perception text (text2), with the sentence numbers for the structured and the unstructured version (those in italics) at the corresponding nodes. (See Appendix C for the corresponding sentences.)

The order of presentation of the sentences was determined by traversing the graph top-down, depth-first, or left-to-right. In order to produce an unstructured version that did not match the superstructure, the order of the nodes of the graph was reversed. Therefore, the sentences that in the structured version of text1 fell into the category *results*, now fell into the category *predictions*, and so on. The resulting text was then checked, and altered if necessary, for microcoherence, by adding, for example, connectives. Here are examples of the resulting sentences, where both the structured and unstructured versions are grouped together for comparison. Thus,

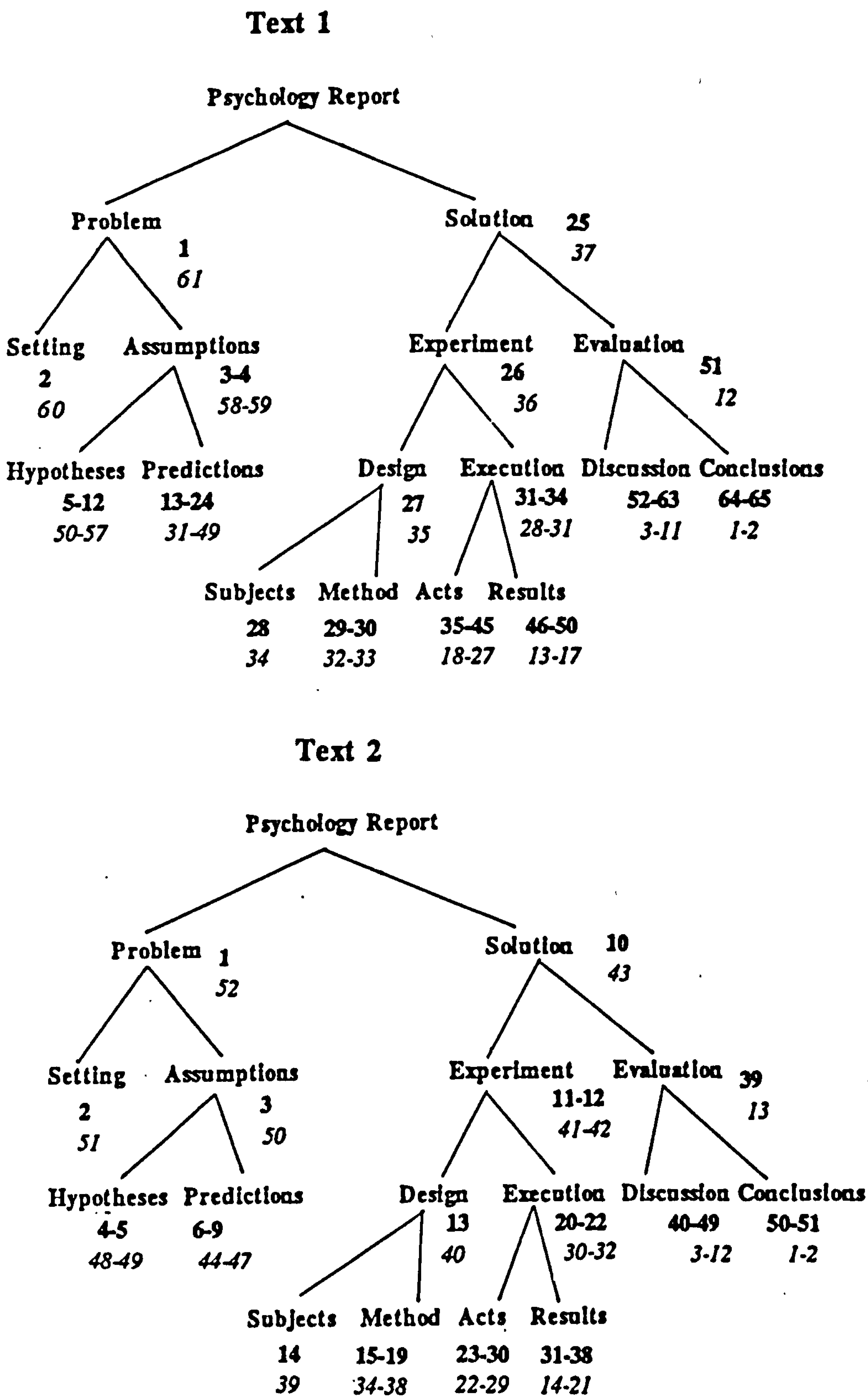

Figure 5.4. Structure for text1 and text2. The numbers in italics represent the sentences for the unstructured version.

sentence 1 in the original (structured) version of text1 becomes sentence 61 (in terms of presentation order) in the unstructured (*re*structured) version, and so on. (Appendix C shows the full text.) The words in italics refer to those which became slightly changed in the unstructured version:

1 *The problems being investigated in this study* are the factors influencing *courteous* behaviour.
61 *It also seems that* the factors *discussed here* are *indeed* those which influence *helping* behaviour.

2 *Researchers investigating prosocial behaviour have found* that it is possible *for a person* to influence the likelihood of receiving help.
60 *It seems, in conclusion,* that it is possible to influence the likelihood of receiving help.

3 Prosocial or altruistic behaviour involves activities where one person assists another but receives no obvious reward for his help.
58 Altruistic or prosocial behaviour involves activities where one person assists another but receives no obvious reward for his help.

4 These kind of behaviours are generally regarded as thoughtful or courteous.
59 These behaviours are generally regarded as thoughtful or courteous.

5 Social contact theory *is* related to helping behaviour.
50 Social contact theory *does appear to be* related to helping behaviour.

6 This suggests that if A had previously had brief verbal contact with B, then B would be more likely to help A than if this initial contact had not occurred.
51 This suggests that if A had previously had brief verbal contact with B, then B would be more likely to help A than if this initial contact had not occurred.

There were a few occasions where sentences from the structured version were either collapsed into one sentence in the unstructured version, or had to be presented twice — i.e., at two differents positions — in the unstructured version, in order to maintain local coherence. (This fact was taken into account when calculating the results.) For example, sentences 35 and 36 in text1 (the social psychology text), became sentence 18 in the unstructured version:

Structured version:

35 In the control condition, the confederate preceded the subject to the exit door, opened the door and shut it behind him.
36 This required the subject to open the door for himself.

Unstructured version:

18 In the control condition, a confederate preceded a subject to an exit door which led to a corridor, opened the door, and shut it behind him, which required the subject to open the door himself.

Sentences 55, 56 and 57 became sentence 6 in the unstructured version:

Structured version:

55 In the contact condition, the subject received intentional help.
56 This was followed by additional contact with the confederate asking for directions.
57 Courteous behaviour did not increase above that in the control condition.

Unstructured version:

6 In conditions with verbal contact, for example, where a subject receives intentional help, followed by additional contact with the confederate asking for directions, courteous behaviour does not increase above that in control conditions.

Sentences 59 and 60 became sentence 8 in the unstructured version:

Structured:

59 The confederate acted courteously, holding the door open for the subject.
60 The subject then reciprocated by answering the confederate's questions.

Unstructured:

8 The confederate acts courteously, for example by holding a door open for the subject, and the subject then reciprocates by answering the confederate's questions.

Finally, in the perception text (text 2), sentence 23 was presented both as sentence 22 and as sentence 33 in the unstructured version:

Structured:

23 Threshold exposure durations were obtained for letter arrays presented at each visual field position.

Unstructured:

22 Threshold exposure durations were obtained for letter arrays presented at each visual field position.
33 Using the method of ascending limits, threshold exposure durations were obtained for letter arrays presented at each visual field position.

(See Appendix A for the context of these sentences.)

The resulting "unstructured" version was also submitted to the same readability

Table 5.2

Readability analysis					
Variable		Text1		Text2	
		Structured	Unstructured	Structured	Unstructured
Readability	Kincaid	12.1	12.7	15.5	15.6
	Automated	11.9	12.6	16.5	16.6
	Cole-Liau	13.4	13.6	15.8	15.8
	Flesch	14.6	14.9	17.0	17.0
Sentence info.	av. sent. length	17.2	18.5	23.2	23.5
	av. word length	5.25	5.27	5.59	5.59
	av. nonfun length	6.91	6.92	7.2	7.16
	short sentences	22%	22%	25%	25%
	long sentences	8%	10%	12%	12%
Sentence types	simple	41%	33%	35%	33%
	complex	55%	58%	55%	60%
	compound	2%	5%	4%	4%
	compound-comp	3%	3%	6%	4%
Verb types	ToBe	28%	28%	33%	32%
	Auxiliary	33%	31%	15%	16%
	Infinitives	16%	17%	24%	23%
	Passives	16%	16%	20%	19%
Word usage	Prepositions	8.5%	9%	16%	14%
	Conjunction	2%	2.1%	2%	2%
	Adverbs	5%	5.2%	2%	1.9%
	Nouns	23.8%	24.1%	27.6%	27.8%
	Adjectives	16.1%	16.5%	24.5%	24.2%
	Pronouns	3.8%	4.1%	2.2%	2.6%
	Nominalisations	1%	1%	3%	3%
Sentence openers	Prepositions	13%	12%	16%	12%
	Adverbs	11%	10%	4%	4%
	Subject	69%	70%	63%	69%
	Verbs	0%	0%	0%	0%
	Sub conj.	6%	5%	12%	10%
	Conjunctives	0%	0%	0%	0%
	Expletives	2%	3%	6%	6%

analysis as the original text. As Table 5.2 shows, the restructuring did not make any substantial difference to the readability of the text, at least according to these indices.

The next factor involved the instructions given to the subjects for comprehending the text. This manipulation was included to examine the effects of making the structure of the text explicit to subjects. The instructions given to subjects in the structured and unstructured text conditions were exactly the same, except that in the unstructured condition, subjects were told that the text was unstructured. Instructions concerning the format of psychology reports was either given before reading the text (condition B), or after reading but before recall (condition A), or subjects were not informed of the format of psychology research reports (condition N). In all conditions, when asked to recall the unstructured text, subjects were asked to recall so that it took the form of a research report. Thus, here are the instructions for reading and recall in the structured and unstructured conditions, given *before* reading:

Structured text — Instructions for reading:

> A number of sentences will appear one at a time, on the screen in front of you. You are asked to read each sentence carefully, in order to understand it, and then to press the RETURN key once, when you are ready to move on to the next sentence.
>
> The text concerns a piece of research in psychology. In order to help you to organise the text in your mind, you are asked to bear in mind the following points concerning the reporting of psychological research:
>
> A research report usually begins with a statement of the *problem* to be investigated. There then follows an outline of the research conducted previously, whose findings are relevant to the current problem; this is the *setting* or *background*. A number of *assumptions* are then outlined, upon which various *theories* or *hypotheses* are based. From these hypotheses the researcher can draw various *predictions*. These predictions lead to a proposed *solution* to the problem outlined previously. This solution consists of an outline of the *experiment* to be conducted. The *design* of the experiment is outlined, with details of *subjects* and *method* of testing used. A description of the *execution* or procedure of the experiment is then given, with details of *acts* performed, and *results* obtained. These results are then *evaluated*, with a detailed *discussion* of the findings, from which various *conclusions* are drawn.

In summary, then, a psychological research report usually takes the following form:

Statement of problem
Setting or background of previous research
Assumptions
Theory or hypotheses
Predictions
Proposed solution to problem
Outline of experiment
Design used
Subjects
Method for testing
Procedure or execution of experiment
Details of acts performed
Results obtained
Outline evaluation of experiment
Detailed discussion
Conclusions

When you are ready to begin reading, press the RETURN key.

Instructions for recall:

Please write down as many of the sentences as you can recall, word for word if possible, and in the order in which you have read them. You may have as much time as you like.

Unstructured text — Instructions for reading:

A number of sentences will appear one at a time, on the screen in front of you. You are asked to read each sentence carefully, in order to understand it, and then to press the RETURN key once, when you are ready to move on to the next sentence.

The text concerns a piece of research in psychology. The text is taken from an actual research report, but has been slightly reworded and reordered. In order to help you to organise the text in your mind, you are asked to bear in mind the following points concerning the reporting of psychological research:

A research report usually begins with a statement of the *problem* to be investigated. There then follows ... *etc.*

When you are ready to begin reading, press the RETURN key.

Instructions for recall:

You are asked to rewrite the text you have read, without altering the actual content as far as possible, so that it takes the form of a research report, as outlined to you previously. You have as much time as you like.

(The full instructions are given in Appendix D.)

PROCEDURE

The text was organised into "chunks" of similar size, according to sentence or major phrase boundaries (cf. Jarvella, 1971), and presented, one sentence at a time, on a computer-controlled screen. Subjects were instructed to read each sentence carefully, taking as long as they liked, and to press RETURN as soon as they had understood the sentence, at which point the screen cleared and the next sentence was presented. At the end of the presentation, subjects were asked to leave the experimental room for 15 minutes, to get a cup of coffee. They were then brought back to the room, and each subject was instructed to write down as much as they could recall, word for word if possible. Subjects in the unstructured condition were asked to recall in the order conforming to a psychology research report. Reading times for the sentences were recorded by the computer.

In summary, the experimental design consisted of the within-subject factor, position of the sentence in the hierarchy, and four between-subject factors: Text type (social psychology text versus perception text), structure (original versus reordered version), knowledge ("high" versus "low"), and instructions (given either before reading, after

Table 5.3

Experimental Design				
	Text1 (social psychology report) N = 37		Text2 (perception report) N = 33	
	Structured N = 20	Unstructured N = 17	Structured N = 16	Unstructured N = 17
High Knowledge N = 39	A (4) B (4) N (3)	A (3) B (3) N (3)	A (3) B (3) N (3)	A (3) B (4) N (3)
Low Knowledge N = 31	A (3) B (3) N (3)	A (3) B (3) N (2)	A (2) B (3) N (2)	A (2) B (3) N (2)

reading, or not at all). (See Table 5.3 — the letters A, B, and N refer to the instructional conditions: after reading, before reading, and no instructions concerning the form of a research report.)

RESULTS

Reading Times[4]

Table 5.4

Mean Reading Time Per Sentence (secs)					
		Text1		Text2	
		Structured	Unstructured	Structured	Unstructured
High Knowledge	A	8.72	7.83	11.72	19.34
	B	4.37	8.26	13.98	14.03
	N	9.86	8.81	16.82	11.05
Low Knowledge	A	7.64	13.98	17.33	22.07
	B	10.14	11.15	18.03	19.10
	N	13.57	10.65	13.58	30.63

Table 5.4 shows the overall mean reading times per sentence for each of the experimental conditions (text, structure, knowledge, and instructions). Analysis of variance revealed significant main effects for text and for knowledge, but no other significant main effects and no interactions. Reading times for the perception text (text2) were longer than those for the social psychology text (text1) ($F(1, 46) = 29.850$, $p < .01$); reading times for the low-knowledge group were longer than for the high-knowledge group ($F(1, 46) = 9.010$, $p < .01$). Figure 5.5 shows the difference between mean reading times per sentence for the perception text and the

4. Raw data are available on request from the experimenter.

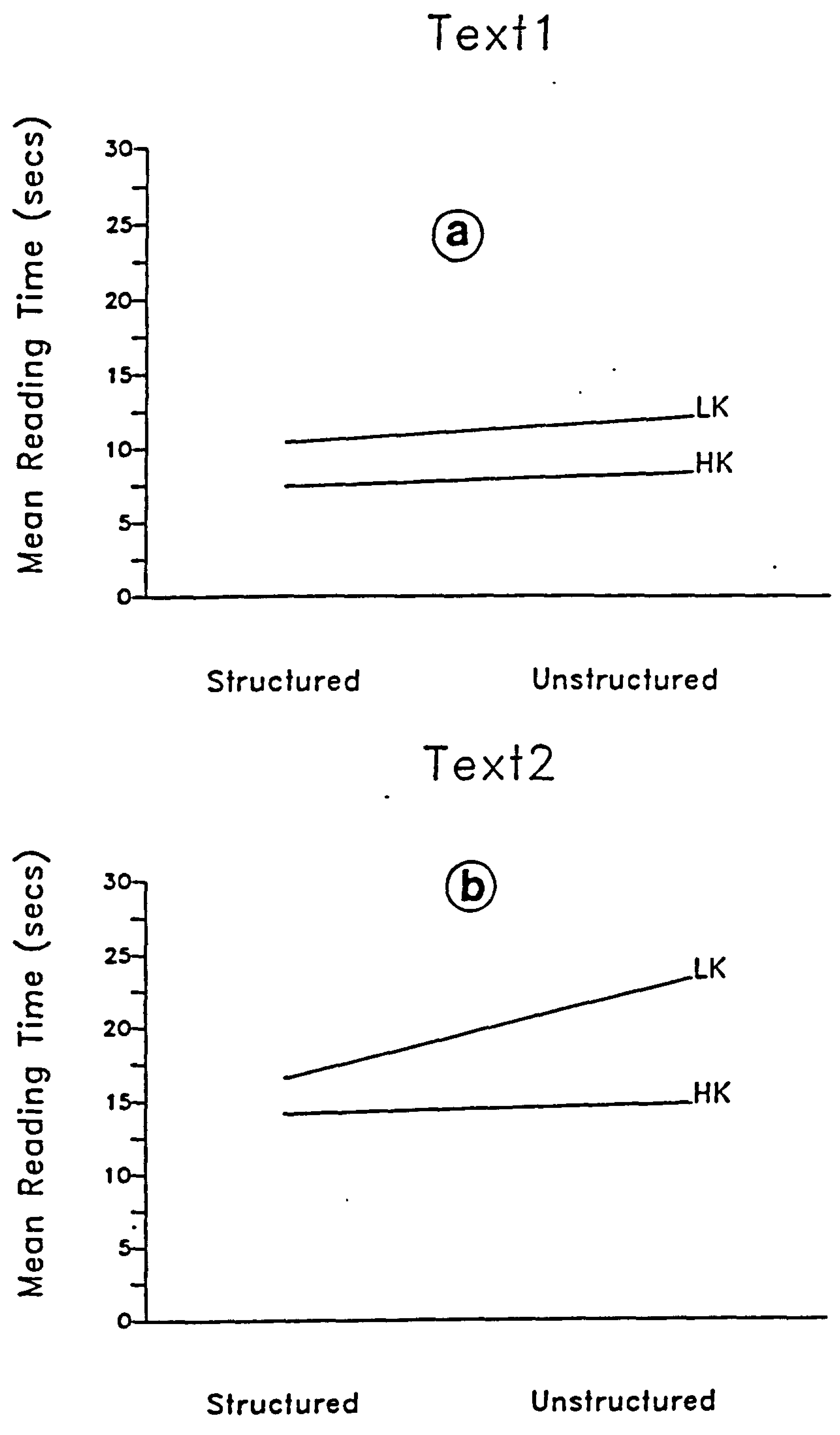

Figure 5.5. Mean reading time per sentence (secs).

Table 5.5

Mean Reading Time Per Sentence (secs)				
	Text1		Text2	
	Structured	Unstructured	Structured	Unstructured
High Knowledge	7.45	8.30	14.17	14.73
Low Knowledge	10.45	12.09	16.56	23.24
F table				
	df	*MS*	*F*	*p*
Text	1	950.92	30·48	0.00**
Knowledge	1	287.02	9·20	0.00**
MSE	62	31·19		

social psychology text. Table 5.5 shows the means and *F* values. The apparent knowledge by structure interaction for the perception text (text2) was not significant.

These results are for actual mean reading times. However, these are not entirely comparable, since sentence length varied both within and across texts. This factor was not controlled experimentally since this study concerns naturalistic materials. Therefore, the reading time data were adjusted post hoc in order to control for sentence length.

Some researchers have used number of words per sentence as a predictor of reading time (e.g., Haberlandt, 1980), while others have used number of syllables per

Table 5.6

Mean Reading Time Per Word (secs)					
		Text1		Text2	
		Structured	Unstructured	Structured	Unstructured
High Knowledge	A	0.54	0.42	0.51	0.82
	B	0.27	0.44	0.58	0.61
	N	0.60	0.47	0.69	0.46
Low Knowledge	A	0.49	0.76	0.72	0.90
	B	0.62	0.62	0.77	0.79
	N	0.83	0.56	0.56	1.28

sentence (e.g., Olson, Mack & Duffy, 1981). Table 5.6 shows the mean reading times per sentence, after dividing the raw data by number of words per sentence. Analysis of variance revealed significant main effects for text ($F(1, 46) = 6.760, p < .05$) and for knowledge ($F(1, 46) = 10.605, p < .01$), with the perception text once again having longer reading times than the social psychology text, and low-knowledge subjects having longer reading times than high-knowledge subjects.

Figure 5.6 shows the difference between the means. Table 5.7 shows the means and F values. Once again, the apparent knowledge by structure interaction was not significant.

The same pattern of results held when the data were adjusted instead for number of

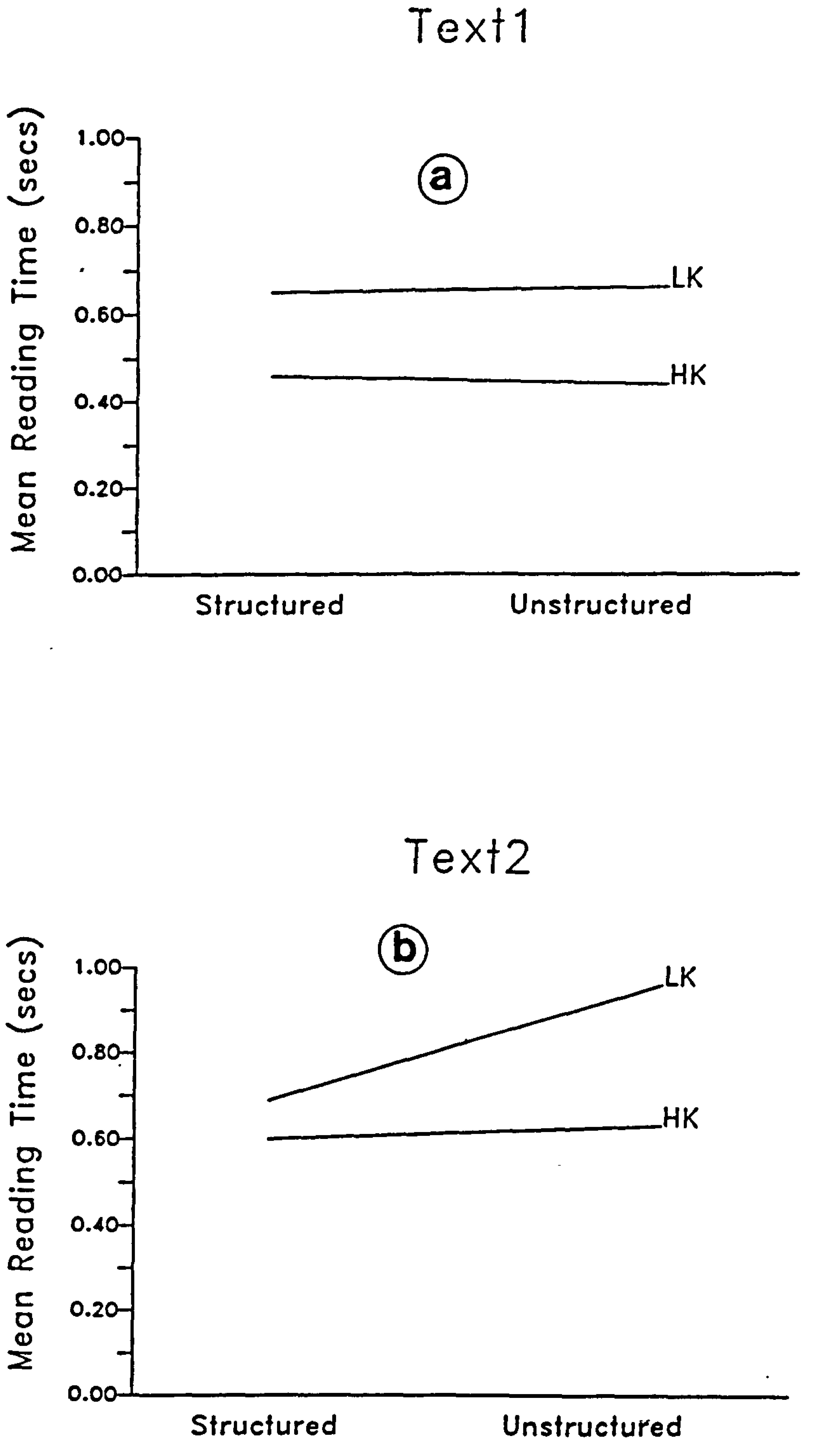

Figure 5.6. Mean reading time per word (secs).

Table 5.7

Mean Reading Time Per Word (secs)				
	Text1		Text2	
	Structured	Unstructured	Structured	Unstructured
High Knowledge	0.46	0.44	0.60	0.63
Low Knowledge	0.65	0.66	0.69	0.96
F table				
	df	*MS*	*F*	*p*
Text	1	0.45	6.76	0.01*
Knowledge	1	0.71	10.61	0.00**
MSE	62	0.07		

Table 5.8

Mean Reading Time Per Syllable (secs)					
		Text1		Text2	
		Structured	Unstructured	Structured	Unstructured
High Knowledge	A	0.30	0.24	0.27	0.44
	B	0.15	0.24	0.31	0.33
	N	0.33	0.27	0.37	0.25
Low Knowledge	A	0.27	0.43	0.38	0.48
	B	0.35	0.35	0.41	0.42
	N	0.46	0.31	0.30	0.69

syllables per sentence, as Tables 5.8 and 5.9 and Figure 5.7 show. The perception text had longer reading times than the social psychology text ($F(1, 62) = 4.76$., $p <$ .05). and the low-knowledge group took longer to read the sentences on the whole

Table 5.9

Mean Reading Time Per Syllable (secs)				
	Text1		Text2	
	Structured	Unstructured	Structured	Unstructured
High Knowledge	0.25	0.25	0.32	0.34
Low Knowledge	0.36	0.37	0.37	0.52
F table				
	df	*MS*	*F*	*p*
Text	1	0.10	4.76	0.04*
Knowledge	1	0.22	10.90	0.00**
MSE	62	0.02		

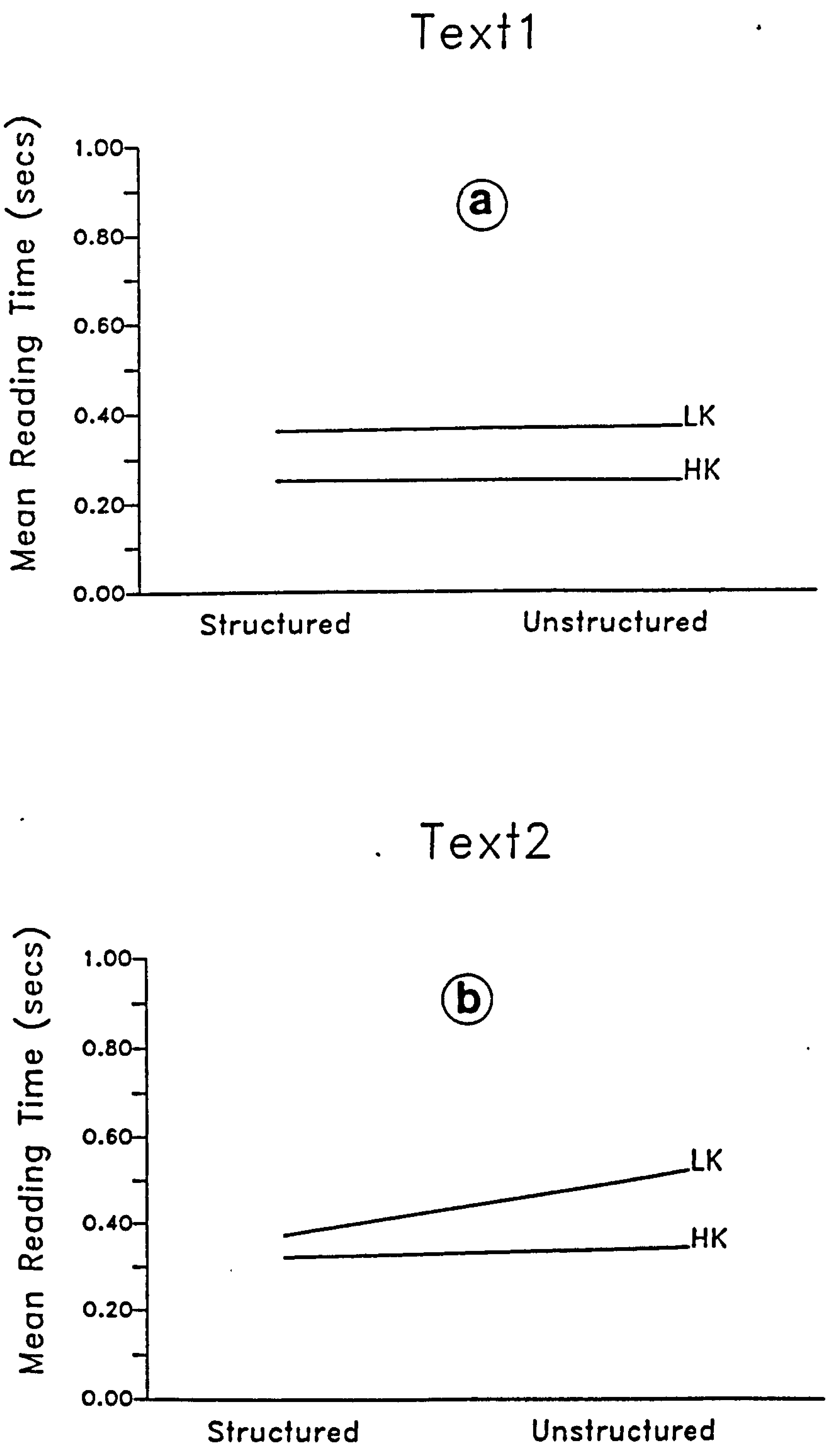

Figure 5.7. Mean reading time per syllable (secs).

than the high-knowledge group ($F(1, 62) = 10.90$, $p < .01$).

However, simply dividing by number of words or syllables is a rather crude control, and assumes a directly proportional relationship between reading time and number of words or syllables. Therefore, the data were subjected to a regression analysis to further examine the relationship between reading times and these predictors.

The two predictors of number of words and number of syllables per sentence are not independent, so one cannot use both of them to adjust the reading times. A choice was made, therefore, on the basis of the largest amount of variance (R^2) accounted for by either number of words or number of syllables, in a simple regression analysis.

This analysis was performed for those subjects who did not receive instructions before reading (i.e., instructional conditions A and N — see Appendix D), in case this affected the results of the regression. For each experimental condition (i.e., knowledge, structure, and text) means were entered into a regression analysis, with either number of words per sentence, or number of syllables per sentence as the independent (predictor) variable. The percentage of variance (based on R^2) accounted for by these variables is shown in Table 5.10. The table shows that a larger percentage of variance was accounted for by number of syllables, so this was chosen as the factor to control for sentence length. The table also shows that more variance was accounted for in the unstructured condition than in the structured condition, that more variance was accounted for in the perception text than in the social psychology text, and that more variance was accounted for in the low-

Table 5.10

% Variance Accounted For By No. of Words				
	Text1		Text2	
	Structured	Unstructured	Structured	Unstructured
High Knowledge	22.77	25.86	42.69	46.54
Low Knowledge	8.93	14.63	40.56	48.79
% Variance Accounted For By No. of Syllables				
	Text1		Text2	
	Structured	Unstructured	Structured	Unstructured
High Knowledge	39.22	52.19	48.03	54.52
Low Knowledge	16.21	37.68	45.13	56.17

knowledge condition than in the high-knowledge condition.

It was decided to adjust raw scores individually, that is, by deriving regression equations for each subject's reading time and using that to adjust the scores individually. Table 5.11 shows the mean proportion of variance (R^2) accounted for (expressed in percentages) in each condition. Analysis of variance for R^2 revealed a significant main effect for structure ($F(1, 46) = 14.406$, $p < .01$), more variance being accounted for in the unstructured version, and for text ($F(1, 46) = 20.264$, $p < .01$), more variance being accounted for with the perception text. There was also a structure by text interaction ($F(1, 46) = 9.062$, $p < .01$), less variance being

Table 5.11

% Variance Accounted For By Step1 (Syllables)				
	Text1		Text2	
	Structured	Unstructured	Structured	Unstructured
High Knowledge	13.83	33.57	33.26	30.64
Low Knowledge	11.98	28.23	33.53	37.76

accounted for in the structured version of the social psychology text.

Table 5.12 shows the mean residual reading times, produced by subtracting reading times predicted by the regression with number of syllables as the independent variable, from the actual, or observed, reading times. Analysis of variance revealed no significant differences between the actual residuals, but analysis based on the size of the difference between observed and predicted reading times (i.e., the extent to which the residuals departed from zero) revealed significant main effects for text

Table 5.12

Mean Residual Reading Time (Step1) (msecs)				
	Text1		Text2	
	Structured	Unstructured	Structured	Unstructured
High Knowledge	0.15	-0.12	-0.02	-0.14
Low Knowledge	-0.02	-0.24	0.01	-0.57
F table				
	df	*MS*	*F*	*p*
Text	1	72.51	8.09	0.00**
Knowledge	1	63.25	7.06	0.00**
MSE	62	8.96		

($F(1, 62) = 8.09$ $p < .01$), the perception text producing larger differences than the social psychology text. There was also a significant main effect for knowledge ($F(1, 62) = 7.06$, $p < .01$), the low-knowledge group showing larger differences than the high-knowledge group. This is, in effect, a measure of the variability of reading times within each condition, after controlling for sentence length.

Figures 5.8a-d and 5.9a-d show the mean reading times per sentence for each condition (that is, before any adjustment). Since there appears to be some effect of presentation order, or serial position, across the whole text, the adjusted scores (that is, those scores derived from a regression analysis with number of syllables as the predictor) were subjected to a further regression with serial position as the predictor. That is, a forward stepwise multiple regression was performed on the raw data, with number of syllables and serial position as independent variables, entered in that order.

Table 5.13

% Variance Accounted For By Step2 (Serial Posn.)				
	Text1		Text2	
	Structured	Unstructured	Structured	Unstructured
High Knowledge	16.92	11.80	20.89	8.03
Low Knowledge	19.06	11.06	30.67	18.15

Table 5.13 shows the percentage of additional variance accounted for by the regression using serial position as a predictor variable, and Table 5.14 shows the total percentage of variance (cumulative R^2) accounted for by the two-step regression. When the contribution made by this second step in the regression was examined, there was a significantly higher amount of variance accounted for in the structured condition ($F(1, 46) = 7.204$, $p < .05$). In terms of total variance

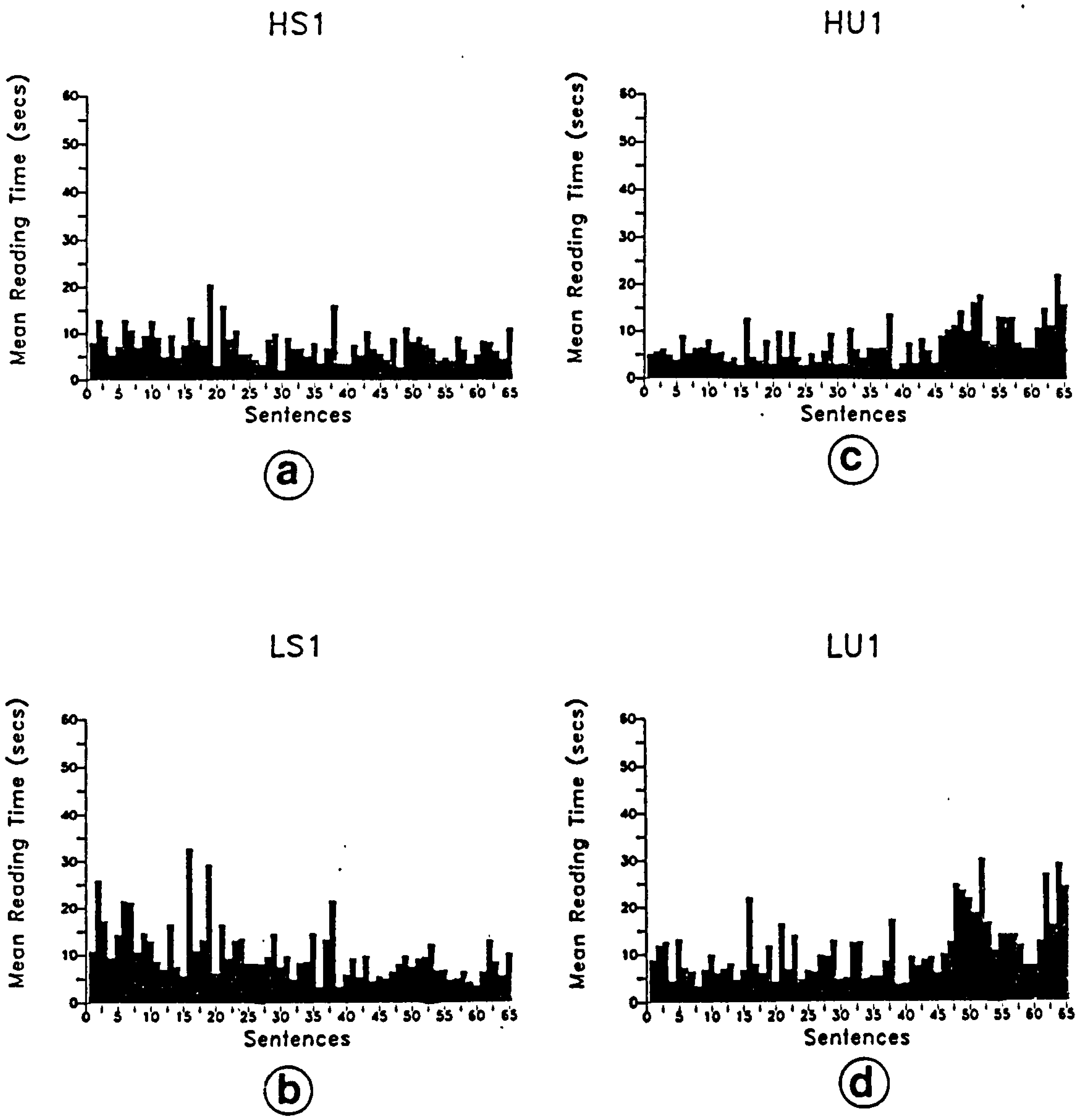

Figure 5.8. Mean reading times per sentence (text1). (The order of sentences for the unstructured conditions — i.e., HU1 and LU1 — corresponds to the order for the structured version.)

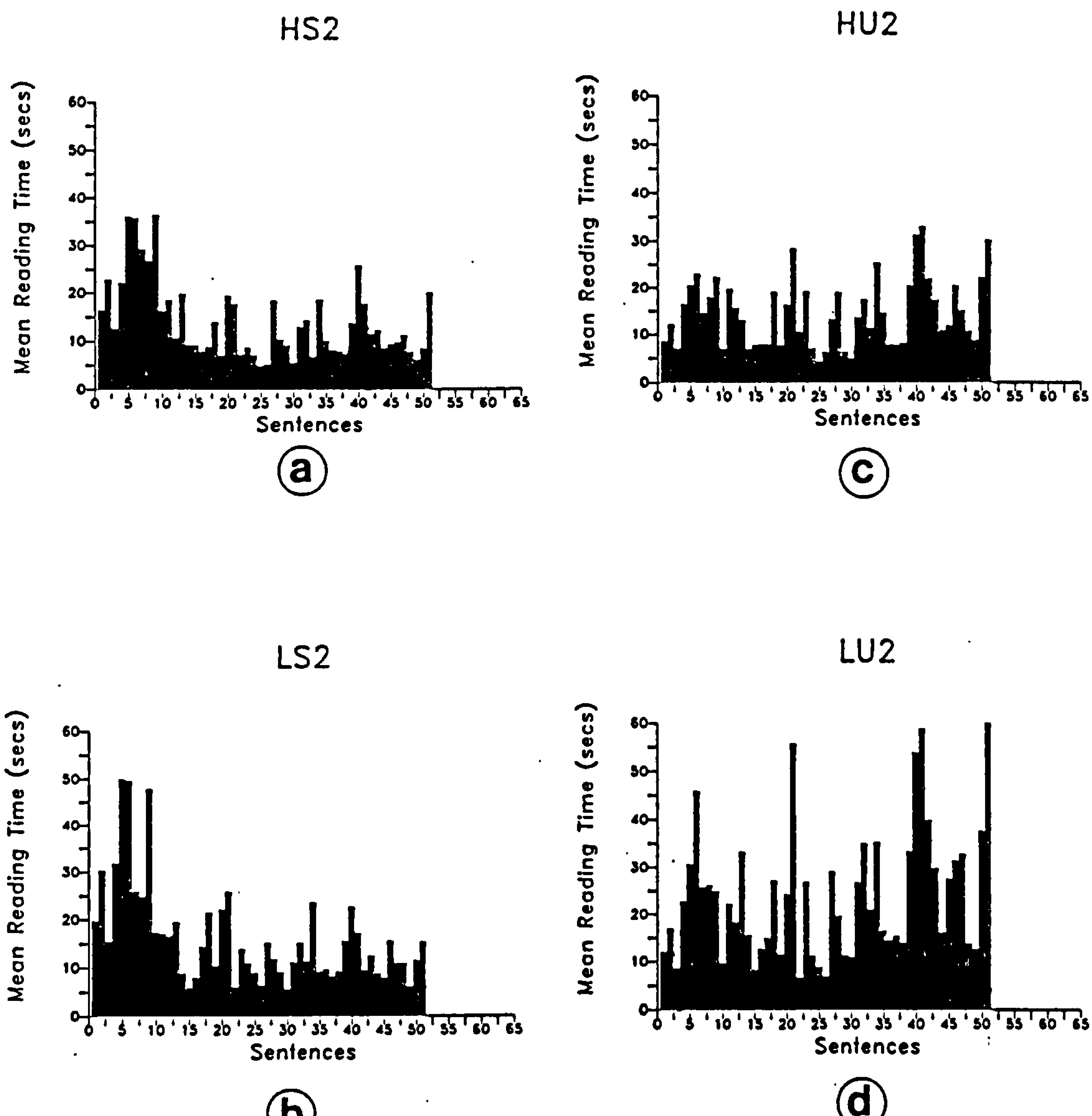

Figure 5.9. Mean reading times per sentence (text2). (The order of sentences for the unstructured conditions corresponds to the order for the structured version.)

Table 5.14

% Total Variance Accounted For				
	Text1		Text2	
	Structured	Unstructured	Structured	Unstructured
High Knowledge	30.75	45.37	54.15	38.67
Low Knowledge	31.04	39.29	64.20	55.91

(cumulative R^2), it can be seen that more variance was accounted for in the perception text, but analysis of variance on these percentages revealed significant interactions with knowledge ($F(1, 46) = 4.588, p < .05$) and structure ($F(1, 46) = 9.372, p < .01$), as well as a significant main effect for text ($F(1, 46) = 15.748, p < .01$).

Residual means were derived by subtracting means predicted by the two-step regression from the actual reading times. Table 5.15 shows the resulting mean residuals. Analysis of variance revealed no significant effects between these residuals, but when the size of the difference between observed and predicted reading times was examined, significant main effects were found for text ($F(1, 46) = 8.859, p < .01$) and knowledge ($F(1, 46) = 6.683, p < .05$), but there was a slight four-way interaction between text, knowledge, structure, and instructions ($F(2, 46) = 4.378, p < .05$).

Table 5.15

Mean Residual Reading Time (msecs)					
		Text1		Text2	
		Structured	Unstructured	Structured	Unstructured
High Knowledge	A	-0.51	0.49	0.12	0.61
	B	-0.74	-0.36	-0.60	-0.45
	N	0.77	0.13	0.16	-0.79
Low Knowledge	A	-0.27	0.13	0.43	0.73
	B	-0.47	-0.17	0.18	0.16
	N	0.11	0.003	0.91	0.25

F table

	df	*MS*	*F*	*p*
Text	1	56.23	8.86	0.00**
Knowledge	1	42.41	6.68	0.01*
TSKI	2	27.79	4.38	0.02*
MSE	46	0.02		

Having examined reading times for the text as a whole, the patterns of reading times across sentences within the text were examined. Figures 5.10a-d and 5.11a-d show the distribution of mean residual reading times for each condition. These figures confirm the finding suggested by amount of variance accounted for in the regression, that reading times were more variable in low-knowledge conditions and with the unstructured texts, especially in the case of the perception text (text2). Table 5.16 shows the proportion of residuals falling between -3.0 and 3.0 for each condition.

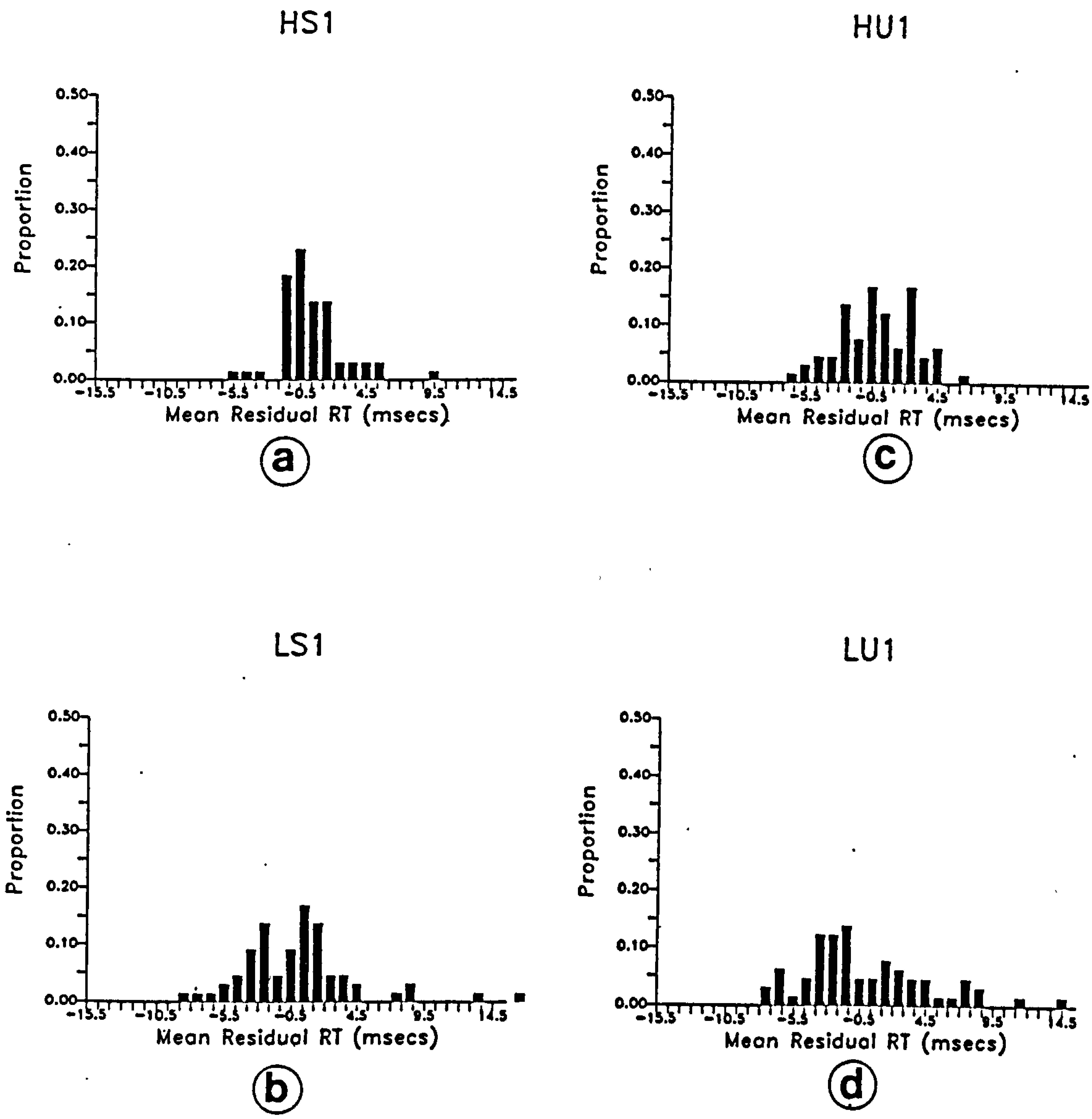

Figure 5.10. Distribution of mean residual reading times per sentence (text1).

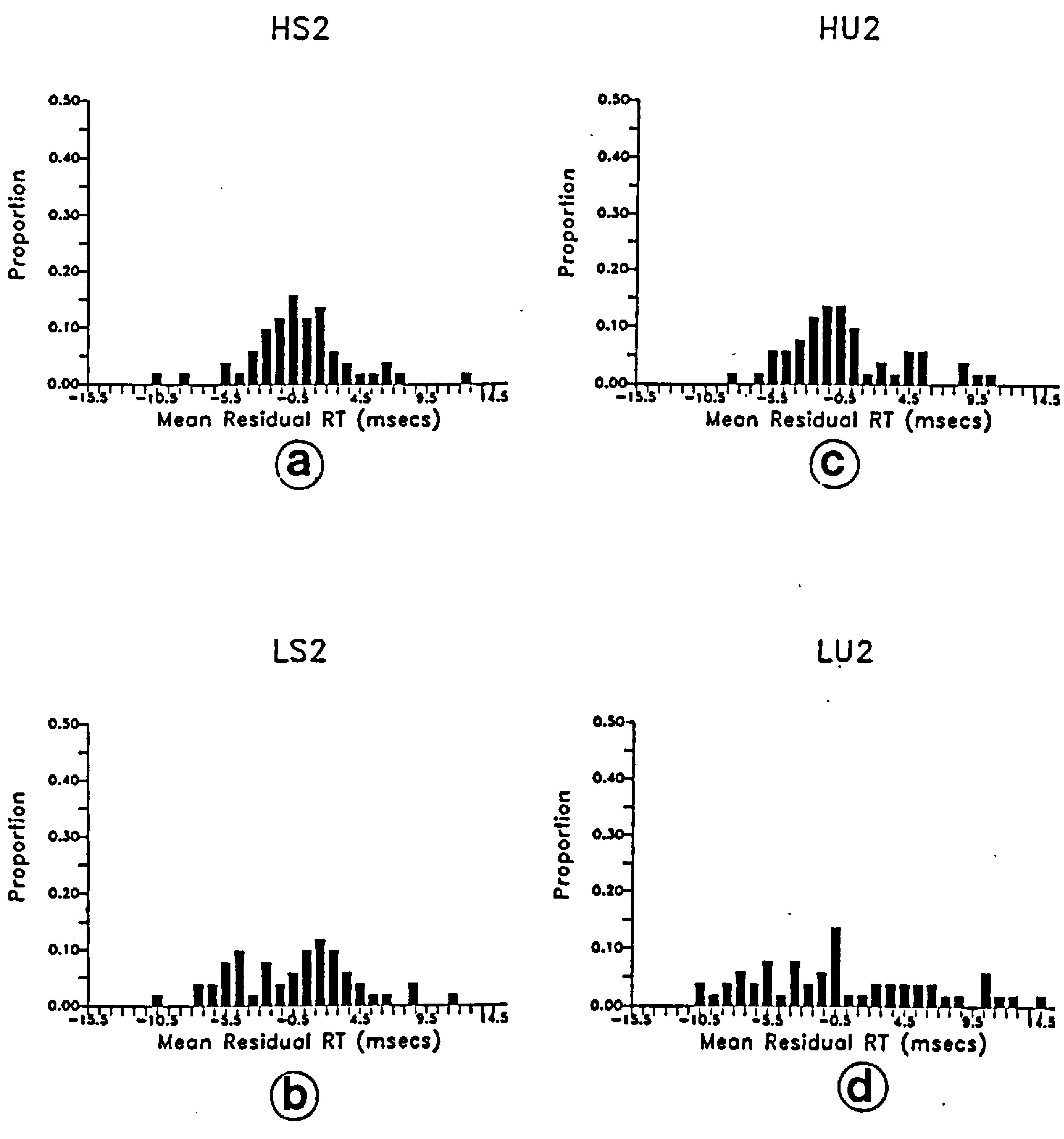

Figure 5.11. Distribution of mean residual reading times per sentence (text2).

The mean residual reading times per sentence for each condition were submitted to a correlational analysis, in order to examine the relationship between the pattern of

Table 5.16

% Residual RTs Between -3.0 and 3.0				
	Text1		Text2	
	Structured	Unstructured	Structured	Unstructured
High Knowledge	84.6	73.8	68.7	54.9
Low Knowledge	62.9	49.2	49.0	31.4

residuals. Table 5.17 shows the resulting correlations. It is important to note that these are correlations between exactly the same sentences in the structured and unstructured conditions for both texts, having adjusted the scores for both sentence length (i.e., number of syllables) and presentation order (i.e., serial position). The data indicate that reading time patterns for high-knowledge and low-knowledge subjects with either the structured or the unstructured version were highly correlated, whereas those between structured and unstructured versions were very

Table 5.17

Correlation Between Mean Residual RTs								
	HS1	LS1	HU1	LU1	HS2	LS2	HU2	LU2
HS1		0.61	0.27	0.20				
LS1	0.61		0.25	0.30				
HU1	0.27	0.25		0.77				
LU1	0.20	0.30	0.77					
HS2						0.72	0.29	0.17
LS2					0.72		0.24	0.02
HU2					0.29	0.24		0.72
LU2					0.17	0.02	0.72	

weakly correlated when they were significant.

These correlations were also performed for those sentences falling within categories, which were not reordered in the structure manipulation — i.e., those sentences in the middle of, and not cutting across, category boundaries. The effect was to

Table 5.18

Correlation Between Mean Residual RTs								
	HS1	LS1	HU1	LU1	HS2	LS2	HU2	LU2
HS1		0.62	0.44	0.30				
LS1	0.62		0.33	0.39				
HU1	0.44	0.33		0.79				
LU1	0.30	0.39	0.79					
HS2						0.60	0.29	0.17
LS2					0.60		0.28	0.05
HU2					0.29	0.28		0.81
LU2					0.17	0.05	0.81	

slightly raise the correlation values for each comparison, as Table 5.18 shows.

Several researchers (e.g., Haberlandt, 1980; Mandler & Goodman, 1982) have used increases in reading times at category boundaries as an indication of the validity of the hypothesised text structure. This analysis was performed for the data presented here. Mean residuals from sentences at the beginning of categories were compared with those at the end and in the middle. It was expected that reading times would be longest for sentences at the beginning, shortest for those in the middle, with the rest falling somewhere in between (sentences at the end of categories may or may not summarise the local topic). Table 5.19 shows the mean residuals for each condition. It indicates that, on the whole, these predictions were not confirmed. The only conditions meeting the predicted pattern occur in those groups receiving the unstructured version of the social psychology text. This may have also been affected by whether or not subjects received instructions before reading the text, so

Table 5.19

Mean Residual RTs (Boundary Comparison)			
	Beginning	End	Middle
HS1	-1.02	-0.86	0.60
LS1	-0.44	-0.55	0.30
HU1	0.20	0.13	-0.12
LU1	0.90	0.12	-0.43
HS2	-0.99	0.98	0.27
LS2	-1.93	2.17	0.44
HU2	-0.42	0.04	0.02
LU2	-0.35	-1.58	0.65

the same analysis was performed between those who had instructions beforehand and those who did not. Table 5.20 shows the results of this analysis. This time the prediction was confirmed for low-knowledge subjects who had received instructions with the social psychology text, both in the structured and unstructured conditions. The effect also appeared in conditions where subjects did not received instructions, but once again only for subjects receiving the unstructured version of the social psychology text.

When sentences appearing at both the beginning and the end of categories were compared with those occurring in the middle, they were not any slower than those in the middle, except for subjects receiving the unstructured version of the social

Table 5.20

Mean Residual RTs (Boundary Comparison) With Instructions (Condition B)			
	Beginning	End	Middle
HS1	-0.23	0.09	0.73
LS1	0.87	-1.49	-0.02
HU1	-0.15	0.81	-0.14
LU1	0.61	-0.16	-0.23
HS2	0.23	0.56	-0.34
LS2	-2.26	0.42	1.25
HU2	-0.13	-0.90	0.23
LU2	0.34	-3.68	1.01
Without Instructions (Condition N)			
	Beginning	End	Middle
HS1	-1.42	-1.33	0.87
LS1	-1.09	-0.06	0.45
HU1	0.37	-0.22	-0.11
LU1	1.05	0.26	-0.53
HS2	-1.60	1.18	0.57
LS2	-1.76	3.04	0.03
HU2	-0.57	0.51	-0.08
LU2	-0.69	-0.53	0.48

psychology text. This pattern held whether or not subjects had instructions beforehand (see Table 5.21).

Table 5.21

Mean Residual RTs (Boundary Comparison) With Instructions (Condition B)		
	Beginning/End	Middle
HS1	-1.39	0.73
LS1	-0.72	-0.02
HU1	0.16	-0.14
LU1	0.77	-0.23
HS2	-0.60	-0.34
LS2	-0.03	1.25
HU2	-0.18	0.23
LU2	-0.63	1.01
Without Instructions (Condition N)		
	Beginning/End	Middle
HS1	-0.12	0.87
LS1	0.02	0.45
HU1	0.20	-0.11
LU1	0.33	-0.53
HS2	0.35	0.57
LS2	-1.30	0.03
HU2	-0.41	-0.08
LU2	-1.1	0.48

Recall

Each of the texts given to subjects was coded into "main idea units" by three independent judges (see Appendix E for instructions given to coders). The resulting divisions were then checked for agreement, and it was decided on the basis of this to exclude any divisions made by only one out of the three judges. For the original texts given to subjects, the divisions made by at least two out of the three judges accounted for between 56% and 69% of the total number of divisions. Thus, on average, about 38% of divisions were discounted (that is, those which only one judge had decided upon). Table 5.22 shows the breakdown of divisions. (Appendix F shows the texts as they were originally divided.) The same judges were then given the recall protocols to divide into "idea units", as before. About 16% of divisions made by judges were decided upon by only one judge, so these were discounted. The texts and the protocols were then divided into units based on agreement between at least 2 out of the three judges. Each unit was marked, and its corresponding sentence number noted.

In order to produce scoring sheets for the original texts, the most frequent divisions made by coders (i.e., where it was more than 1) were used to divide up the sentences. The structured versions were then compared with the unstructured versions, and, where there was a discrepancy in the units (only occasional) a decision was made on where to make the divisions, based on cases where at least one coder made a division, or otherwise based on whether or not it seemed to make sense to do so. Thus, the structured and unstructured versions were divided into the same units. When collating the divisions made by the coders, an arbitrary decision was made about where to place a mark if it was at a boundary with a conjunction (e.g., "and", "but", etc.). Appendix G shows the resulting texts used for scoring. The divided protocols were then given to the same three judges, together with the corresponding original text, numbered according to idea units, and the judges were asked to score each unit of the protocols for recall. (See Appendix E

Table 5.22

Divisions made by coders				
	Structured		Structured	
no. coders	no. divs	%	no. divs	%
tot	152		86	
1	66	43.42%		
2	18	11.84%	18	20.93%
3	68	44.74%	68	79.07%
	Structured		Structured	
no. coders	no. divs	%	no. divs	%
tot	130		78	
1	52	40%		
2	20	15.38%	20	25.64%
3	58	44.61%	58	74.36%
	Unstructured		Unstructured	
no. coders	no. divs	%	no. divs	%
tot	124		85	
1	39	31.45%		
2	12	9.68%	12	14.12%
3	73	58.87%	73	85.88%
	Unstructured		Unstructured	
no. coders	no. divs	%	no. divs	%
tot	118		75	
1	43	36.44%		
2	13	11.02%	13	17.33%
3	62	52.54%	62	82.67%

for instructions for scoring.)

Two criteria were used in order to collate the 3 scores: A *loose* criterion was based on agreement between at least 2 scorers; a *strict* criterion was based on agreement between all 3 scorers. Each subject's score was then transformed from proposition numbers to sentence numbers, in order to make the comparison between sentences presented (read) and those recalled easier. In calculating whether or not a whole sentence was correctly recalled, a sentence was counted as being correctly recalled if at least one proposition was recalled. For purposes of ordering, only the sentence that appeared first in the recall protocol was counted. Each protocol was checked for repetitions so that only those sentences recalled once were used in the analysis.

In terms of number of sentences recalled, there were no significant differences

Table 5.23

		Text1		Text2	
% Recalled (Loose Criterion)					
		Struct	Unstruct	Struct	Unstruct
High Knowledge	A	44.62	37.16	37.91	25.64
	B	21.92	26.23	39.22	22.12
	N	40.00	34.97	23.53	23.08
Low Knowledge	A	30.77	30.60	31.37	24.04
	B	36.41	28.96	29.41	25.64
	N	31.80	30.33	29.41	29.81
% Recalled (Strict Criterion)					
		Struct	Unstruct	Struct	Unstruct
High Knowledge	A	28.85	21.86	18.95	8.33
	B	13.08	14.21	22.22	11.06
	N	25.13	24.04	13.07	11.54
Low Knowledge	A	15.90	15.30	19.61	12.50
	B	25.13	14.75	16.99	12.18
	N	25.13	13.94	13.73	16.35

between high and low-knowledge groups for either version of either text. Table 5.23 shows the mean proportion of sentences recalled (expressed as percentages) for each condition. Analysis of variance, using the loose criterion, revealed a significant main effect for structure ($F(1, 46) = 4.187, p < .05$), those with the structured version recalling more than those with the unstructured version. Using the strict criterion, however, analysis of variance revealed a significant main effect for text ($F(1, 46) = 10.734, p < .01$), those with the perception text (text2) recalling less than those with the social psychology text. There was also a significant main effect for structure ($F(1, 46) = 10.582, p < .01$), those with the structured version recalling more than those with the unstructured version, but no significant interactions involving these factors. However, there was a three-way interaction between text, knowledge and instructional condition ($F(2, 46) = 3.397, p < .05$). Since the N per cell was very low for this analysis, and since the interaction only just reached significance ($p = 0.042$) this analysis, and the succeeding ones for amount recalled, excluded instructional conditions as a factor.

Once again, there was a significant main effect for structure, using the loose criterion ($F(1, 62) = 4.103, p < .05$), those with the structured versions recalling more than those with the unstructured versions (see Figures 5.12a-d and Table 5.24). Using the strict criterion, there was a significant main effect for text ($F(1, 62) = 9.808, p < .01$), those with the perception text recalling less than those with the social psychology text. There was also a significant main effect for structure ($F(1, 62) = 9.669, p < .01$), those with the structured versions recalling more than those with the unstructured versions. The graphs in Figure 5.12d show that there is a slight interaction of structure by knowledge, but this is not significant.

The number of categories or nodes recalled per subject was estimated, based on recalling at least one sentence from a category. Using the loose criterion, there was a significant main effect of text ($F(1, 62) = 4.740, p < .05$), those with the social psychology text recalling more categories or nodes than those with the perception text (see Figures 5.13a-d and Table 5.25). Using the strict criterion, there was a significant main effect for text ($F(1, 62) = 11.037, p < .01$) and structure ($F(1, 62) = 14.445, p < .01$), those with the social psychology text recalling more categories than those with the perception text, and those with the structured versions recalling more than those with the structured versions. There was also a significant three-way

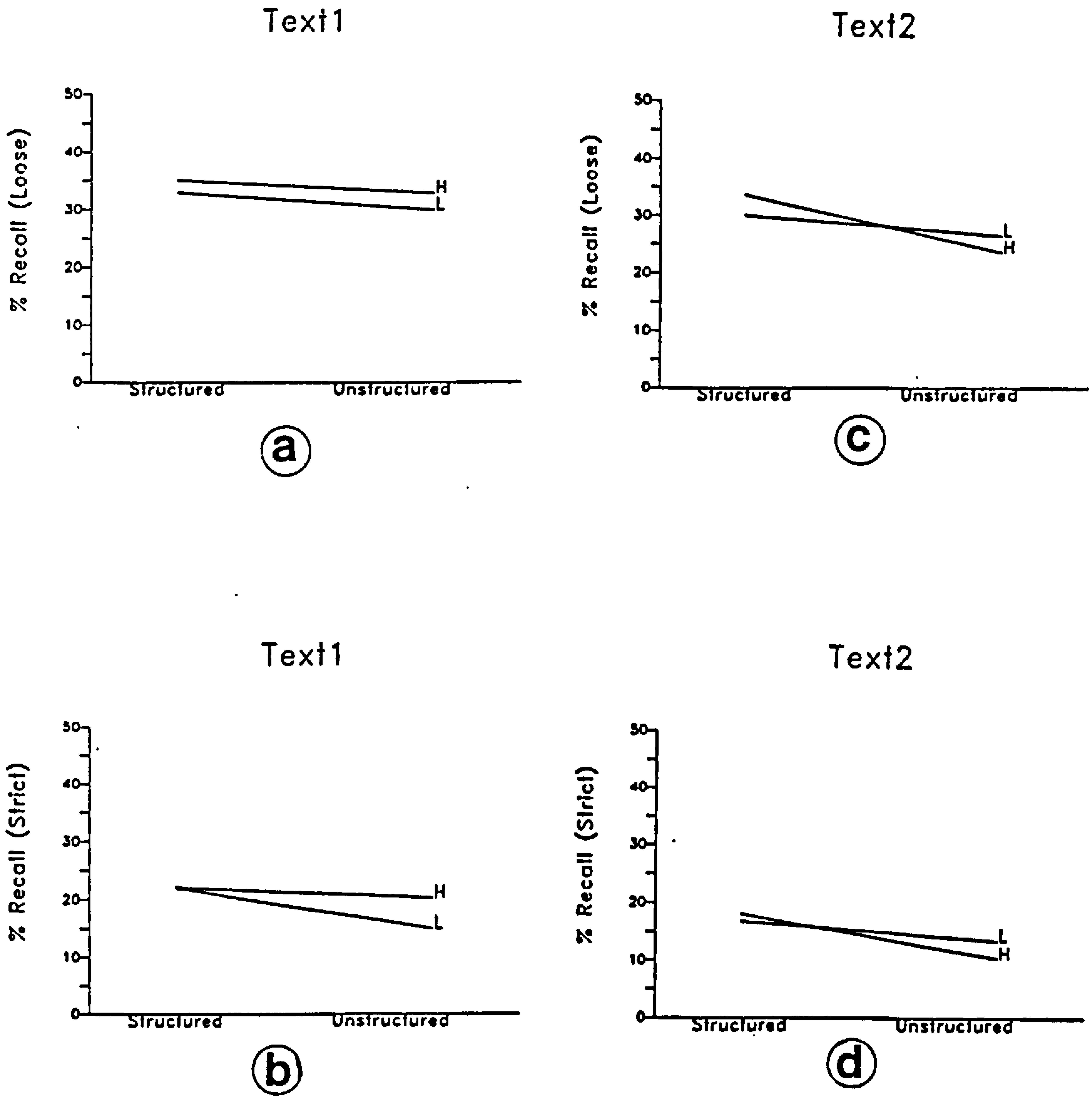

Figure 5.12. Percent recalled.

Table 5.24

% Recalled (Loose Criterion)				
	Text1		Text2	
	Structured	Unstructured	Structured	Unstructured
High Knowledge	35.10	32.79	33.55	23.46
Low Knowledge	32.99	29.92	29.97	26.37
F Table				
	MS	df	F	p
Structure MSE	462.20 110.39	1 62	4.19	0.046*
% Recalled (Strict Criterion)				
	Text1		Text2	
	Structured	Unstructured	Structured	Unstructured
High Knowledge	22.10	20.04	18.08	10.38
Low Knowledge	22.05	14.75	16.81	13.46
F Table				
	MS	df	F	p
Text Structure TxKxI MSE	527.30 519.85 166.86 49.13	1 1 2 62	10.73 10.58 3.40	0.002** 0.002** 0.042*

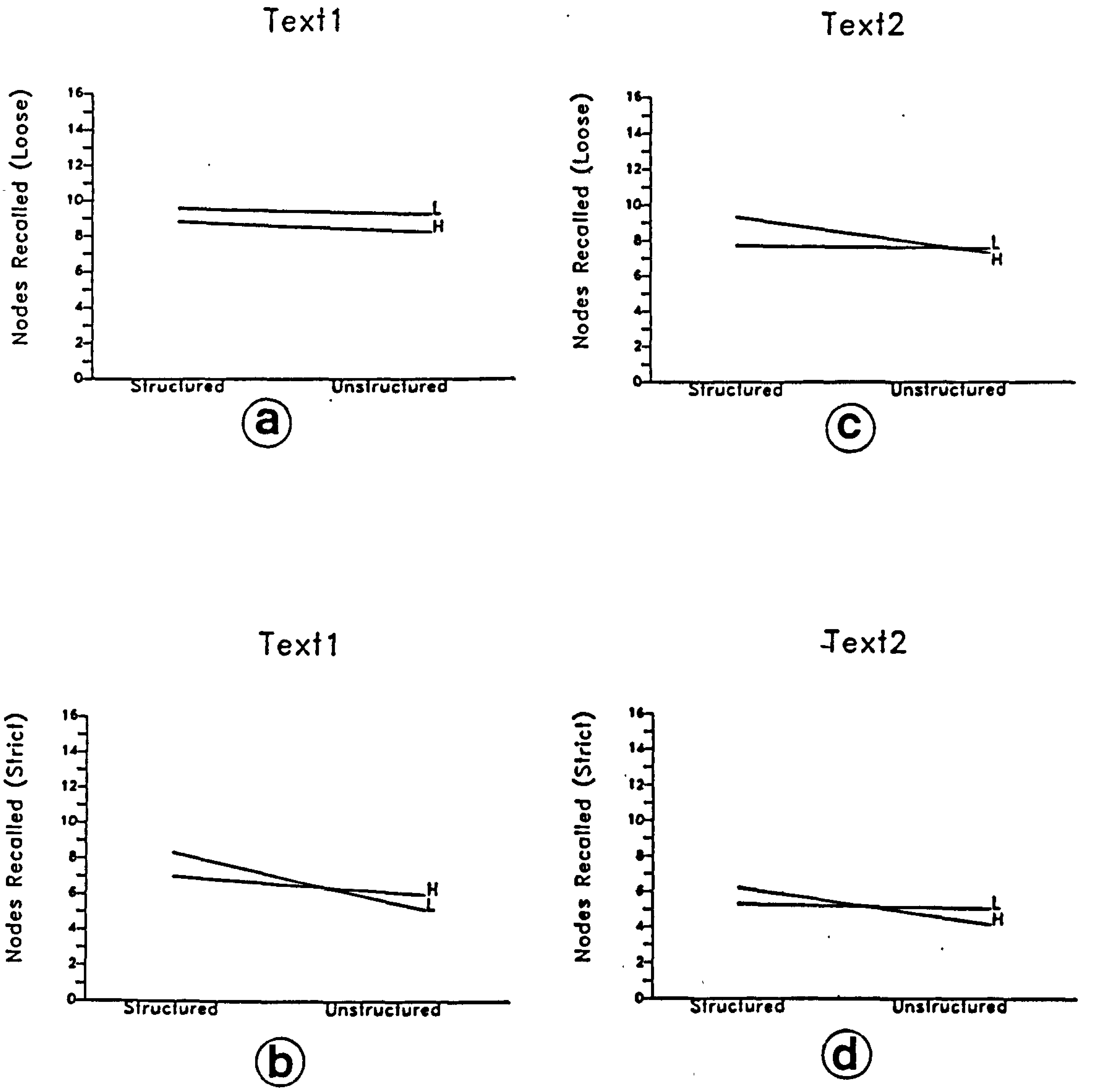

Figure 5.13. Average number of nodes recalled.

Table 5.25

Nodes Recalled (Loose Criterion)				
	Text1		Text2	
	Struct	Unstruct	Struct	Unstruct
High Knowledge	8.82	8.22	9.33	7.3
Low Knowledge	9.56	9.25	7.71	7.57
F Table				
	MS	df	F	p
Text MSE	15.61 3.29	1 62	4.74	0.033*
% Recalled (Strict Criterion)				
	Text1		Text2	
	Struct	Unstruct	Struct	Unstruct
High Knowledge	7.0	6.0	6.22	4.1
Low Knowledge	8.33	5.12	5.29	5.0
F Table				
	MS	df	F	p
Text Structure TxSxK MSE	42.15 55.16 17.83 3.82	1 1 2 62	11.04 14.45 4.67	0.002** 0.000** 0.035*

interaction between text, knowledge, and structure ($F(1, 62) = 4.669, p < .05$).

In order to examine recall patterns, the frequency with which sentences were recalled in each condition was expressed as a proportion (percentage) of the number of subjects. (See Figures 5.14a-d, 5.15a-d, 5.16a-d, and 5.17a-d.) Figures 5.18a-d, 5.19a-d, 5.20a-d, and 5.21a-d show the data expressed as mean frequencies per category. In order to examine the relationship between recall patterns in each condition, intercorrelations were performed between frequencies of recall for each

Table 5.26

Correlation Between % Recall (Loose)								
	HS1	LS1	HU1	LU1	HS2	LS2	HU2	LU2
HS1		0.79	0.60	0.46				
LS1	0.79		0.59	0.45				
HU1	0.60	0.59		0.72				
LU1	0.46	0.45	0.72					
HS2						0.60	0.66	0.44
LS2					0.60		0.47	0.51
HU2					0.66	0.47		0.64
LU2					0.44	0.51	0.64	

sentence in each text. Table 5.26 and 5.27 show the resulting correlation values (all significant at $p < .01$). The data show that patterns of recall for high and low-knowledge subjects with the social psychology text were more highly correlated than with the perception text. When structured and unstructured versions were compared, correlations were lower in general than for knowledge groups, but still higher for high-knowledge subjects than for low-knowledge subjects. Thus, there is some indication of restructuring occurring in recall for high-knowledge subjects, especially for the perception text, using the strict criterion. When knowledge and structure intercorrelations were examined, they were all low, except for low-

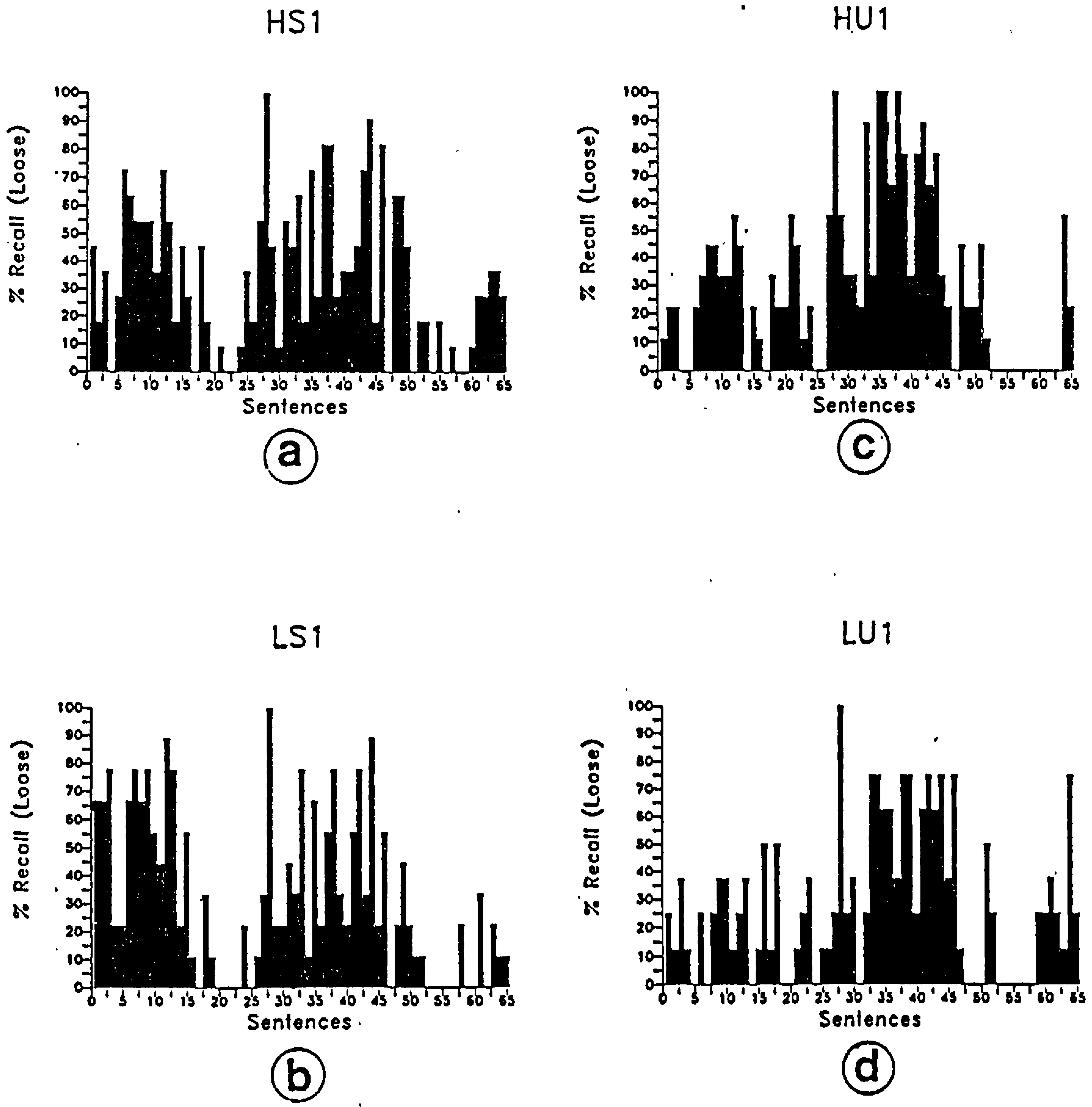

Figure 5.14. Proportion of subjects recalling sentences (Text1: Loose criterion).

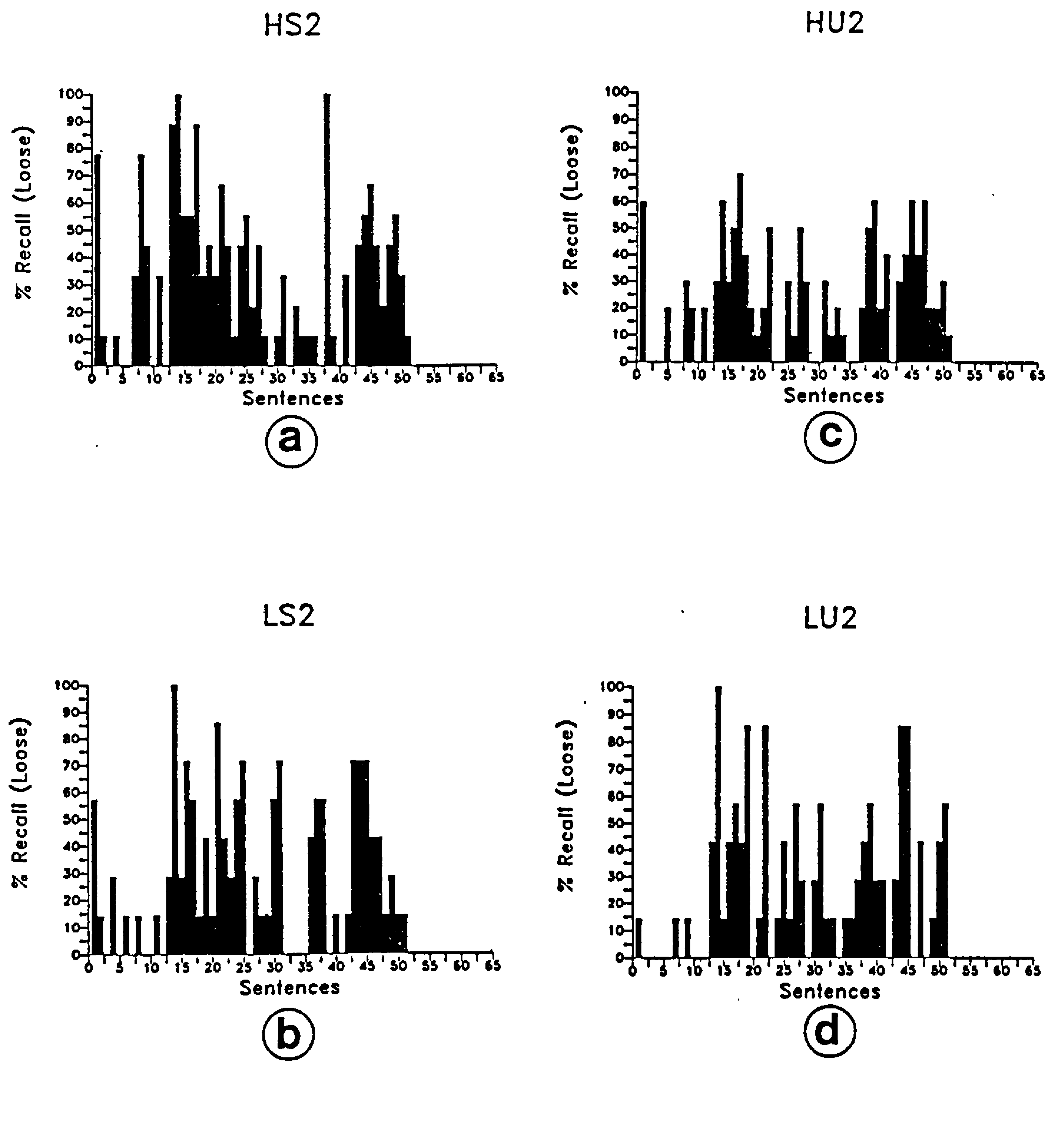

Figure 5.15. Proportion of subjects recalling sentences (Text2: Loose criterion).

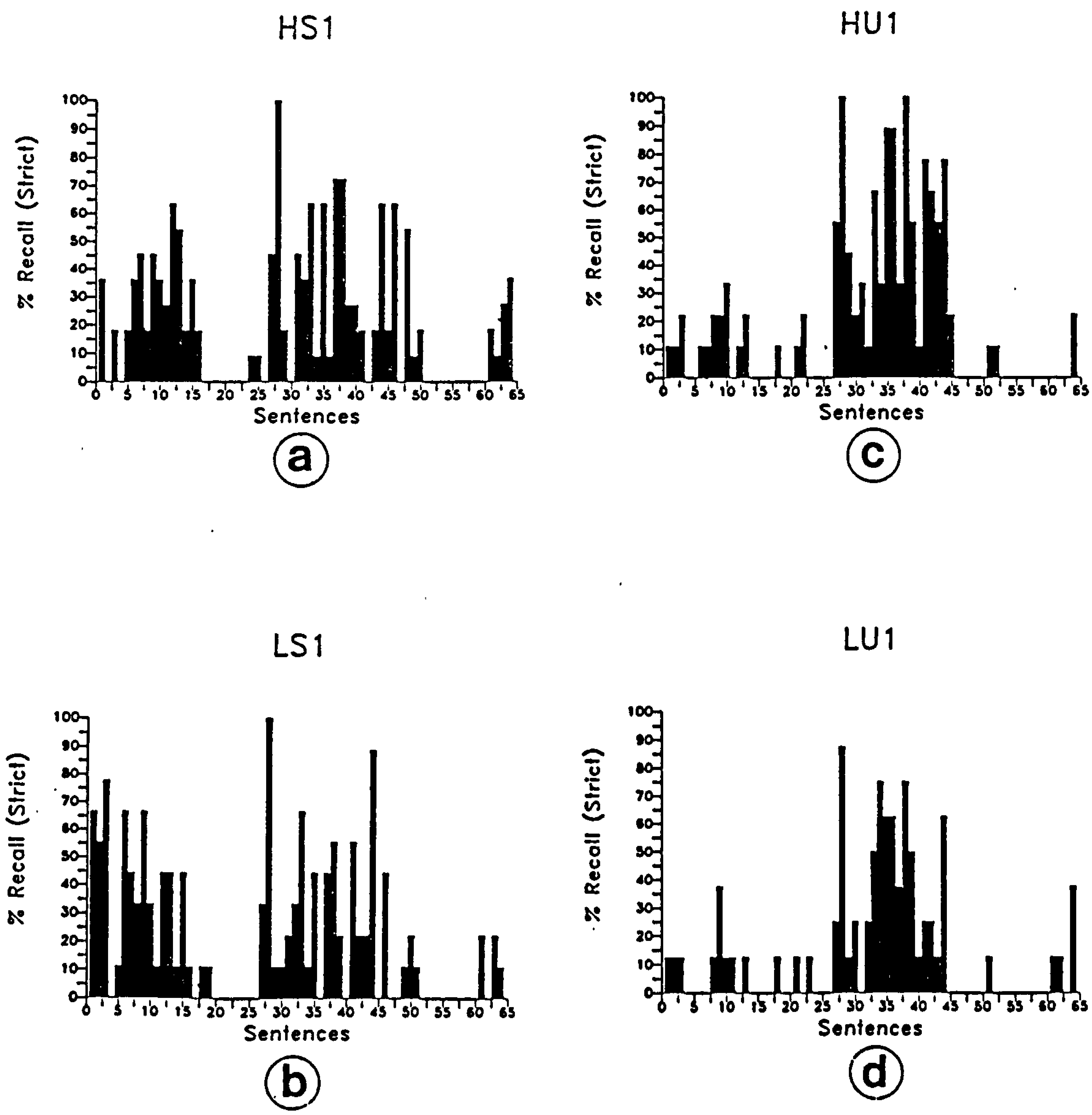

Figure 5.16. Proportion of subjects recalling sentences (Text1: Strict criterion).

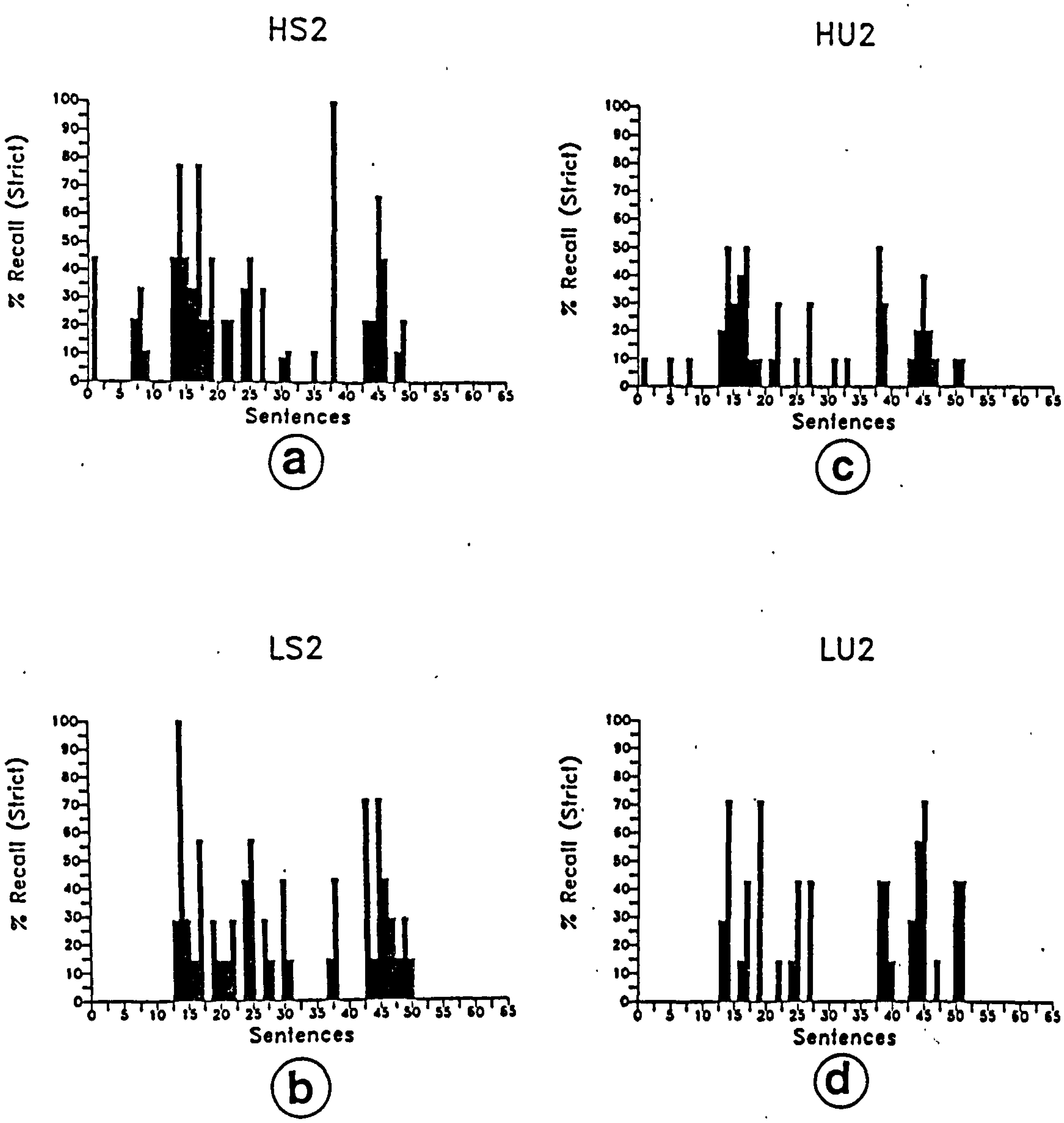

Figure 5.17. Proportion of subjects recalling sentences (Text2: Strict criterion).

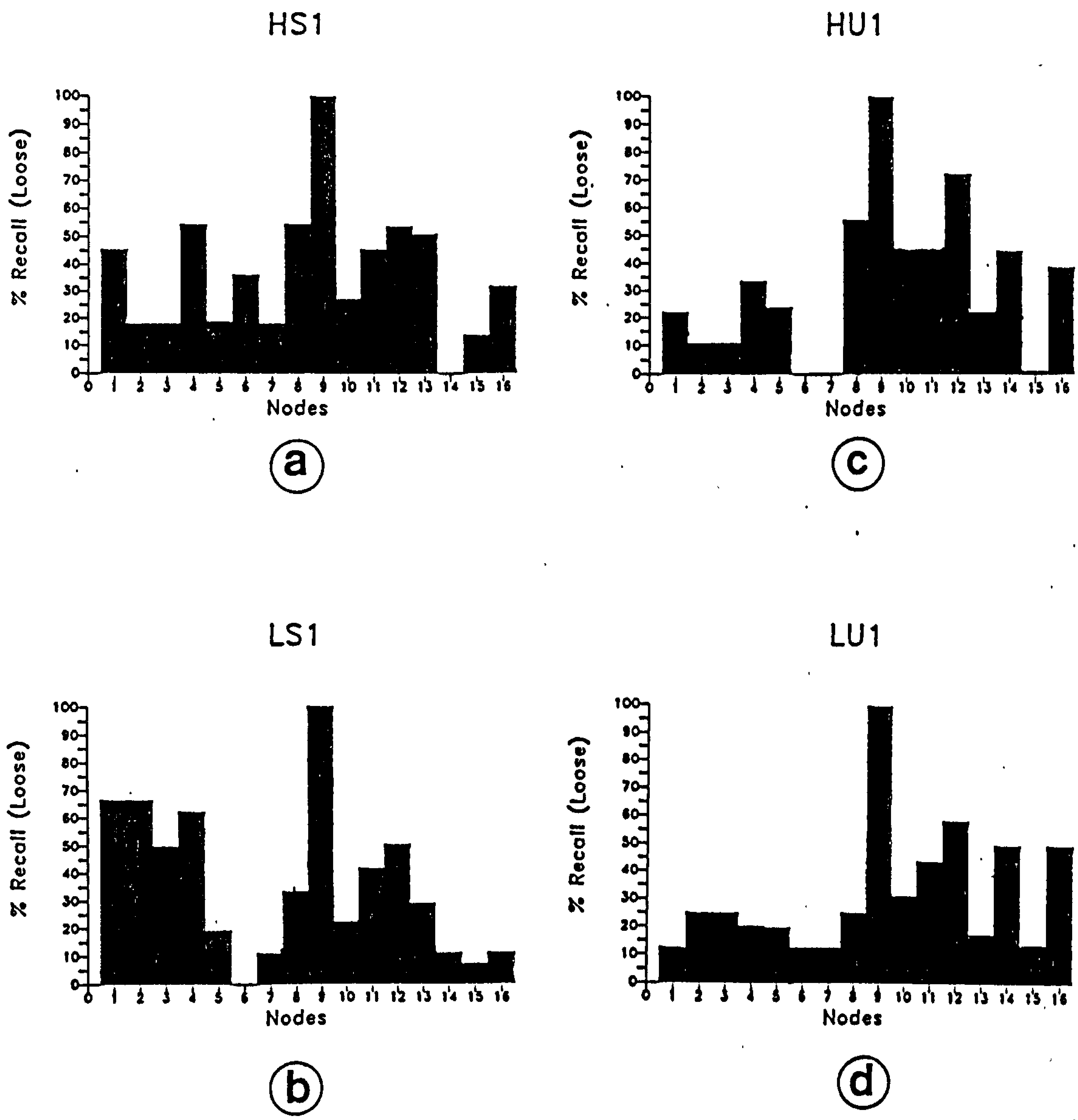

Figure 5.18. Frequency of recall per node (Text1: Loose criterion).

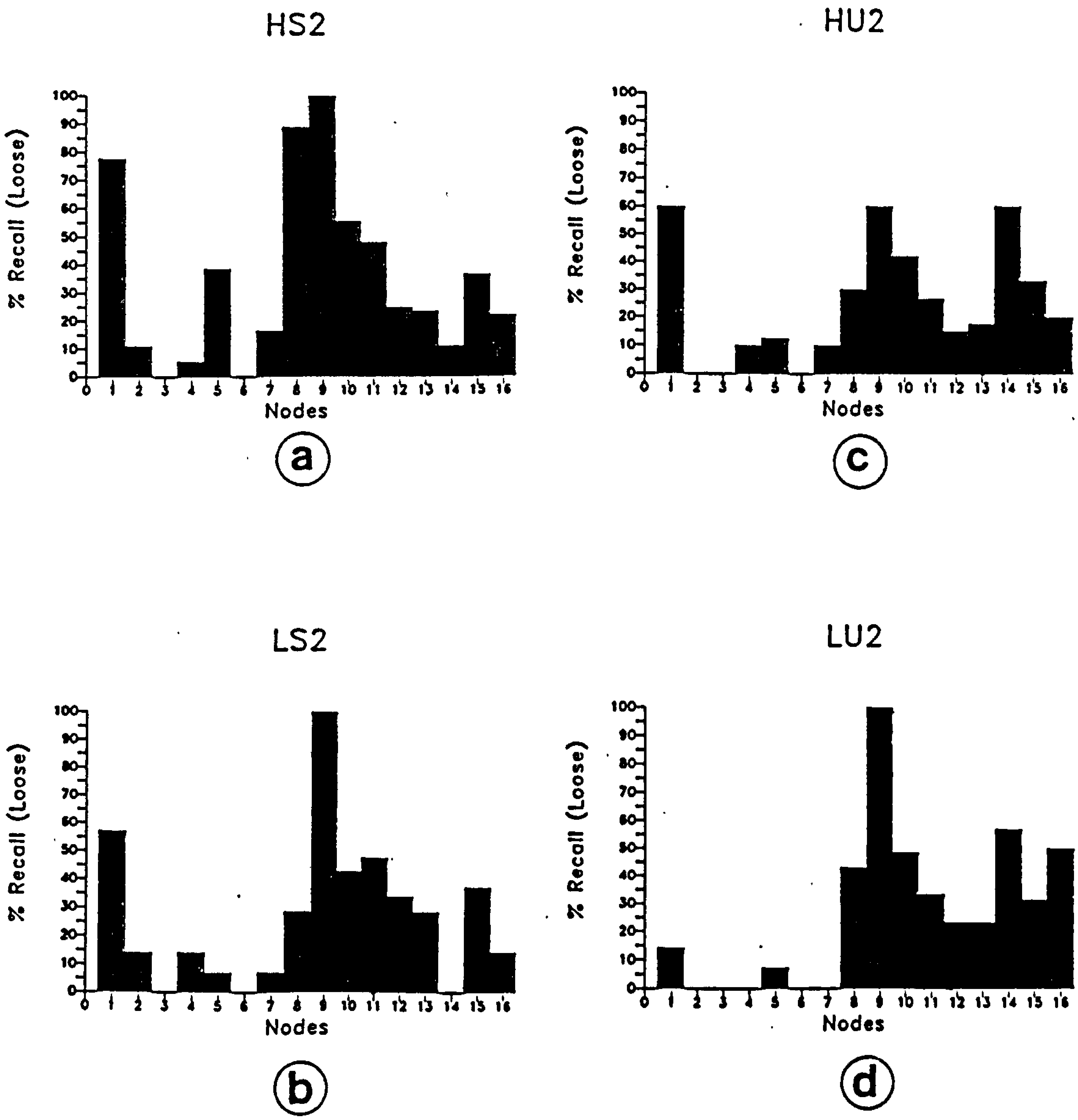

Figure 5.19. Frequency of recall per node (Text2: Loose criterion).

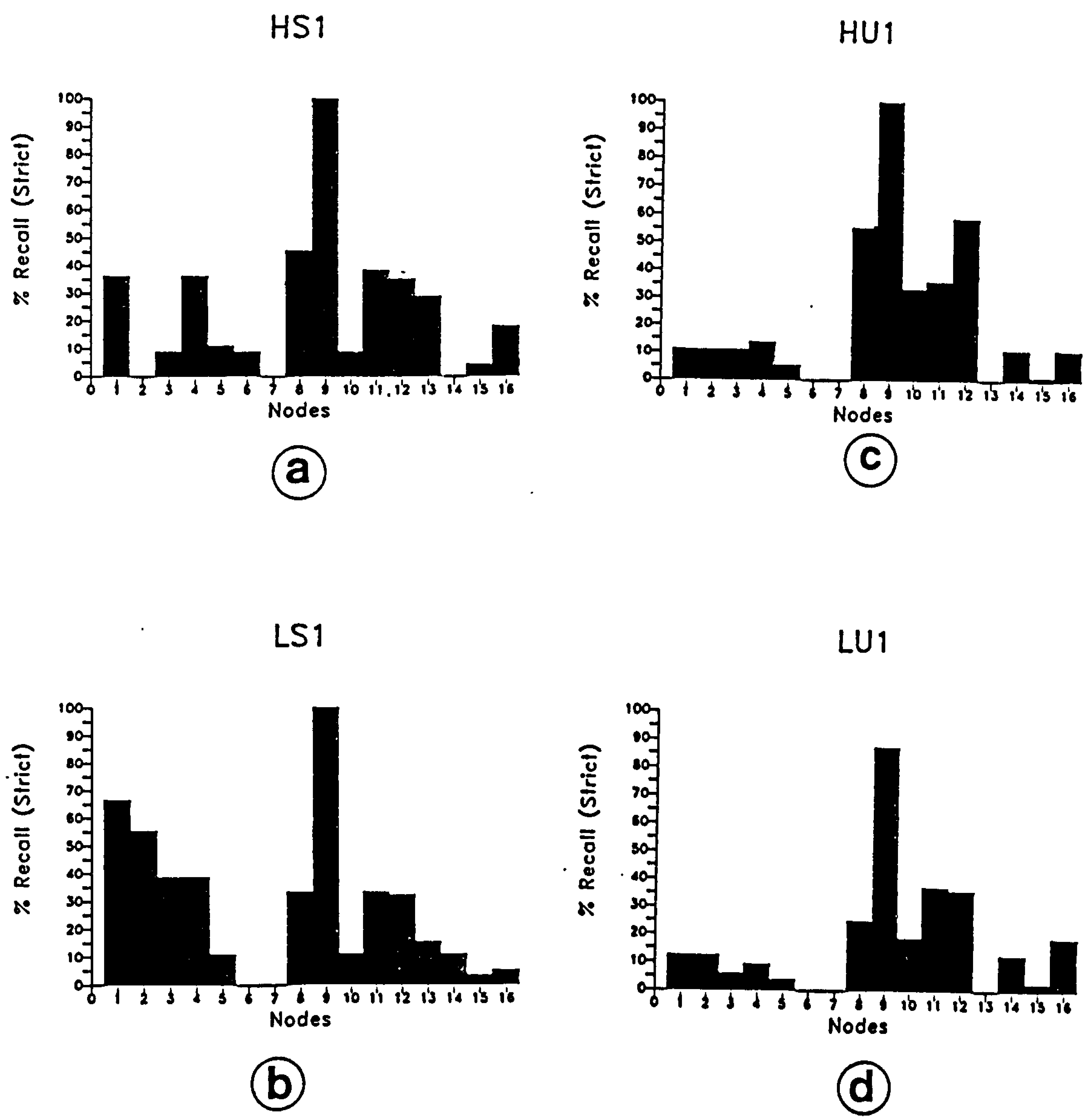

Figure 5.20. Frequency of recall per node (Text1: Strict criterion).

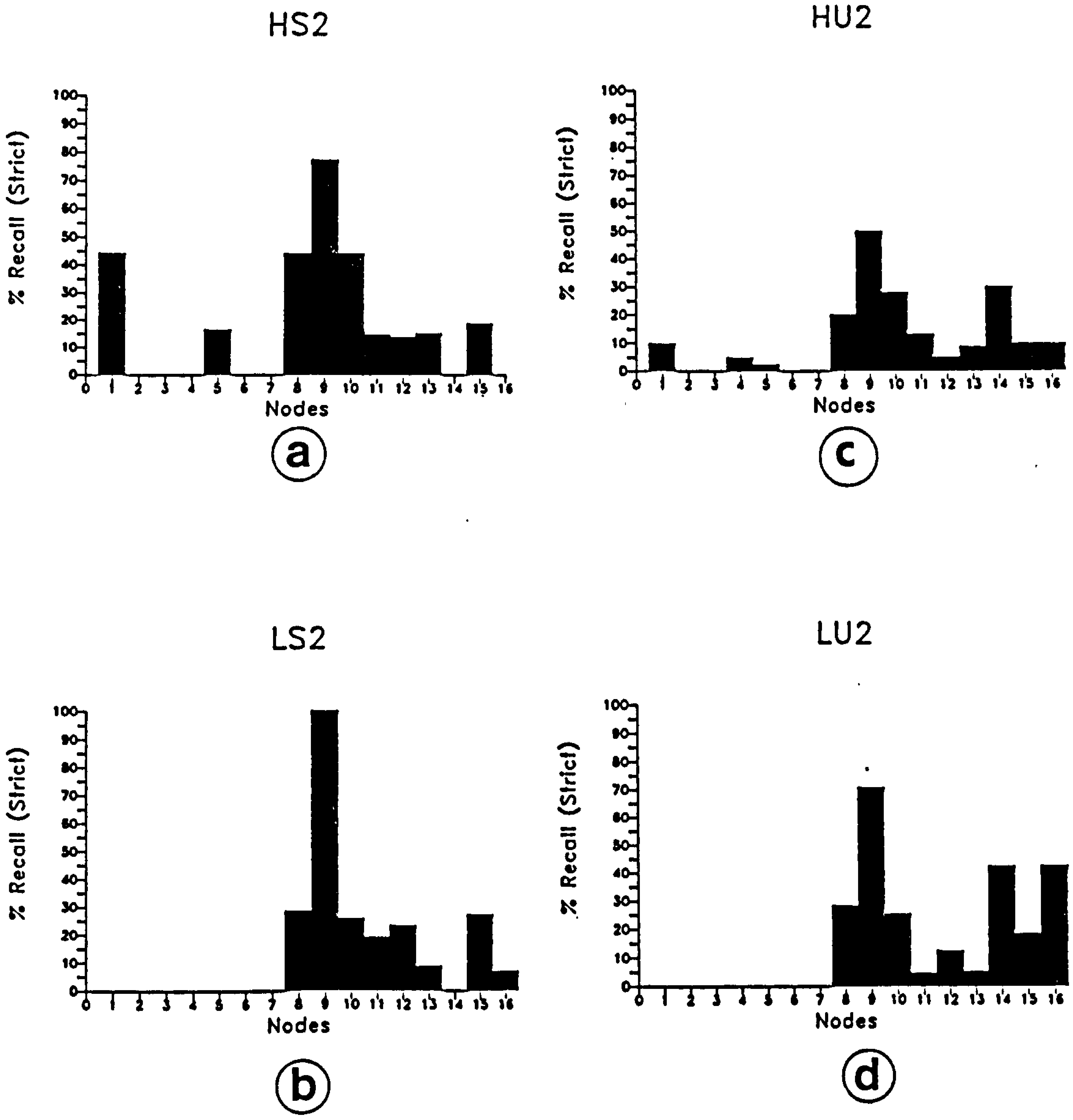

Figure 5.21. Frequency of recall per node (Text2: Strict criterion).

knowledge subjects with the structured text versus high-knowledge subjects with the unstructured text, at least for text1 (the social psychology text). Thus, there is some indication of restructuring for the high knowledge group with the social

Table 5.27

Correlation Between % Recall (Strict)								
	HS1	LS1	HU1	LU1	HS2	LS2	HU2	LU2
HS1		0.74	0.53	0.53				
LS1	0.74		0.54	0.51				
HU1	0.53	0.54		0.84				
LU1	0.53	0.51	0.84					
HS2						0.73	0.80	0.60
LS2					0.73		0.63	0.64
HU2					0.80	0.63		0.67
LU2					0.60	0.64	0.67	

psychology text.

One of the predictions of this study was that there would be differences between the high and low-knowledge groups with the unstructured versions in terms of amount of clustering of the sentences according to categories, and in terms of how much they adhered to the input order in their recall output. In other words, the high-knowledge group were expected to show some attempt at reordering the unstructured text in terms of the hypothesised superstructure organisation, which would be indicated by a lower clustering measure. Two measures were used in order to investigate this: the first was a clustering measure called *RRR* (the Relative Ratio of Repetition). Mandler (1978) used *RRR* in order to measure the amount of structure in free recall of stories. *RRR* measures clustering with respect to the designated categories, and is the ratio of the actual proportion of repetitions (total repetitions divided by number of sentences recalled minus one) over the maximum possible *RR* (number of sentences recalled, minus number of categories recalled, divided by the number of sentences recalled, minus one). This reduces to the number of repetitions, divided by the number of sentences recalled, minus the number of categories recalled. Repetition is defined as the number of times sentences are recalled in a single category, unbroken by sentences from other categories. The measure ranges from 0 to 1.

For example, if we have four categories:

birds
countries
cloths
musical instruments

and the following items are recalled:

wren
oriole
trumpet
piano
flute
Mexico
clarinet
Russia
England
Sweden
hawk
parrot
dacron
harp
silk
rayon

and since $RRR = [R/(N-1)]/[(N-Nc)/(N-1)] = R/(N-Nc)$,
where R = number of repetitions, where repetitions means the number of consecutive occurrences of items from the same category,
N = number of items recalled,
Nc = number of categories recalled.

Therefore, for this example, $R = 7/(16-4) = 0.583$.

Figures 5.22a-d show the amount of clustering as measured by *RRR* for each condition. Analysis of variance revealed no significant main effect or interactions.

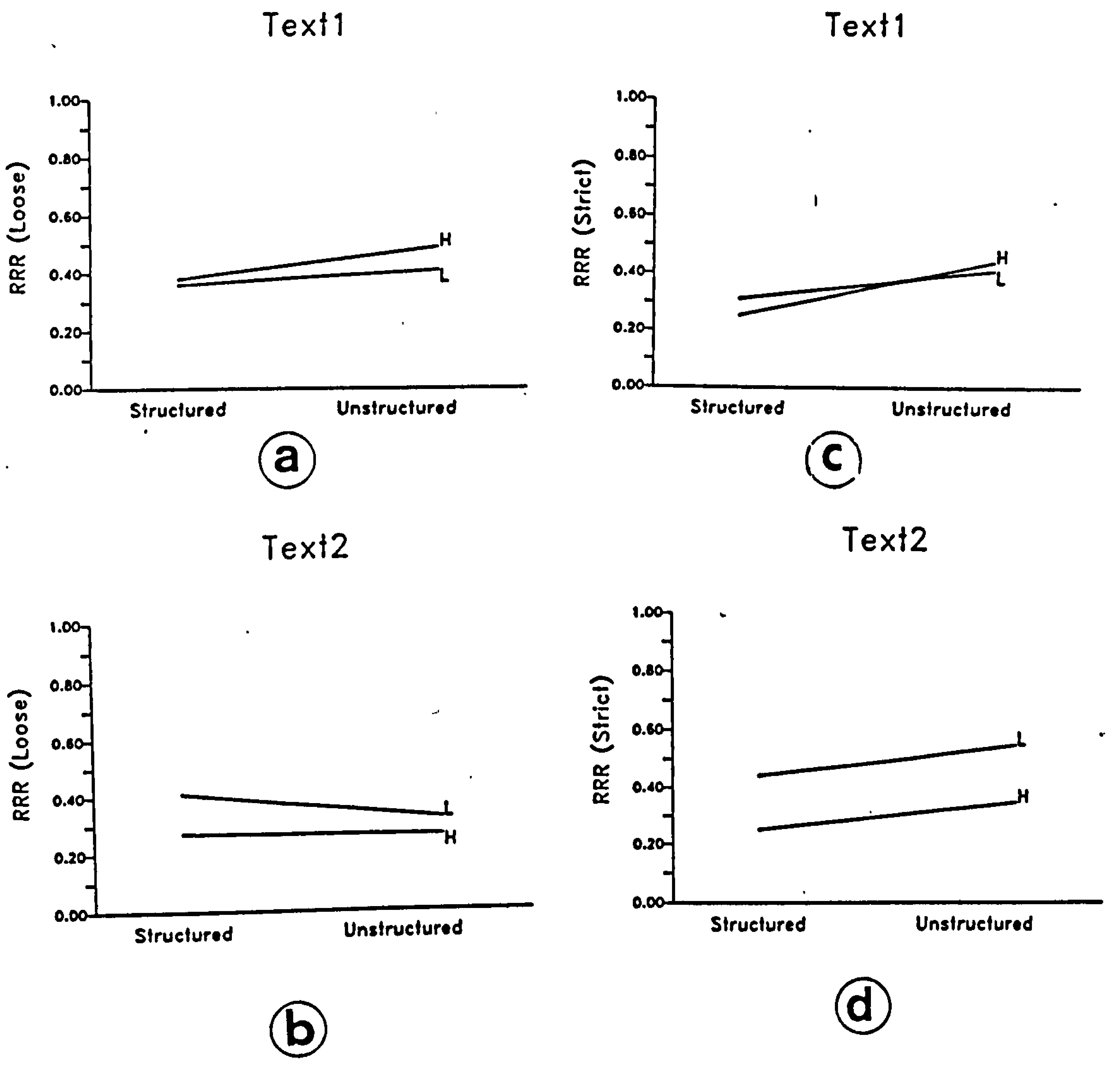

Figure 5.22: Clustering in recall, as measured by *RRR*.

(See Table 5.28 and 5.29.) The figures show that subjects with the unstructured versions have a higher tendency to stick to the input order in recall than those with the structured versions, and that low-knowledge subjects have a higher tendency to stick to input order than high-knowledge subjects. However, one problem with *RRR* is that there is a positive relationship between *RRR* and number of items recalled (cf. Mandler, 1978; Murphy, 1979), although *RRR* is less sensitive to quantity

Table 5.28

Clustering in Recall as Measured by RRR (Loose)				
	Text1		Text2	
	Structured	Unstructured	Structured	Unstructured
High Knowledge	0.38	0.49	0.27	0.26
Low Knowledge	0.36	0.41	0.41	0.32

of recall than many other measures.

Table 5.29

Clustering in Recall as Measured by RRR (Strict)				
	Text1		Text2	
	Structured	Unstructured	Structured	Unstructured
High Knowledge	0.25	0.44	0.25	0.35
Low Knowledge	0.31	0.41	0.44	0.55

Another measure of clustering or ordering tried was the *ITR2* (Bidirectional Inter-item Repetition). Ordering or subjective organisation refers to the invariance in response ordering that develops in a series of free recalls when order of presentation of the stimulus item is changed. The term *ITR* (inter trial repetition) is used to denote this sequential ordering effect. A unit of *ITR* may be said to occur whenever two items appearing consecutively in recall on trial t again appear in sequence on trial $t+1$. Mandler and Dean (1969) adapted the *ITR* measure in order to include bidirectional repetition: *ITR2*. They define *ITR2* as the ratio of observed repetitions to the maximum possible *ITR*s for the particular pair of events (e.g., input and output). The maximum *ITR* value is a function of the number of items common to both sets of events and does not depend on the absolute number of items recalled or presented. It is equal to the number of items common to both events minus one. Thus, the measure ranges from 0 to 1, where a value of 0 indicates no concordance, while a value of 1 represents a complete concordance of the two lists.

Figures 5.23a-d show the trends using this measure. Although there is some indication of less conformity to input order for the low knowledge group with the unstructured version of the perception text, analysis of variance revealed only a significant main effect for text (*Loose*: $F(1, 62) = 14.745$, $p < .01$; *Strict*: $F(1, 62) = 4.816$, $p < .05$), those with the social psychology text showing more tendency to

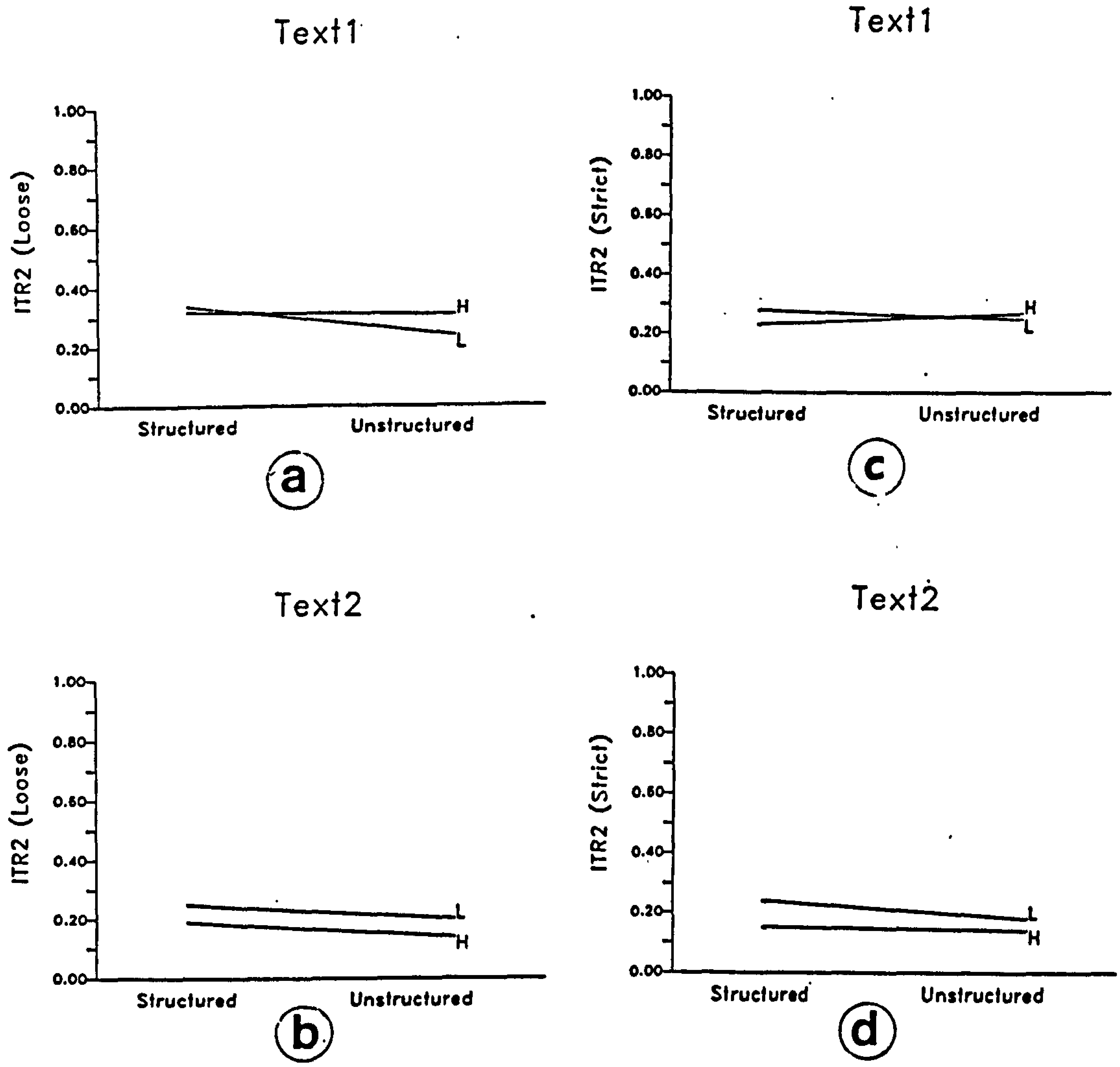

Figure 5.23. Clustering in recall, as measured by *ITR2*.

Table 5.30

Clustering in Recall as Measured by ITR2 (Loose)				
	Text1 Structured	Text1 Unstructured	Text2 Structured	Text2 Unstructured
High Knowledge	0.32	0.31	0.19	0.14
Low Knowledge	0.34	0.24	0.25	0.20
F table				
	df	*MS*	*F*	*p*
Text	1	0.23	14.75	0.00**
MSE	62	0.02		

cluster than those with the perception text. (See Tables 5.30 and 5.31.)

However, neither of the clustering measures appeared satisfactory for giving a full picture of the structuring in recall, since they aren't strictly measures of input-output order. Furthermore, they are slightly susceptible to omitted sentences, albeit not as much as most measures. Therefore, another version of *ITR* was developed, based on whether or not the sentences tended to be grouped in an ascending or descending series in recall. A trend towards an ascending sequence would indicate that recall followed the same pattern as input order.

Table 5.31

Clustering in Recall as Measured by ITR2 (Strict)				
	Text1		Text2	
	Structured	Unstructured	Structured	Unstructured
High Knowledge	0.23	0.27	0.15	0.14
Low Knowledge	0.28	0.25	0.24	0.18
F table				
	df	*MS*	*F*	*p*
Text	1	0.13	4.82	0.03*
MSE	62	0.03		

Analysis of variance revealed a significant main effect for text ($F(1, 62) = 9.838, p < .01$) with the loose criterion, subjects with the social psychology text following input order more than those with the perception text. There was also a significant main effect for structure (*Loose*: $F(1, 62) = 15.310, p < .01$; *Strict*: $F(1, 62) = 10.631, p < .01$), those with the structured version showing more adherence to input order than those with the unstructured version. However, although this effect appears greater for low-knowledge subjects, there was no significant structure by knowledge interaction. (See Figures 5.24a-d, and Tables 5.32 and 5.33.) The drop in *ITR* from around 0.7 to about 0.55 is indicative of some reordering in recall for those with the unstructured version.

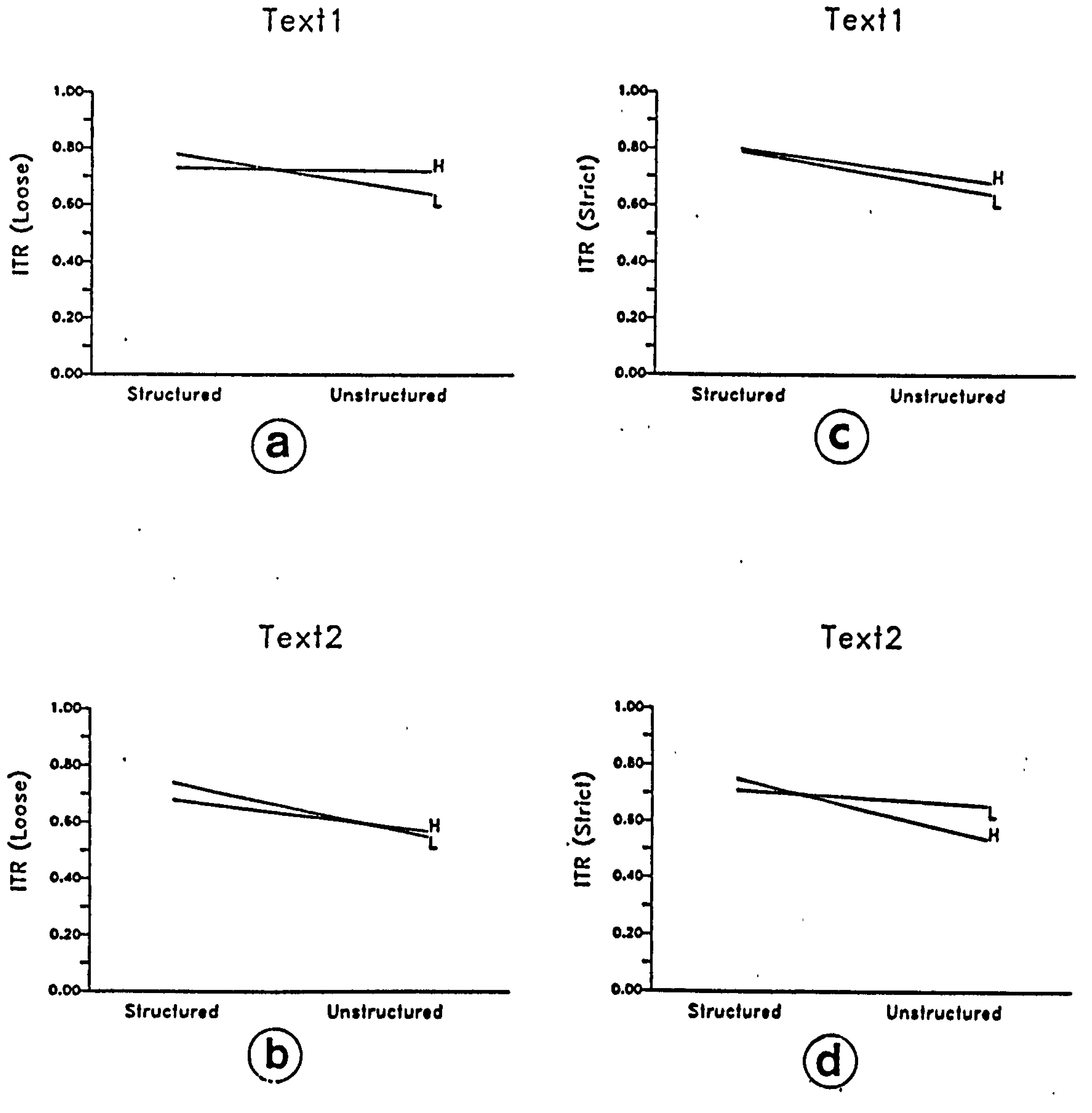

Figure 5.24. Ordering of sentences in recall.

Table 5.33

Ordering in Recall as Measured by ITR (Loose)				
	Text1		Text2	
	Structured	Unstructured	Structured	Unstructured
High Knowledge	0.73	0.72	0.68	0.57
Low Knowledge	0.78	0.64	0.74	0.55

F table

	df	*MS*	*F*	*p*
Text	1	0.14	9.84	0.00**
Structure	1	0.21	15.31	0.00**
MSE	62	0.01		

Table 5.33

Ordering in Recall as Measured by ITR (Strict)				
	Text1		Text2	
	Structured	Unstructured	Structured	Unstructured
High Knowledge	0.80	0.68	0.75	0.54
Low Knowledge	0.79	0.64	0.71	0.66
F table				
	df	*MS*	*F*	*p*
Structure MSE	1 62	0.35 0.03	10.63	0.00**

However, since even this measure may not give the complete picture, in order to reveal any restructuring effects in the ordering of categories, the same analyses as before were performed, but with the categories, rather than the sentences, and based on category recall scores, i.e., the first occurrence of the category in the recall protocol. (See Figures 5.25a-d, showing the ordering of categories in recall.) The figures show that subjects with the unstructured version did not follow the input order of categories, but tended to reverse the order slightly. Analysis of variance revealed a main effect for text (*Loose*: $F(1, 62) = 4.07$, $p < .05$; *Strict*: $F(1, 62) = 6.351$, $p < .05$), those with the social psychology text following input order more than those with the perception text. There was also a significant main effect of structure (*Loose*: $F(1, 62) = 73.301$, $p < .01$; *Strict*: $F(1, 62) = 48.408$, $p < .01$) and, at least for the strict criterion, a significant main effect for knowledge ($F(1, 62) = 4.752$, $p < .05$). However, there were no significant interactions. (See Tables 5.34 and 5.35.)

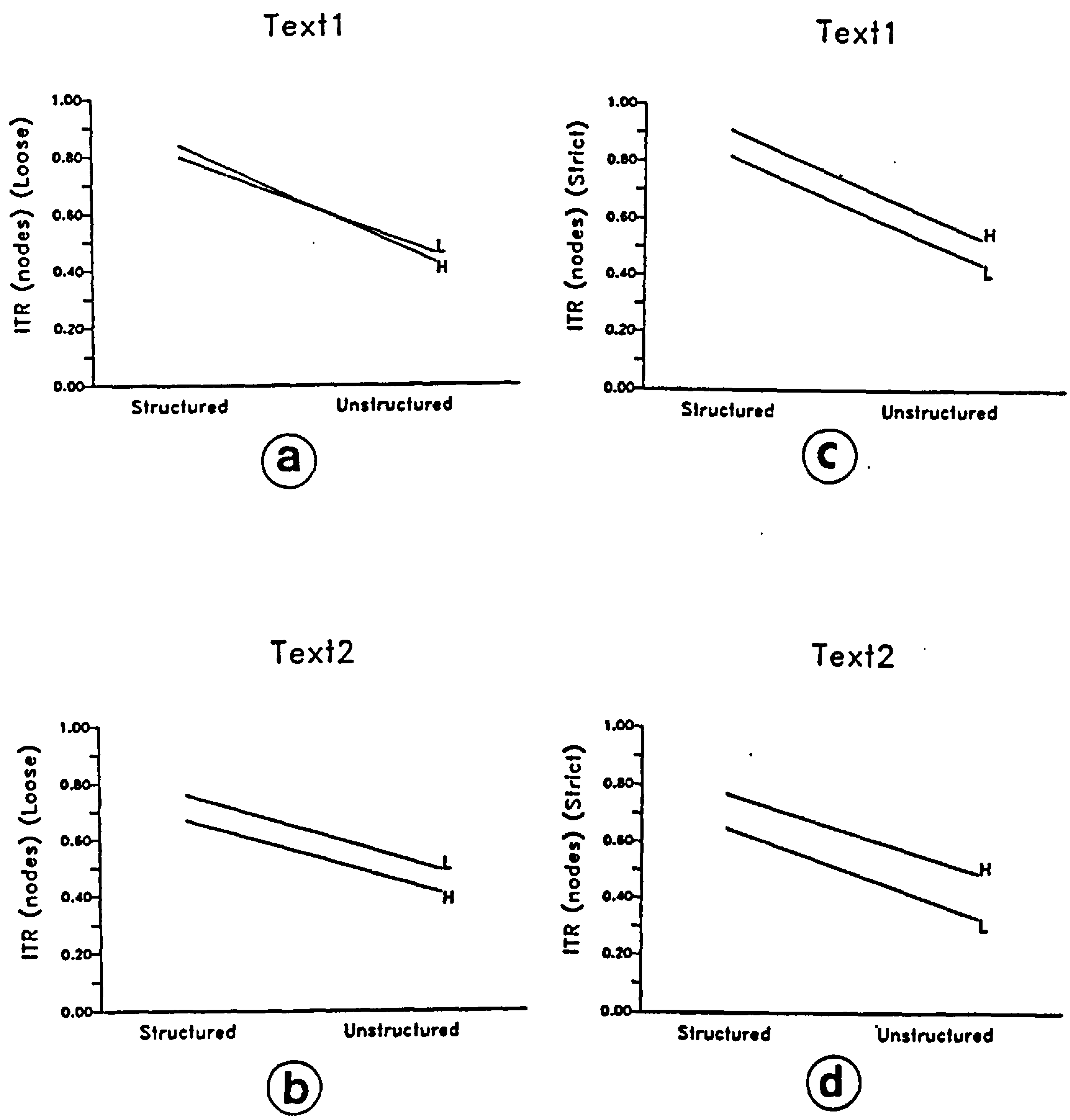

Figure 5.25. Ordering of categories in recall.

Table 5.34

Ordering of Categories in Recall (Loose)				
	Text1		Text2	
	Structured	Unstructured	Structured	Unstructured
High Knowledge	0.84	0.43	0.67	0.41
Low Knowledge	0.80	0.46	0.76	0.49
F table				
	df	*MS*	*F*	*p*
Text	1	0.11	4.07	0.048*
Structure	1	1.89	73.30	0.000**
MSE	62	0.03		

Table 5.35

Ordering of Categories in Recall (Strict)				
	Text1		Text2	
	Structured	Unstructured	Structured	Unstructured
High Knowledge	0.91	0.53	0.77	0.49
Low Knowledge	0.82	0.44	0.65	0.33
F table				
	df	*MS*	*F*	*p*
Text	1	0.28	6.35	0.014*
Structure	1	2.14	48.41	0.000**
Knowledge	1	0.21	4.75	0.033*
MSE	62	0.04		

When the analysis was run again to include instructional condition, however, there was a significant knowledge by instruction interaction (*Loose*: $F(1, 46) = 3.580, p < .05$). The means for this are shown in Tables 5.36 and 5.37. This shows that the tendency to stick to input order was less for high-knowledge subjects in instructional conditions B and N. However, it was greater than for low-knowledge subjects in condition A. There was no significant structure by knowledge by instruction interaction, although there was a significant interaction between text, instructions, and knowledge (*Strict*: $F(1, 46) = 5.845, p < .01$), and between text structure, knowledge and instructions (*Strict*: $F(1, 46) = 4.353, p < .05$).

Thus, there is some evidence that subjects with the unstructured text were reordering categories in recall. However, the only case where the expected interaction seemed to appear was not significant. Thus the effect seems to be the same for high and low-knowledge subjects.

Table 5.35

Ordering of Categories (Loose Criterion)						
		Text1		Text2		
		Structured	Unstructured	Structured	Unstructured	
High Knowledge	A	0.94	0.54	0.69	0.48	
	B	0.71	0.43	0.67	0.38	
	N	0.90	0.32	0.65	0.39	
Low Knowledge	A	0.78	0.43	0.63	0.44	
	B	0.81	0.44	0.83	0.64	
	N	0.83	0.53	0.79	0.33	

F table

	df	*MS*	*F*	*p*
Structure	1	1.89	72.48	0.00**
I x K	2	0.09	3.58	0.04*
MSE	46	0.03		

The reordering of the texts was done between categories, and not within categories. Thus, we would expect reordering to occur to a greater extent in the former case. A comparison of Figures 5.24a-d and 5.a25-d show the expected difference between- versus within-categories, since there is a bigger drop in input-output order for categories than for sentences, at least with the loose criterion, although not for the strict criterion. Analysis of variance, including the factor of ordering within- versus between-categories, showed main effects for text (*Loose*: $F(1, 62) = 11.493, p < .01$; *Strict*: $F(1, 62) = 12.539, p < .01$), the effect for the social psychology text being greater than for the perception text. There was also a main effect for structure (*Loose*: $F(1, 62) = 80.350, p < .01$; *Strict*: $F(1, 62) = 77.551, p < .01$), the extent to which input-output order was maintained being higher for the structured text. There was a main effect for whether or not ordering was within or between categories (*Loose*: $F(1, 62) = 7.829, p < .01$). That is, there was more of a tendency to stick to input order within categories than between categories. There was also an interaction between structure and within- versus between-ordering (*Loose*: $F(1, 62) =$

Table 5.37

Ordering of Categories (Strict Criterion)					
		Text1		Text2	
		Structured	Unstructured	Structured	Unstructured
High Knowledge	A	0.94	0.55	0.72	0.67
	B	0.83	0.69	0.77	0.40
	N	1.00	0.35	0.83	0.44
Low Knowledge	A	0.80	0.44	0.58	0.13
	B	0.82	0.24	0.72	0.59
	N	0.86	0.73	0.63	0.14

F table

	df	*MS*	*F*	*p*
Text	1	0.28	7.70	0.008**
Structure	1	2.14	58.67	0.000**
Knowledge	1	0.21	5.76	0.02*
T x I x K	2	0.21	5.85	0.005**
T x I x S x K	2	0.16	4.35	0.019*
MSE	46	0.04		

22.159, $p < .01$; *Strict*: $F(1, 62) = 7.499$, $p < .01$). This tendency to stick to input order between categories was much lower for the group with the unstructured version, but mainly for the social psychology text, the three-way interaction between text, structure, and within- versus between-ordering just reaching significance (*Loose*: $F(1, 62) = 4.335$, $p < .05$).

In summary, the results of the recall analysis show that reordering the text significantly reduces the number of sentences recalled, the number of categories recalled, and the ordering of the sentences and categories in recall. The order results show that subjects with the restructured text did not follow the input order to the same extent as the subjects with the structured text, thus providing some evidence for restructuring the text according to the hypothesised superstructure.

DISCUSSION OF RESULTS

In general, the results of this study have shown that subjects familiar with the structure of psychology research reports found the texts easier to comprehend, as evidenced by faster and less variable reading times than those unfamiliar with the structure of such texts.[5] The topic manipulation (that is, the perception text versus the social psychology text) made a difference to ease of comprehension, by raising reading times for both groups, although high-knowledge subjects still found the perception text easier to read than low-knowledge subjects. The effect of restructuring the text was to increase reading times for both groups, but high-knowledge subjects still found the text easier to read than low-knowledge subjects with either version (structured or unstructured). The low-knowledge group with the unstructured version was worse off in general, but the fact that there was no effect of structure on high-knowledge subjects with the more difficult text (the perception text), and a much greater effect for low-knowledge subjects, indicates that familiarity with the topic was not the factor responsible for the effect. It suggests that familiarity with the structure is more likely to be responsible. This is also substantiated by the fact that, when patterns of reading times within categories were examined, they were found to be more highly correlated than patterns of reading times that included category boundaries. In other words, since restructuring the text only involved categories, and not sentences within categories, the lower correlations for comparisons including category boundaries confirms that the effects of this manipulation are due to text structure, and not to local comprehensibility.

Evidence for the use of structure, as indicated by boundary effects (i.e., higher reading times at boundaries) did not appear, except for subjects with the unstructured version of the social psychology text. One could account for this by arguing that the high-knowledge subjects were expecting a different ordering of categories, and that the higher reading times are due to violations of these expectations. However, the effect was also similar for low-knowledge subjects. This may be due to some of the low-knowledge group having instructions before reading, as indicated by a boundary effect when this condition was examined. However, the effect also appeared for low-knowledge subjects who had received no

5. This is not to imply that the relationship between reading time and comprehension is a simple one; increases in reading time can only provide some indirect pointers to increased processing load.

instructions before reading. Since the effects that were found for these subjects were not very systematic, and since they had a much higher variability than high knowledge subjects in terms of overall reading times, the category boundary effect here is difficult to interpret.

However, as Van Dijk (1980a) points out, whether or not one predicts an increase in reading time as the topic (global or local) changes, depends upon whether one believes the process of macrostructure formation to be serial or parallel. If one adheres to the serial processing hypothesis, then it would be predicted that the transition across boundaries would be slow because of time taken up with the formation of maropropositions, or instantiation of the superstructural schema. However, a parallel processing hypothesis would predict no difference. That is, if one postulates a process in which macrostructure formation proceeds in parallel with microstructure formation, shifts in topic should make little difference to reading times.

The fact that reading times were much more variable for low-knowledge subjects and for those with the unstructured versions, is also indicative of less ease of comprehension. These patterns held even when factors involving the length of sentences and their serial position were accounted for. However, several other factors not controlled for here may have also contributed to this variability, including content word frequency (Mitchell & Green, 1978); referential overlap; imagery value of sentences (Kieras, 1974); presence of new argument nouns (Kintsch, Kozminsky, Streby, McKoon & Keenan, 1975; Kieras, 1978; Cirilo & Foss, 1980; Graesser, Hoffman & Clark, 1980); the nature of the referential relations, and the directness of antecedence (Haviland & Clark, 1974); and the importance of the sentence (Cirilo & Foss, 1980). However, since there was far less evidence of variability in the high-knowledge groups, it seems less likely that one can attribute these differences *solely* to such text variables.

Although high-knowledge subjects had faster reading times than low-knowledge subjects, there were no significant differences in the amount of material recalled for these groups, except with the unstructured version, in which case high-knowledge subjects recalled more than low-knowledge subjects. There is, of course, a tradeoff between reading times and recall. Although, on the one hand, higher reading times may indicate less ease of comprehension (or more difficulty with the text) they may also indicate more effort being expended in processing, leading perhaps to deeper encoding. Thus, it is hard to predict recall from reading times alone. Higher reading times may be due to difficulty with the material, leading to a prediction of low recall; on the other hand, deeper encoding should lead to higher recall. In this case, it is possible that the longer reading times for low-knowledge subjects enabled them to make up for lack of familiarity with the material (content as well as structure) by deeper encoding. Thus, they were able to recall as much as high-knowledge subjects, who spent less time on reading the text.

Although there were no significant differences in reading times between structured and unstructured versions (although weak evidence that low-knowledge groups with the unstructured version had longer reading times), the group with the structured version were able to recall more than those with the unstructured versions. Moreover, the high-knowledge subjects with the unstructured version recalled more than the low-knowledge subjects with the unstructured version. Although low-knowledge subjects with the unstructured version of the perception text had the highest, or longest reading times, they recalled far less than any of the other groups. Thus, the longer reading times and the lower rates of recall for low knowledge subjects indicate difficulty of comprehension. This pattern held for both texts, although effects were greater for the perception text. Although the high-knowledge groups had longer reading times, therefore possibly more difficulty, and recalled less with the perception text, the differences were greater for low-knowledge subjects (although this text by knowledge interaction was not significant). Thus, the pattern of means suggests that the effects are due to familiarity with the structure, rather than with the topic or content.

The major source of evidence for the use of knowledge of text structure is to be found with the pattern of recall, rather than the amount recalled. Thus, it was expected that high-knowledge subjects with the restructured version should be able to reorder the text at recall, while low-knowledge subjects would conform much more to input order, with both the structured and the unstructured versions. Moreover, low-knowledge subjects who were given instructions concerning the structure of the text should show some evidence of attempting to reorganise the unstructured version.

In general, the pattern of recall was similar for high and low-knowledge subjects with structured versions, and for high and low-knowledge subjects with the unstructured versions. When the same sentences were compared for structured and unstructured versions, the patterns for the low-knowledge subjects were shown to be quite different, whereas the patterns for the high-knowledge subjects were more highly correlated. This provides preliminary evidence for some ability to restructure the text on the part of the high-knowledge subjects. In fact, the intercorrelations showed that high-knowledge subjects with the unstructured version of the social psychology text had similar recall patterns to low-knowledge subjects with the structured version. Thus, the ability of high-knowledge subjects to restructure the text seems to be greater, at least with the social psychology text.

The clustering and input-output ordering measures indicate that those with the unstructured versions had less tendency to stick with input order, especially the order of categories, thus indicating some restructuring. There was also some evidence that high-knowledge subjects had less tendency to stick to input order than low-knowledge subjects, although this knowledge by structure interaction failed to reach significance. The fact that the tendency to stick to input order was higher when sentences, rather than categories, were examined suggests that the effect is genuinely to do with restructuring the text as a whole, rather than any local restructuring. The fact that low-knowledge subjects also showed some tendency to restructure the text, although not as much as high-knowledge subjects, suggests that instructions may have had an effect. As the results indicate, there was a significant interaction between knowledge and instructions, the high-knowledge group with instructions given after reading showing more tendency to stick with input order than the low-knowledge subjects. This pattern was reversed for instructions given before reading. This is some indication that restructuring effects may be encoding specific — that is, subjects may have more difficulty restructuring the text if they tried to encode it with the wrong structure. Although there was no significant interaction between structure, knowledge, and instructions, the fact that high-knowledge subjects given no instructions also showed this pattern gives some indication that it is an encoding strategy.

The size of the decrements in performance between structured and unstructured versions was higher for high-knowledge subjects than for low knowledge subjects in terms of amount recalled, as well as in terms of ordering, although in general high-knowledge subjects recalled more than low-knowledge subjects. Furthermore, recall for portions of the text such as the discussion category, presented very early on in the unstructured version, was much lower for high-knowledge subjects, indicating that violations of their expectations led to a greater decrement in performance than for low-knowledge subjects. This indicates, once again, that familiarity with the structure, rather than the content, is leading to these effects.

GENERAL DISCUSSION

In general, the effects found here were not as great as those found in similar studies with narratives (e.g., Haberlandt, 1980; Mandler & Goodman, 1982). First, it may be argued that this study used particularly difficult texts, which may have damped the effects of structure and confounded familiarity of structure with text difficulty. Reading times were rather high in general, and recall was rather low (thus leading to some difficulties in obtaining higher ordering effects with the measures chosen). However, it has been noted that expository texts are in general read slower than, and recalled less well than narrative texts (e.g., Thorndyke, 1977; Graesser, Hauft-Smith, Cohen & Pyles, 1980; Graesser, Hoffman & Clark, 1980). An apparently conflicting finding is that more inferences are made in narrative than in expository texts. Reiser and Black (1982), and Graesser and Goodman (1985) have suggested that this apparent paradox may be due to differences in processing. They argue that the reader of an expository text attempts to build or expand knowledge structures, while processing in narratives consists of instantiating existing knowledge structures and constructing causal links between them. More inferences are made for narratives since there is a more complete existing structure to instantiate — that is, a content structure based on world knowledge as well as a superstructural framework. This uses more cognitive capacity, leading to better depth of processing and therefore better recall. However, since the instantiation process is easier than the construction process involved in expository texts, narratives texts are read faster.

Although this argument may account for less effects with the present materials, the findings reported here still suggest that, when readers are unfamiliar with the content, they rely more on their knowledge of the text structure to help them recall it, as evidenced by a greater effects of structure with the perception text for high-knowledge subjects. Furthermore, the results suggest that if readers are unfamiliar with both content and structure, restructuring the text makes less difference to performance, as indicated by less of a decrement in performance for the low-knowledge group. Even though their performance was low overall, in terms of amount recalled and ability to reorder the text, the effect of restructuring was greater for high-knowledge subjects. Moreover, there is some evidence that this is due to an encoding strategy adopted by high-knowledge subjects, since there subjects did worse in terms of reordering the text when instructions were given after, rather than before reading (i.e., encoding). Evidence for the use of structure as an encoding strategy for expository texts has also been provided by, for example, Meyer and Rice (1982).

Finally, another caveat concerning the relative weakness of the familiarity effects found here relates to the use of undergraduates in text comprehension studies. As Johnson and Kieras (1983) point out, it could be that such students are practised at using powerful elaboration strategies for memorising information that is not very familiar. Although the subjects in the study reported here were not told before reading that they would have to recall the text, they may still have expected to have to do so.

CONCLUSIONS

This study has provided evidence that familiarity with the structure of scientific reports enables readers to comprehend and recall texts even when the content of the material is unfamiliar to them. The study thus supports the theory that certain text types have a schematic structure that serves to organise the content of the material, and lends further support to the arguments of proponents of story grammars, that readers are able to use the structure of the text in comprehension and recall, rather than relying solely on general world knowledge. One can conclude from studies such as the one reported here that such effects are due to form rather than content, something which has been difficult to demonstrate in studies using narratives.

CHAPTER 6

STRUCTURES FOR ACCESSING INFORMATION

The previous chapters presented a discussion of the problems of separating issues concerning the content of the text from those concerning the form or structure in which the content is expressed. It has been argued that previous attempts to resolve this controversy have failed in general because theorists have neglected to specify clearly the kinds of structures that play a role in text comprehension, and how they are presumed to operate. Experimental evidence concerning the role of text structure has been conflicting, due to the particular text types used for analysis — i.e., narratives — which tend to confound the form-content distinction. The study described in the previous chapter was an attempt to clarify this issue by using a text type that does not involve this confounding to such a great extent. It was demonstrated that, when readers are unfamiliar with the content of a text, they are able to use their knowledge of the text superstructure — i.e., the functional categories of the text — to help them comprehend and later recall it. Furthermore, this effect appears to be the result of an active encoding strategy — in other words, acting as an "advance organiser" for understanding the text.

Presumably, readers who were using their knowledge of the superstructure of the text were able to pick up and identify implicit cues to the structure of the text, as well as being able to use more explicit signals, such as particular phrases (e.g., "since", "because", "however", etc.), in order to confirm their expectations concerning the functional categories to which the particular sentences belonged. They were able to do this despite the fact that the text was presented a sentence at a time, so that they did not have access to the explicit spatial cues that normally exist in printed forms of these texts.

This study, and much of the previous discussion, has concerned the way readers use a structure of which they already have knowledge, to guide their processing and facilitate their comprehension, by picking up implicit as well as explicit cues in the text. However, highly formalised or conventionalised structures are to be found in only a few text types. Other types of text, such as expository and technical prose, do not contain such a stereotyped structure, though they often contain expected components such as aims, conclusions, etc. Moreover, not only is the form of such texts often unfamiliar to their readers, but so is the content, since these texts are generally read for the purposes of informing the reader and creating new knowledge structures, rather than instantiating existing schemas.

It has been argued (Johnson & Mandler, 1980) that with such texts, there is a greater need for the use of a well-organised structure to facilitate comprehension. However, the study reported in the previous chapter suggests that the structure of the text does not make much difference when the reader is unfamiliar with both the form and the content. The instructions given to low knowledge subjects failed to improve their performance significantly. However, it must be noted that although these subjects were told about the canonical form of the text, they were not taught explicitly how to use it to identify the functional categories in the text. It may be that a combination of "advance organiser", and the use of more explicit structural cues in the text, can improve comprehension and retention for those who are unfamiliar with the content. In other words, if readers are taught to use the structure, rather than just being told about it, this strategy may be more successful.

Research on the effectiveness of advance organisers has in general produced mixed results (Hartley & Davies, 1976; Mayer, 1979). Frase (1969) showed that recall is enhanced by informing the reader of the structure of a passage prior to reading. However, Hartley and Davies (op. cit.) suggest that such advance organisers are useful only in certain circumstances.

Providing explicit structural cues in the text can also facilitate comprehension and recall. For example, headings help readers to integrate information as it is being read (Dooling & Lachman, 1971). They also help the reader in scanning and selecting items (Hartley, Kenely, Owen & Trueman, 1980; Hartley, 1981), as well as in comprehending the whole text (Klare, Shuford & Nichols, 1958; Wright, 1977). Typographic cues can also facilitate comprehension (Shebilske & Rotondo, 1981; Waller, 1979, 1982). However, Hartley (1981) points out that, although typographic cues can help, they can also hinder, especially when multiple cueing is used (Hershberger & Terry, 1965; Glynn & DiVesta, 1979).

Spatial cues can also serve to indicate text structure and facilitate comprehension. Rothkopf, Fisher and Billington (1982) found that readers encode many of the display characteristics of the material they read, and that they can subsequently use this information when searching through the text. Hartley and Burnhill (1976) found that leaving an empty line between paragraphs is more effective a cue than just indenting the first line of the paragraph, especially when the text is typographically complex (Hartley, Burnhill & Davies, 1978). Frase and Schwartz (1979) found that technical prose was searched and understood more quickly when it was "meaningfully segmented" and indented. It has been found that pauses while reading a text mark effective segmentation or chunking points for skilled readers (Aaronson & Scarborough, 1977). The assumption is that these units represent points at which the reader engages in special encoding and storage activities. Frase and Schwartz (1979) used subjective judgements of phrase boundaries in their texts. However, Hartley (1980) had subjects segment a technical passage, having studied one of Frase and Schwartz's passages, and did not find the same uniformity. He argued that Frase and Schwartz's results were due to the subjects having prior knowledge of the sentences before verifying them. However, this may not be so much an objection as an indication that prior knowledge of the structure — or knowing how to use it — is important. In this case Frase and Schwartz's subjects presumably also had prior knowledge of the content.

Various methods have been proposed for combining the use of advance organisers with typographical and spatial cues in order to explicitly signal text structure. One such approach is "information mapping" or "structured writing", developed by Robert Horn (1975, 1982). Information mapping consists of a set of rules or procedures for classifying, organising, and presenting information in the text. It involves chunking information into units and labelled blocks. The presentation format is kept consistent, with various labels, and typographical and spatial forms, and the information blocks are functionally sequenced according to their content into superordinate structures, or information maps. The labelling of blocks serves as advance organisers, and the information map as "ideational scaffolding" (cf. Ausubel, 1968). However, although there is some empirical support for the facilitative effects of Horn's technique (cf. Horn, 1982, for a review), there have also been conflicting findings. Jonassen and Falk (1980) compared information mapping with programmed instruction and found them to be equally facilitative in recalling text, although retrieval (not the same as recall) benefitted from information mapping. However, in another study, Jonassen (1981) found information mapping to be superior to programmed instruction in facilitating recall. The technique has not received wide acceptance, and some researchers are critical of it as a "once and for all" solution, arguing that existing methods serve just as well (e.g., Hartley, 1982).

Not only is it difficult to make comparisons between different versions and layouts of a text (cf. Hartley & Trueman, 1981), but different layouts can be good for different purposes. Furthermore, unless the reader is informed or taught in advance how to use the explicit cues, they often have little effect (Foster, 1979; Hartley, Bartlett & Branthwaite, 1980). This also goes for advance organisers. Studies have in fact shown that comprehension and recall can be improved if subjects are taught how the technique of mapping or flowcharting is related to the content of the text, and if they are explicitly taught how to use it (e.g., Armbruster & Anderson, 1980, 1982; Geva, 1981). Kieras and Bovair (1984) found that the provision of a "mental model" before learning instructions only improved performance over rote learning if subjects were able to infer the procedures for operating the device from this model. Otherwise, the provision of a model made no difference, and in fact impaired performance slightly.

Wright (1977) has emphasised that importance of considering the reader's purpose, putting the stress on usability rather than readability (Wright, 1983). She argues that the form of expression that is right for one situation may not be appropriate for another. This is particularly important in the case of instructional manuals. Rothkopf (1982) also argues that, from the point of view of instruction, the important factor concerning the text is not just the way it is structured, but the relationship between the text and the instructional goals.

There is a good deal of evidence that the structure and sequencing of instructional text affects the performance of tasks. Clark and Clark (1968) showed that people found it easier to remember sentences when the order of mention corresponded to the order in which events were carried out. Enkvist (1981) uses the term "experiential iconicism" to refer to the way in which information ordering in the text (e.g., a recipe or repair manual) reflects the ordering of events in the world. In such cases the text becomes "mimetic" of experience. This "use-order" principle, or "experiential iconicism" suggests that, for procedural or instructional text, comprehension should be easier if the order of information within an instructional step is congruent with the order in which it is used to construct a plan for carrying out the procedure.

Dixon (1982) postulated that plans are hierarchically organised, with actions to be performed at the top level and the consequences linked to actions at a subordinate level. Thus, at the top level of the hierarchy, plans consist of lists of actions, with consequences linked to each of those actions. In this model, antecedent conditions necessary for carrying out an action are linked to actions in a similar subordinate role as consequences, even though in practise they have to be executed first. Dixon found evidence that plans are organised around actions rather than in a strict temporal sequence. Sentences were read faster when the action was stated first and the condition second, regardless of whether the action had to be performed immediately or was to be remembered. However, Spoehr, Morris and Smith (1983) found, in contrast, that instructions following a temporal order sequence, in which the antecedent condition came first, produced faster reading times, whether reading was for immediate execution or for verbatim recall. There was no evidence that the "action first" principle proposed by Dixon had any advantage. Spoehr et al. attributed their conflicting results to the fact that Dixon used the same object as the referent for both antecedent and consequent, thereby averaging the effect of presenting the antecedent and consequent first and last, whereas Spoehr et al. separated these by having the antecedent and consequent refer to different objects. They argue that the best order in which to present the components of an instruction is the order in which the corresponding slots appear in the schema for the instructional step.

However, the studies just described concern what might be called the "microlevel" of instructional text — that is, the representation of a single step. Just as episodes and events in stories are clustered according to the goal or purpose they are intended to achieve, so single steps in instructional text are clustered or hierarchically arranged in terms of super- and subordinate goals (Smith, Spoehr & Collins, 1982). Graesser (1978) also found evidence for a hierarchical plan-based representation of procedures. This kind of analysis is also consistent with the analyses of Moran (1983), Robertson and Black (1983), Norman (1984), and Riley and O'Malley (1984).

Smith and Goodman (1982) found that hierarchical instructions were superior to linear instructions in terms of reading time, execution accuracy, and memory for the instructional steps. However, Kieras, Tibbits and Bovair (1984) compared linear or step-by-step instructions with hierarchical ones, and found that the latter was superior only if the subjects were familiar with the type of device for which the instructions were intended — in fact performance was sometimes worse if they were not familiar with it. Clement (1984) has interpreted Smith and Goodman's (1982) findings as indicating that the hierarchical instructions (both structural and functional) provided causal connections between the procedures described in the linear instruction, the goals to be achieved, and the structure and function of the device. This is consistent with the findings of Kieras (1982b), who showed that users' knowledge of a device was hierarchically organised in terms of functions, operations, and what he called "how it works" knowledge, and that the provision of a mental model of the device facilitated performance only when it enabled subjects to infer procedures from the model (Kieras & Bovair, 1984).

The notion behind the facilitative use of mental models is that they provide a causal model that enables the user to simulate or "envision" how the device might function. (Cf. for example, Gentner, 1982; De Kleer & Brown, 1983; Johnson-Laird, 1983, for discussions of mental models.) The provision of such mental models becomes particularly important when instructional text is applied to the domain of computer use. In this case there is no "device topology" to be examined by the user, except what can be inferred through the "system image", or what is visible to the user on the screen (cf. Norman, 1984; Riley & O'Malley, 1984), or via the documentation. Furthermore, when the documentation itself is online, users have the additional problem of accessing the information that will assist them in understanding how the system works and in using it to accomplish their tasks.

The Provision of Access Structures in Human-Computer Interaction

The discussion that follows stems from two observations, or rather two views of one observation: first, that users prefer to consult other people rather than to use the manuals and other types of help provided for them; second, that the people that users consult may be characterised not simply by how much they know, but also by their ability to get access to the relevant information.

The perspective of "information flow" has been proposed recently to emphasise the importance of a unified approach to the provision of adequate computer help systems (cf. Norman & Draper, in press). This perspective includes information sources other than manuals, such as other people, system displays, system messages, prompts, and so forth. This perspective, therefore, means paying attention to a variety of types of information that must reach users during the execution of their tasks — from low-level feedback, such as the echoing of characters on the screen to the higher level information necessary to plan the next activity. It involves considering not just what information users actively seek, but also what they pick up by chance (Owen, in press); it involves considering not just the information that can be obtained from the computer system and the "official" documentation, but also what help can be obtained from other people (Bannon, in press).

A distinction has been made traditionally between "data" and "information", which is captured in the notion of "informativeness", and defined as some relation between the perceiver and the scene, rather than as only a property of the scene. Woods (1984) argues that displays present data, and that data becomes information only when used to answer a question or to perform some task. However, the information flow perspective assumes that anything can be information, but it only takes on meaning when it is picked up by some receiver. This definition is in keeping with that used by Dretske (1981, 1983). Information flow analysis is, therefore, based on the premise that one can analyse information requirements by specifying the content and sources of information, and the media or channels through which they may be transmitted to some receiver. However, it is argued here that information flow will only succeed to the extent that the effort required to find the information is reduced (and ideally, abolished). In other words, the structure of the information is a vital factor in determining how much effort is required to discover if the relevant information is even present, and, if it is, to make use of that information. I argue in this chapter that many of the problems in existing methods for providing help can be viewed as problems in getting access to the relevant information.

Despite the variety of forms of help that are widely available, it is commonly observed that users tend not to use manuals or other forms of help, but prefer to consult the local expert or other users (Sticht, 1972; Wright, 1979; Lang, Lang, & Auld, 1981; Lang, Auld & Lang, 1982; Scharer, 1983). There are several possible reasons why users give up on manuals or help systems: Common complaints include the fact that they can't find the information, that when they do they can't understand it, or that there's too much information to wade through, they don't have the time to spend searching, or that they just want to check up on some detail, and so on. It isn't just that the manual entries are badly written. A recent study, for example, showed that even when manuals were given to experienced technical writers to revise, they still didn't improve matters (Duffy, Curran & Sass, 1983).

Wright (1983) has coined the phrase "manual dexterity", emphasising the importance of the usability of documentation. She noted that the fact that users don't often read manuals has a good deal to do with problems of getting at the information they want — in other words, they don't read manuals because there are quicker and easier ways of getting the information (such as asking a colleague). However, when users get help from other people, it is not always the case that the other person knows the solution. Often it is more the case that the other person knows where to find the information that will answer the user's question. One of the characteristics of the "local expert" in a user community is an ability to use the documentation and manuals. This is just as likely to be the reason why such people are frequently consulted, rather than the fact that they "know more" than other users.

This simple observation bears closer analysis: Users may go to the "expert" because they are unfamiliar with the way the available information is organised, and it is less effort to ask someone who is familiar with it. Knowing something about the way information is organised means being able to ask the right question to get at the required information. So, another way of articulating this observation is that "experts" are able to point users towards the information they need because they are able to formulate the question better than users can.

Users have trouble finding the answer themselves because there is usually a huge gap between the initial internal (mental) form of the query and the information they need, as expressed by the system. Other people are often better at closing parts of that gap, by helping users translate their intentions into specific questions that can then be mapped onto the form in which the information is presented and organised in the system or the documentation. In fact, there are two problems involved in bridging the gap between the user's initial query and the information as represented in the system: One problem has to do with translation from the user's form of the query to the input language understood by the system. The other problem is where the user needs help in actually formulating a question in the first place.

Waller (1979, 1982) uses the term "access structures" to refer to the spatial and typographical features of printed text that help the reader in giving an overview of the topic, in planning active and selective reading strategies, and in explicitly representing the structure of the text. The process of bridging the gap between the computer user's intention and the information needed can also be seen as a process of constructing appropriate "access structures" with which to get the required information.

The provision of computer-based text display turns reading into information retrieval (Kommers, 1984) and question-answer dialogue (Wright, 1978). However, much of the design of information systems is still based on the linear text model and hasn't yet caught up with the technology (Bork, 1983). Weyer (1982) notes that electronic information systems have been regarded as panaceas, but actual systems often do not venture beyond the imitation of traditional media, functioning merely as automated page turners.

In using the phrase "access structures" I am not simply concerned with what we normally understand by information retrieval — that is, databases, and the query systems used to access them — but also with issues normally thought of as presentational, such as where to place windows, menus, and so on. By expanding this view of information access we incorporate all the problems to do with information clutter as well. That is, just as it is not enough to make information available in the manual (one has to provide the users with easy ways of getting access to the information), so it is not enough to present information on the screen — one has to present it in such a way that it is usable as well as useful. People normally think of textual access structures in the same limited way that they think of information retrieval — that is, query languages, contents lists, indices, and so on. In practise, there are many other structures and mechanisms used to access information. For example, Sticht (1977) studied the use of manuals and found that contents lists and indices were only used 27% of the time where their use was relevant, whereas the activity of leafing through the manual was used more than 90% of the time. This might suggest not just that these access mechanisms were badly designed, but that they were not appropriate for the needs of those particular users. There are many more routes to information than have been exploited to date. Designers need to be conscious of alternative means of obtaining information if they are to design really usable and useful information facilities.

Fischer, Lemke and Schwab (1985) distinguish between explicit and implicit communication channels for the transfer of information. The first kind (explicit) includes such things as windows, menus, pointing devices, icons, etc.; the second (implicit) implies the existence of a shared knowledge base. One can also distinguish between explicit and implicit access structures: Database retrieval systems provide explicit access structures, in the form of query languages, menus, commands, etc. Other examples of explicit access structures for printed materials are tables of content, indices, etc. However, there are other kinds of mechanisms we use to access information that are much more implicit. The organisation of the content can act as an access structure (e.g., "information mapping", Jonassen, 1981; Horn, 1982), as has been discussed with respect to printed material. Furthermore, as the previous chapters have discussed, stereotypical or conventional textual forms can also act as access structures.

Access structures are important not just for static displays. Dynamic and concurrent information can also benefit from the use of spatial and other cues (cf. Green & Payne, 1982). Woods (1984) discusses the importance of ensuring the structuring and integration of information across successive displays. He argues that even with simultaneous displays (e.g., "windows") display systems can become serial rather than parallel presentation media. He uses the notion "visual momentum" as a measure of the user's ability to extract and integrate information across displays, arguing that when visual momentum is high it is analogous to a "good cut" in film editing from one scene to another. Low visual momentum is like a bad cut in film editing — one that confuses the viewer, or delays comprehension. Low visual momentum, in Woods terms, can lead to memory bottlenecks due to increased mental workload, cognitive tunnel vision, inability to locate important data, and so on. He outlines several techniques for providing "access structures" across displays. These techniques work by building a spatial framework (form) that reflects the semantic structure of information (content) — that is, its relationship to user tasks.

One such example is the use of a fixed format, where information is assigned to specific screen locations (e.g., menus). Another example is the use of "long shots". These provide an overview of the display structure, and summarising information, together with some local detail representing "important" information. For example, Furnas (1983) describes a system called FISHEYE, which provides, quite literally, a "fisheye" view of the information. Furnas suggests that the most useful parts of a large structure to show (i.e., the most "interesting") is a combination of two components: The first is a component independent of the current interaction, reflecting parts of the structure that are of a priori global importance. The second is meant to capture the contribution specific to the current focus of interaction, and is approximated by simple distance from the current focus, measured by "degree of interest" (the more distant parts being intrinsically less interesting to the current interaction). The "fisheye view" thus represents a combination of local detail and global context.

Other features, such as "perceptual landmarks", help the user to integrate across successive displays by providing a recognisable feature which anchors the transition and provides a relative frame of reference. Overlapping displays can also be useful, especially if the overlap sections do not contain the same level of detail as the main part of the display frame. Overlap can also be achieved through overlays — e.g., in map-making several layers of information are presented on top of a common geographical framework. (See the online manual described in Feiner, Nagy & Van Dam (1982), for an example of overlays in computer displays.)

Wright (1979) argues that to answer questions successfully depends on the relationship between the way the information is expressed in the question and the way it is represented in the mind of the respondent. Similarly, access structures are successful to the extent to which they provide a good mapping between the desired information and the query of the user. As has been suggested, one of the reasons users go to local "experts" rather than use the manual is not because the experts know more, but because they have the ability to access and use the information contained in the documentation. In other words, the people who appear to make use of manuals (the "experts") are those who know enough about the information that is available to use those "access structures" to find a solution. Novices and casual users may often be limited in their abilities, not only because they know less about the system than the experts, but because they are unfamiliar with the available documentation and how to access the information it contains.

Current help systems are not sensitive enough to the variety of levels of user expertise, nor to the variety of contexts in which online help is required. The form in which the information is accessed and presented doesn't match the form in which the user initially makes the query, nor the type of task for which the information is required. Documentation should be provided at several different levels — not simply because there are different kinds of users — but because at various times, and for various types of task, users need different forms of information.

DIFFERENT TASKS REQUIRE DIFFERENT KINDS OF HELP

A Study of Users' Documentation Needs

This argument was examined by a study conducted in co-operation with researchers at the Institute for Cognitive Science, University of California, San Diego. The study considered the online help facilities for UNIX, and led to proposals for three different forms of documentation to replace one existing system: namely, a quick reference facility, a task-specific help system, and a facility for providing full explanation. The first step was the design of a quick reference facility, followed by an evaluation study.

The UNIX online manual contains separate entries for each program, which is accessed by typing the command *man* with the program name as the argument. This produces several screenfuls of text in a standardised format, with the name of the program, a short synopsis, a longer description of the program and how to use it, some examples, and some diagnostic information. Although the manual entries generally follow a standard format, because they are written by many different programmers, there are large variations in style, verbosity, and differential emphasis on the kinds of information given about the program. Some entries rely on using many examples but not much explanation, other entries are verbose, but contain few examples of use. Most entries also assume a good deal of knowledge on the part of the user. (There are also two other programs that give short descriptions of what the command does: *whatis*, and *apropos*. However, the usefulness of these programs is limited to reminding a user of what the command does, but gives no information about syntax, options or flags, and requires that the user know the name of the command.)

Design and Procedure

The use of this online manual was monitored by means of built-in system accounting information, and then compared with the use of the system commands. The system accounting information was obtained from records kept automatically, which indicate, amongst other things, the user identification number, the commands and arguments used, and the time at which each command was executed. Thus, by looking at each user identification number, one can build up a rough picture of the user's pattern of activity. However, since patterns of use do not give any indication of particular problems encountered, this method of data collection was supplemented by a facility by which users could furnish online comments concerning the information they were trying to obtain. They were asked to give their reasons for using the manual, and whether or not they were successful in finding the information they wanted. Thus, the system accounting information was used to identify patterns of use, and users' comments were used to identify common problems and to suggest reasons for the patterns found in the system accounting data.

Results and Discussion

Three categories of information seeking were identified from these data. The first category was the need for quick reference information. This was identified mainly from users' comments: Users said that they had received too much irrelevant information, that it took them a long time, when all they wanted was to check on the appropriate command, its syntax, and its options or flags. The second kind was a need for more task-specific help. This was identified by requests for help for groupings of commands that perform similar operations on similar objects. Users' comments indicated that they were trying to find information concerning a task for which they didn't know the name of the appropriate command or command line. All other kinds of information requirements were classified under the broad heading of full explanation, which included tutorial help. Users said that they wanted to learn how to use a command they had heard about, or to get a fuller understanding of some program.

Analysis of the data revealed that 50% of the use of *man* was for task-specific help, 35% for quick reference help, and 15% for more detailed descriptions and explanations (see Figure 6.1). The feedback provided by users indicated that the

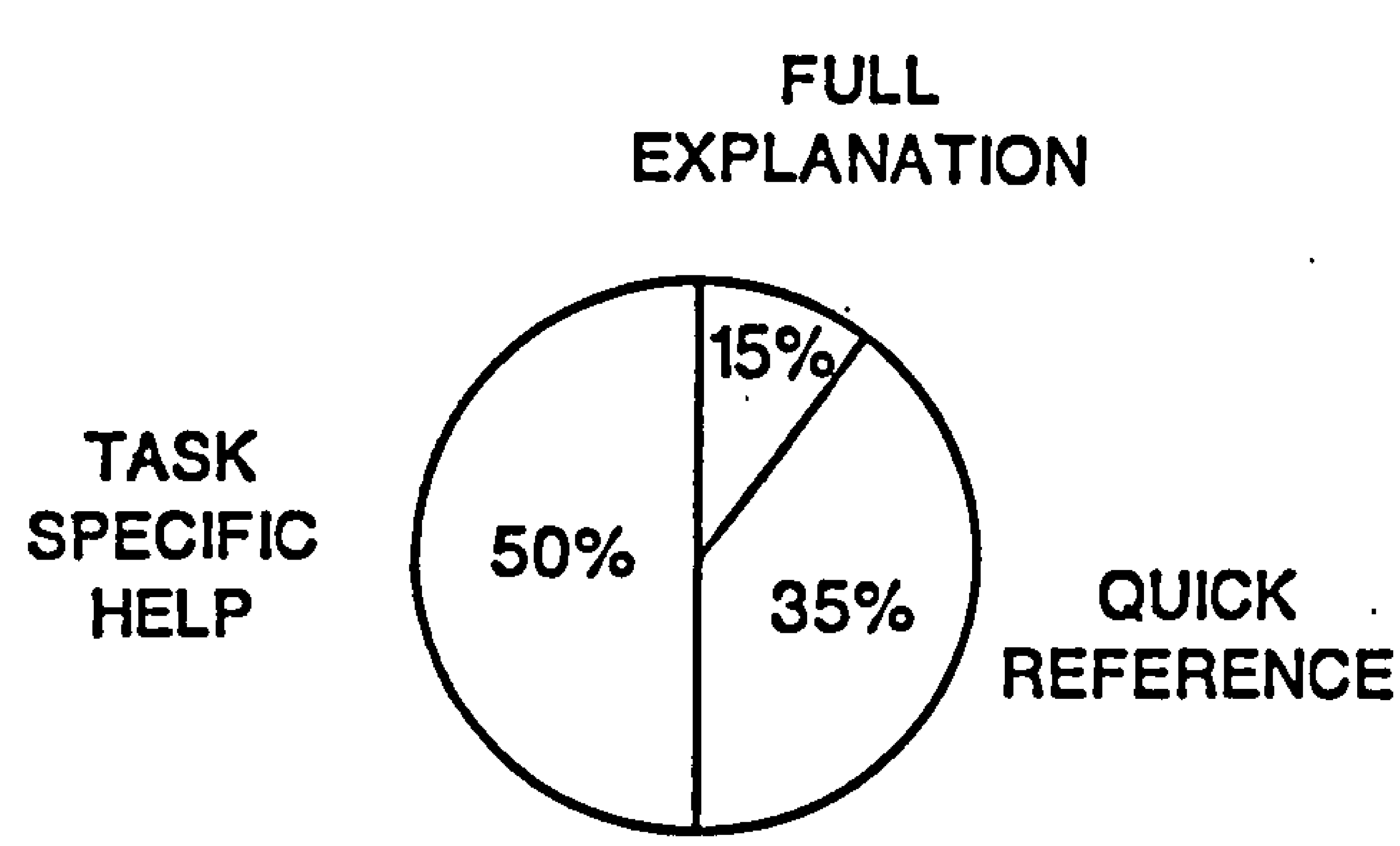

Figure 6.1. Use of existing online manual.

existing manual was unsatisfactory, particularly with respect to the first two kinds of help. In order to get task-specific help users have to do a lot of searching through many manual entries, since most of the time they don't even know the name of the program they want. There is no real quick reference help, since users have to get the full manual entry in order to check up on flags and syntax, which takes up a good deal of time.

DESIGN AND EVALUATION OF A QUICK REFERENCE HELP SYSTEM

In order to address the need for quick-reference help, a prototype system was developed for the domain of programs associated with printing tasks, since these were a subset of commands widely used by all members of the user community. Figure 6.2 gives an example of a display produced by this system. The system gives

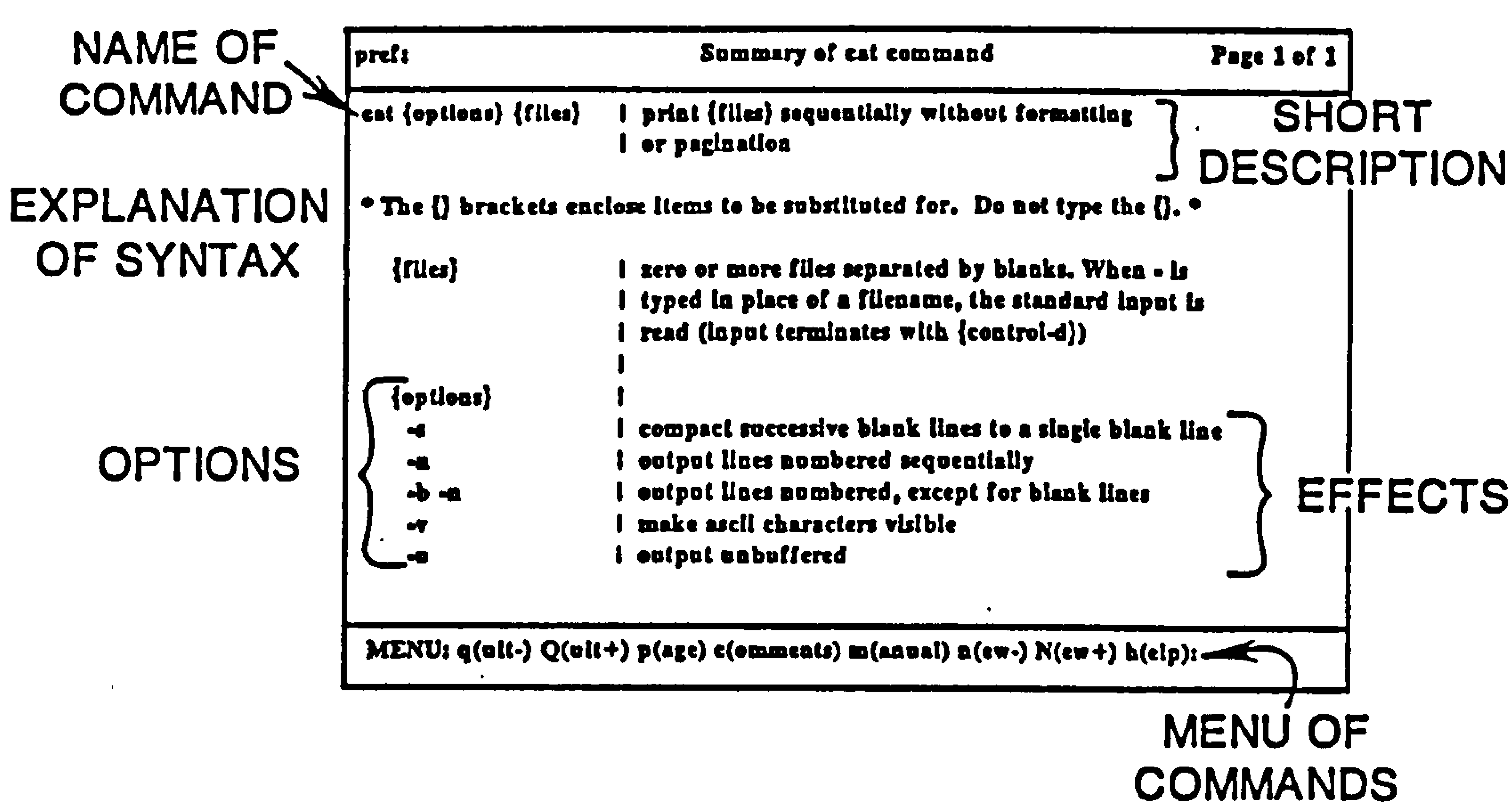

Figure 6.2. Example of a Quick Reference display.

the name of a command, with a brief explanation of syntax. It lists the various options available and gives a short description of their use and effects. At the bottom of the screen is a one-line command menu.

Design and Procedure

A means of evaluating this facility was incorporated into its design, so that users could indicate their success or failure to get the information they wanted by a simple menu of commands. In order to quit the program, the user could type upper case *Q* if the facility was useful, and lower case *q* if it was not. (The reason for using upper case to indicate success was to lessen the probability of response bias, since lower case is the normal form of a command, and it is easier to type.) Users were also given the option of providing more detailed feedback, by typing *c*, which put them in the editor where they could type their comments and then return to the quick reference entry.

Users were also able to get a more detailed explanation of the command from within the quick reference facility, by calling the regular online manual, without having to quit the program, by using the command *m*. They could also specify new arguments quickly and easily, by typing lower case *n* if the entry just seen was unhelpful, and upper case *N* if it was helpful. Users could also get online help for the use of the facility, by typing *h*.

Results

It was predicted that if the quick reference facility was serving its intended purpose, a subsequent change in patterns of usage of the online manual would be found. The use of the online manual prior to implementation of the facility was compared with its subsequent use, over a period of fifteen weeks.

When the system accounting data were examined, a difference of 58% was found in the use of *man* (for commands covered by the quick reference facility) before and after the introduction of the facility (see Figure 6.3). When the use of the manual before and after the implementation of the new facility was compared for those who used it and those who didn't, it was found that the drop in the use of the manual was greater for those who used the quick reference facility than for those

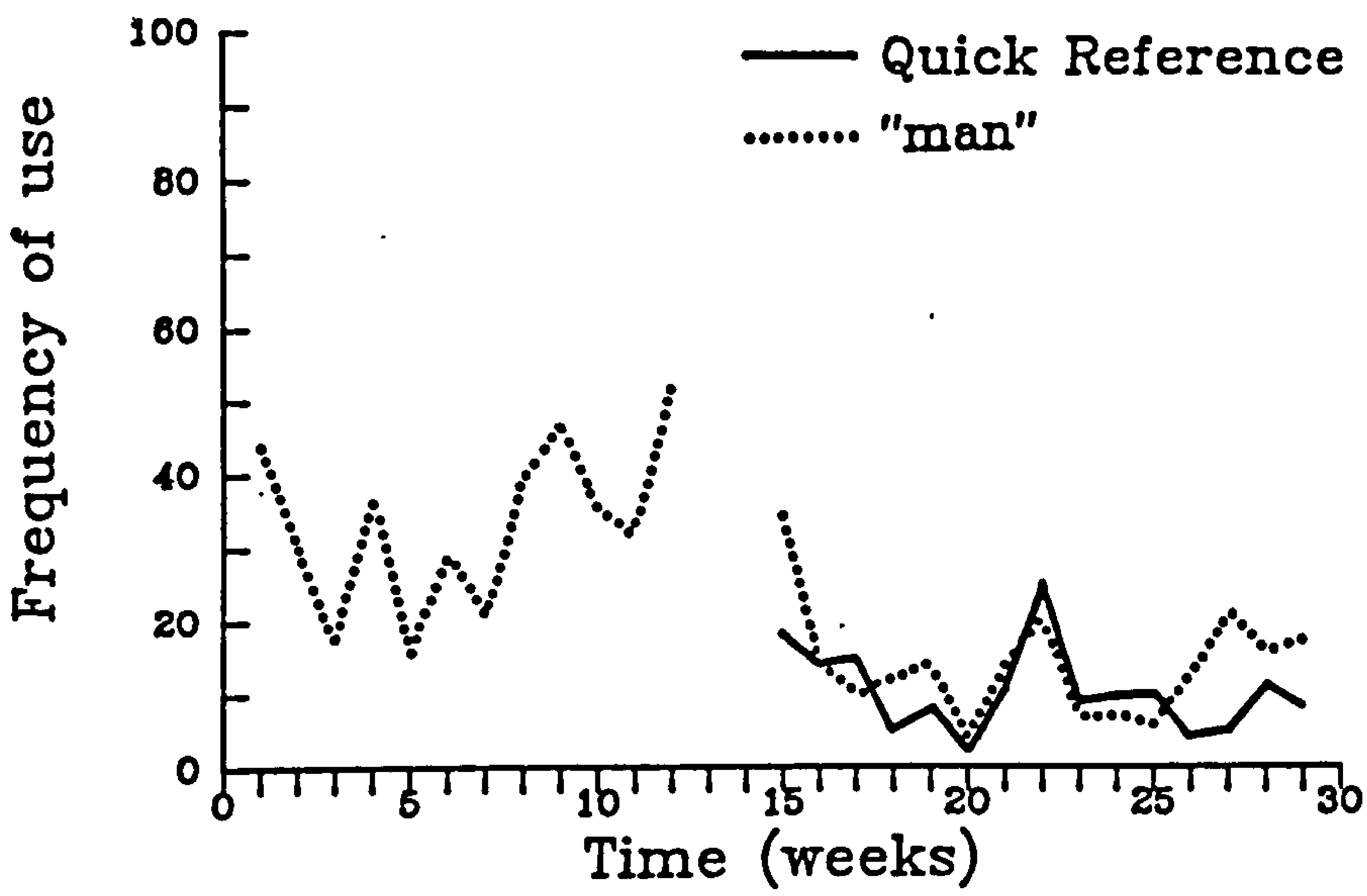

Figure 6.3. Comparison of frequency of use of *man* and Quick Reference.

who did not (see Table 6.1 and Figure 6.4). Analysis of variance revealed a significant interaction between the period covered (i.e., before or after the introduction of the facility) and whether or not the user had used the facility ($F(1, 25) = 15.116, p < .01$). The possibility was ruled out that the difference in the use of the online manual was due simply to a difference in the actual use of the printing commands for that period, firstly, because the difference was much greater for those who used the program (63%, as opposed to 36% for those who did not use it), and secondly, although there was a slight drop in the actual use of the printing commands (about 27%), when the data on the use of *man* were normalised, by expressing them as a ratio of the use of the corresponding print commands, the difference in the use of *man* was still much greater for those who used the program, than for those who did not (see Table 6.2). So, the effect of the introduction of the quick reference program on the use of the pre-existing online manual appeared

Table 6.1

Mean Use of *man* Per Week			
	(12 weeks) Before QR	15 (weeks) After QR	% Difference
QR users $N = 23$	27.58	10.27	62.76
QR non-users $N = 14$	5.67	3.6	36.51
Total (users + non-users)	33.25	13.87	58.29

F Table

	df	*MS*	*F*	*p*
Bef/Aft	1	1252.38	28.06	0.00**
QR/nonQR	1	2440.17	47.58	0.00**
B x Q	1	775.21	15.11	0.00**
MS	25	51.29		

to be a genuine one.

When the use of the upper and lower-case commands for the facility was examined, a success rate of about 55% was found, indicated by comparing the number of upper-case quit commands with the number of lower-case quit commands (see Figure 6.5). This analysis ruled out the possibility that users were simply always typing upper or lower-case, since, even though there was a significant difference between *Q* and *q* ($t\,(32) = 2.504$, $p < 0.01$) there was no corresponding significant difference between *N* and *n* ($t\,(32) = -1.141$, $p < 0.26$).

Examination of the mean time elapsed from calling the facility to quitting it showed that those who found the facility unhelpful (lower case) took longer to find the desired information (see Table 6.3). The use of the lower-case *n* was higher than the use of the upper-case *N*, which possibly reflects the fact that if users needed to

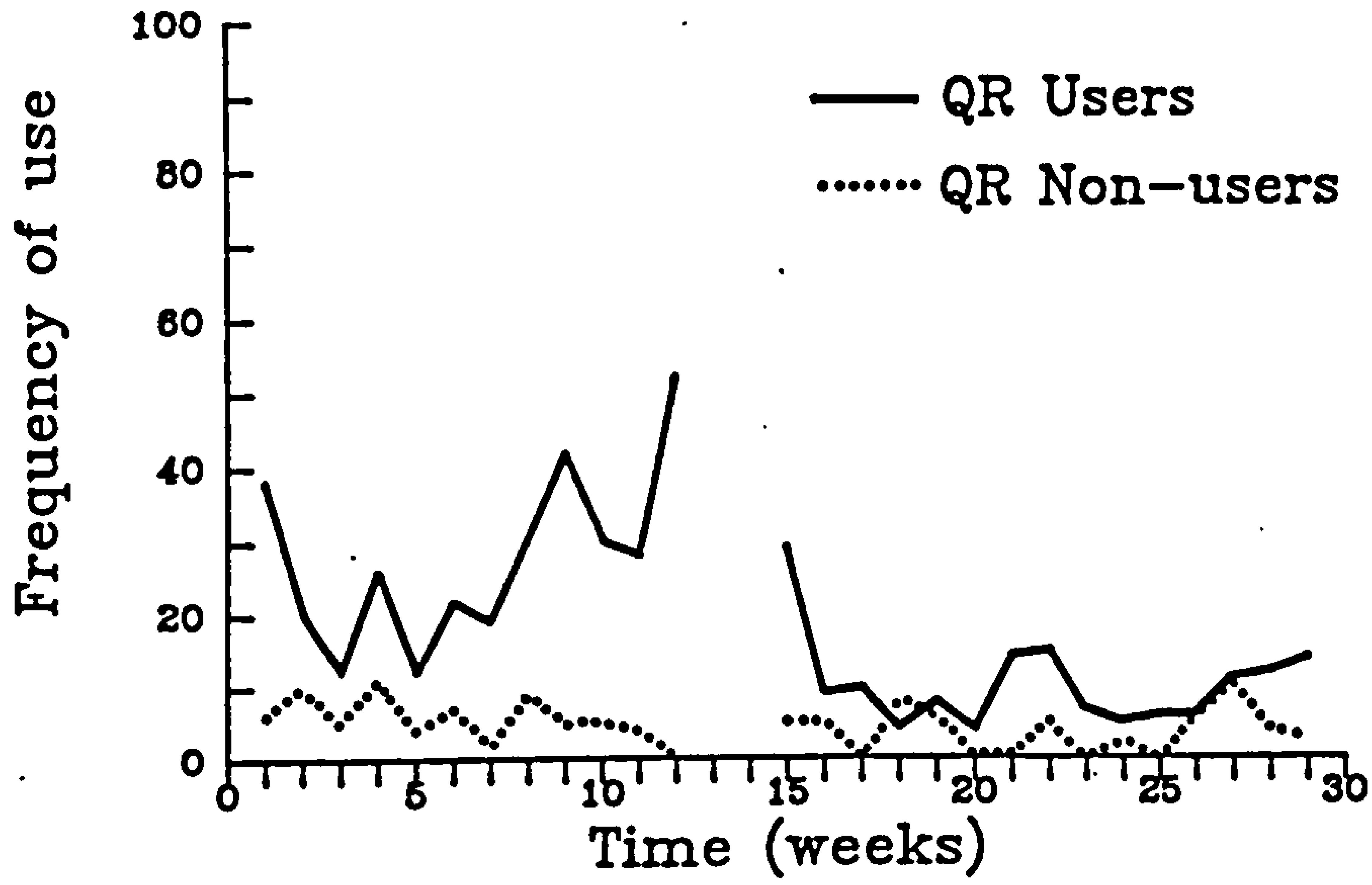

Figure 6.4. Comparison of the use of *man* for "users" and "nonusers".

Table 6.2

Mean Use of *man* Per Week			
	Before QR	**After QR**	**% Difference**
QR users	0.61%	0.31%	49.18%
QR non-users	0.13%	0.11%	15.38%

select a new command, they had found the previous one unhelpful. (The data on the length of time to find information provide an internal consistency check on the evaluation method.)

Finally, the data were examined to see whether or not the use of the facility

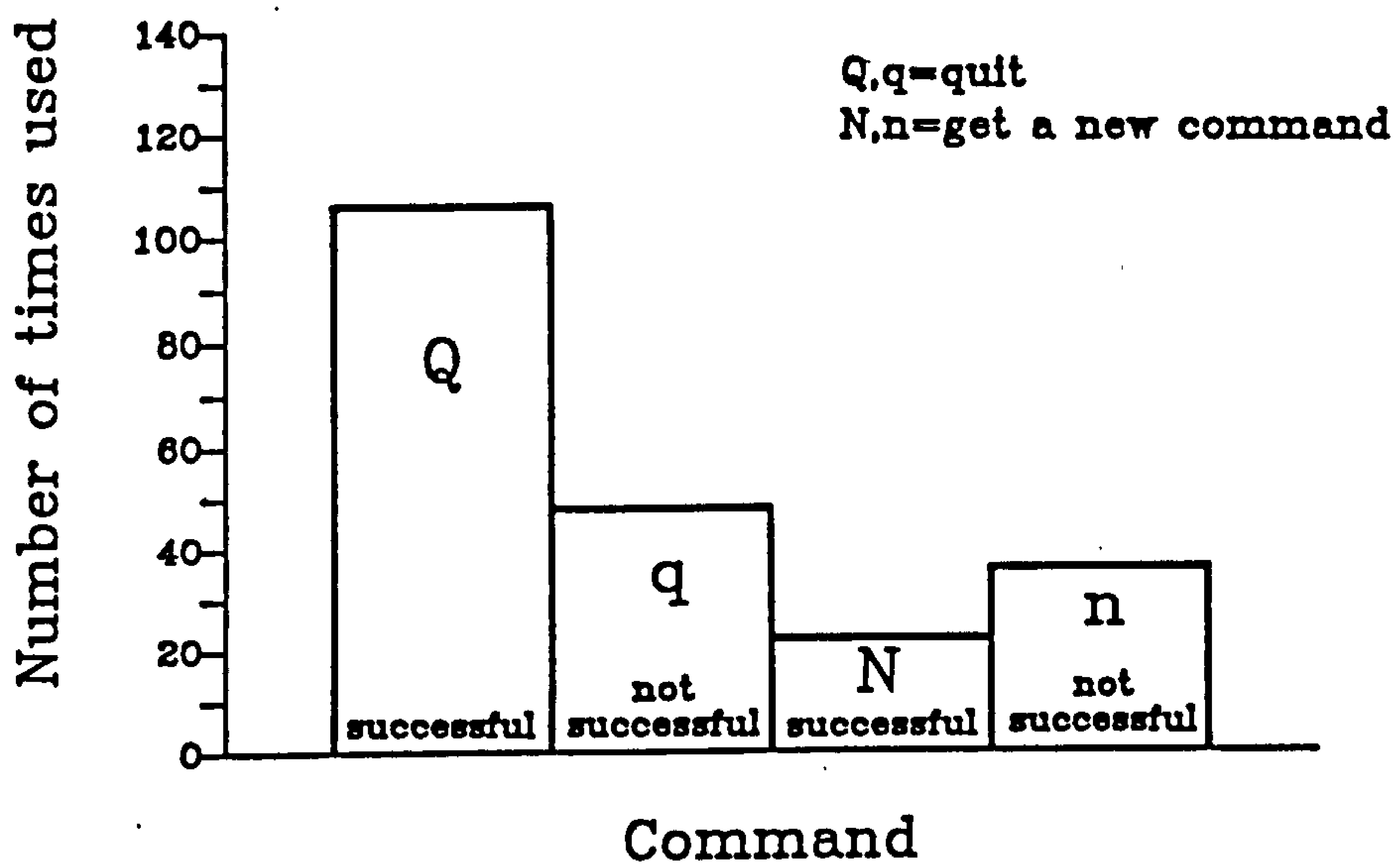

Figure 6.5. Use of upper and lower-case commands.

followed a different pattern between commands selected as arguments than did the use of the manual. It might have been hypothesised that even though the new facility appeared useful, it was only useful for certain commands and not for others. Use of the manual (after implementation of the facility) was examined for each command, and compared with the same data for the use of the facility. There was a high correlation between these sets ($r = 0.89$, $p < .01$), indicating that the pattern was no different with the introduction of the quick reference facility.

Table 6.3

Frequency of use of Quick Reference		
Command	**Frequency**	**Mean time (secs)**
Total	154	
Q	106	86
q	48	141
N	22	42
n	36	32
m	12	146
h	29	91
c	10	66

Discussion and Conclusions

The results suggest that the prototype was reasonably useful to people. The online feedback obtained from users also seemed to confirm this, indicating that users found the program a positive and useful addition to the existing facilities. Their comments also indicated, however, that there were improvements to be made concerning the design of the display, and so on.

In summary, the evaluation provided a number of measures supporting the hypothesis that a quick reference facility was needed. There was a drop in the use of *man* for printing commands after the facility was introduced, and online evaluations of the facility were, on balance, favourable. Despite the apparent usefulness of the facility, however, the design actually used glosses over two distinct kinds of quick reference help that is needed: It does not make the distinction between quick reference as a reminder for users already familiar with the use of the program, and examples of use for those not familiar with the use of the program. (Such distinctions are made, for example, in the SPSS documentation.)

However, there was a tendency, over a longer period of time, towards a slight reduction in the use of the facility (cf. Bannon & O'Malley, 1984). One of the factors responsible for this may have been that the novelty of using the facility wore off after a certain time. Another possibility is that users may have learned the information (which is reasonably brief) enough not to need to use the facility as much. Nevertheless, the important point concerning documentation is not so much how heavily it is used, but how useful it is when the occasion to use it arises. For example, our use of a dictionary entry may not be frequent, but the facility may nevertheless be very useful on occasion. This observation highlights the importance of considering several measures in evaluating documentation, especially the contributions that can be made by users' feedback.

The quick reference facility may be seen as serving a function of reminding people of what they actually know, but have forgotten. It can also be seen as a facility for structuring the information that is already actually contained in the appropriate manual entries, but which requires, in its present form, considerable effort on the part of users to select out the information. Thus, this facility can be seen as a kind of "access structure". However, this facility is only appropriate for certain types of task — that is, tasks for which the user simply wants to check up on information they actually already know, and just need reminding on. Another kind of access structure is needed for tasks that involve a certain amount of problem solving or planning.

Evidence had also been found in the study of the UNIX online manual that users needed some facility that would give them specific help on tasks for which they did not know the appropriate commands to use, nor how to use them. The problem here is one of providing some means of access for users to the information, but in terms that match their conceptualisation of the task. In addition, this information should be provided in a way that enables users to carry out tasks without additional problem solving.

A system like the UNIX operating system is extremely powerful and flexible, based on a philosophy of modular design. This modularity is a positive feature: Users can tailor their environment to their own needs and combine commands in many varied ways. However, this modularity can also lead to problems in ease of learning and use. The UNIX system in this study contains, in addition to approximately 300 standard commands or programs, about 400 other commands specific to the particular computing environment. This means that there are around 700 commands for users to learn. No one is expected to learn the full set, and in fact there is a certain amount of evidence that the average user's command set is a much smaller subset of this. Furthermore, users' working sets seem to overlap marginally with each other, so that there is a good deal of variety in what different people know about and can do on the system (cf. Draper, 1984).

Before any action can be carried out it has to be translated into an executable command-line. A single command-line may contain several separate programs with their various associated flags or options. In such cases, one kind of organisation of information will not necessarily work for all types of problem. Users' tasks, more often than not, cut across the boundaries forming the organisation of the information.

For example, on the particular UNIX system involved in this study, if you wanted to print a document with tables and equations on the laser printer, you might have to type the command line:

tbl filename | eqn -s8 -f1 | ltroff -mlcsl

This single command-line contains three separate programs with their various associated flags or options. In order to get help for this task you would have to access information via the online manual from three different entries for the programs *tbl*, *eqn*, and *ltroff* — and you would have to spend some time on each query to find the relevant information, assuming you knew the names of the appropriate commands.

The point is that users often have to break down a single task into several components, and if help is required in subsequently constructing the command line, it has to be obtained for each of the components separately. The documentation is structured at the level of the system modules (commands) and not at the level of the task, nor at the level of a legitimate executable command line. This is not a complaint specific to one particular system, it is a problem for any system that offers users the power of combining units.

Accordingly, a documentation system was designed that attempted to address the problem of mapping users' task conceptualisations and descriptions to system modules. This involved a formal analysis of tasks, and a means for formalising task descriptions based on a system of attributes and values — cf. Smolensky, Monty and Conway (1984). The provision of more "task-based" help is one way of improving the mapping, and reducing the amount of translation needed between the user's problem and the information provided by the system. However, this also presupposes being able to characterise tasks in some unique way, and although this may be possible for some domains, it is likely to fail in others, where the task is loosely defined, or has to be negotiated, or cannot be specified in advance (cf. Fikes, 1982), and may vary considerably with context and with users.

DIFFERENT TYPES OF QUESTIONS REQUIRE DIFFERENT KINDS OF HELP

Not only do different types of task require different forms of help, but even within one task various different types of questions can crop up, and should ideally be answerable by the system. There are several possible reasons for users needing to ask questions: They may have a goal to accomplish but don't know how to go about achieving it, or they may need a description of a term or concept. However, they also may just need confirmation or verification of a solution they are considering, or help in testing a hypothesis. They may also be considering several alternative solutions and need help in making a decision between them.

For example, in the course of asking the question *"how do I get a formatted printout of a letter?"* the user may need to ask several more specific questions because of the subtasks involved. For example, *"how do I get a letter heading?" "how do I get the date printed on the right hand side?" "how do I indent the paragraphs?"* etc. These questions are all of the same general type — that is, *"how?"* questions, which are handled more or less adequately by systems that can give some assistance on procedures for performing a task. However, in order to obtain the required information the user may also have to ask several different types of question: for example, "how *do I indent paragraphs?"* "where *do I place the .pp?"* "should *it be on a separate line?"* "what if *I didn't want the paragraph indented?"* etc. Thus, there are often cases where the mapping between users' intentions and the information contained in the system is one-to-many, and where the initial question leads to many different expressions of queries being generated.

Questions themselves can provide many cues to the kinds of information required, since they reveal at least what the user does not know, and therefore, by implication, something of what they do know. However, the match between a person's knowledge structures (and the gaps in that knowledge) and the questions they might ask is not straightforward. Miyake and Norman (1979) found that the amount of questioning dropped as a function of how much was known about the topic — novices tend not to ask many questions on material that is too difficult.

Some research within artificial intelligence has also concerned question types and question answering. However, in general it lacks systematicity and completeness (although Lehnert's (1978) analysis is one of the more thorough attempts), and much of the research has yet to be tested empirically (although some work, e.g., Swartout (1983), at least is based on informal observations of users). A good deal of research has been done in psychology on answering simple questions (e.g., McCloskey & Glucksberg, 1979; Reder, 1982b), however, most of the questions that users might ask are likely to be much more complex than those that involve simple truth verification or straightforward solutions. It is only recently that research has begun on more complex questions. Some research has been conducted on question types (Graesser & Murachver, in press), on the problem of how people know when they have a question (Glucksberg & McCloskey, 1981), and on what people need to know to be able to formulate a question (Miyake & Norman, 1979), however, little research has been done to date on the topic of questions with respect to human-computer interaction.

The task of the addressee of a question — and the task, therefore, of any help system can be seen as one of bridging the gap between the user's initial question and that question that is going to lead (in one step) to a specific answer. There are some types of question that are reasonably well supported by traditional designs such as menu-based help, command-language, or keyword systems, and so on. However there are other kinds of questions that are not possible to ask, let alone answer, with traditional help systems.

Some questions are fairly straightforward, such as "*what is grep?*" The user needs a description of the command, which can be very brief, or may involve a longer exposition — especially if it is a concept they are asking about (e.g. a buffer) rather than a command. Although most systems are able to provide descriptive help of this kind more or less effectively, there is a variant of the *"what is?"* question that is often not taken into account: the question of the form *"what is the difference between...?"* Here the user needs to know the difference, or the relationship, between one or more things.

Another common type of question that help systems often address is the *"how?"* question — that is, requests for information about how to accomplish some goal. Most systems, however, assume that the user has specified this question well enough for the system to produce an immediate solution in the form of a procedure to carry out. In many cases, the *"how?"* questions of users may require further specification, which means that they will need some help in planning how to accomplish their higher level goal.

"What if?" questions are a special case of *"how?"* questions. These are questions about hypothetical situations, rather than being requests for recipes for performing some task. The answer to this type of question requires that the system be able to simulate the hypothetical case in some way. (Cf. Rich (1982) and Coombs and Alty (1984) for similar suggestions.)

"Why?" questions are usually interpreted as being a form of the question *"what caused...?"* and are viewed as requiring explanations based on the system providing a trace of the steps it went through to produce some result. Other interpretations of *"why?"* questions involve wanting to know why an undesirable result occurred. In other words, there is a difference between the question *"where did I go wrong?"* and *"why is this wrong?"* There is also a need to provide justification following answers to *"how?"* questions. This may imply the need for some form of question-answer dialogue. The problem of providing such dialogues, though important, is outside the scope of this thesis. There are possibilities, however, for supporting these needs, to some extent, without necessitating the use of sophisticated artificial intelligence systems.

A simple way of supporting users in trying to work out why something went wrong is to show them where to look, without necessarily giving any further explanation. In other words, if the problem is due to a simple slip, the user can figure out what to do once it is pointed out. However, there may be cases where the system interprets the command without any problem, but where the output doesn't "look right" (e.g., a formatted letter). In this case, the user needs to know where the unwanted effect originated. For example, users should to be able to point to the area that is wrong and have the system indicate somehow what caused the failure, or at least, what alternatives are possible. The only difference is in whether the system or the user detected the existence of a problem: In both cases, a simple backtrace to the origin of the problem, without any deeper explanation, would suffice.

There are other cases, however, where even when the source of the problem has been pointed out, the user still doesn't understand why something is wrong — perhaps because of some fundamental misconception. The distinction made here is related to the distinction made between "slips" and "mistakes" (cf. Lewis & Norman, in press). Slips (or questions of the type "*where did I go wrong?*") can be handled by drawing the user's attention to the location of the error. Mistakes (or questions of the type "*why is this wrong?*") are cases where there is a "bug" in the user's plan or model (i.e., "in the head"). In these cases some "deeper" explanation is required, since the erroneous action may still "look right" even when pointed to as the source of the problem.

This is the third kind of documentation that was identified in the original analysis described earlier — i.e., the need for what was called full explanation, that includes documentation aimed at a conceptual understanding of the system, both at the level of the novice and at the level of the expert in the domain.

A Study of Users' Questions and Explanations

It was decided to find out from users themselves what kind of conceptual information they might need from documentation. Collecting this kind of information usually requires lengthy protocol studies or interviews. However, it was felt that there were several problems inherent in using traditional protocol analysis. Therefore, in co-operation with other researchers at the Institute for Cognitive Science, University of California, San Diego, some exploratory studies were undertaken, involving an observational technique called Constructive Interaction.

This method is a variant of think-aloud protocol methods, the difference being that Constructive Interaction involves studying two people jointly trying to solve a problem. The method is based on the work of Naomi Miyake (1982). This technique was applied to aspects of the UNIX user interface, by running several pilot studies, in which two participants were videotaped as they discussed some part of the system.

Constructive Interaction is similar to other types of protocol study in which subjects are asked to think aloud while solving some problem. However, while these kinds of studies have proved useful, for example in problem solving, and in studies of novices learning how to use a text editor, it was felt that there were several problems inherent in the method. One of these is the doubtful nature of the connection between verbal reports and mental processes — people tend to use implicit theories about themselves to infer causally their behaviour. It is also possible that having to make a verbal report changes the subject's task significantly and thereby invalidates any generalisation of the findings to more naturalistic situations.

Both of the objections outlined above hinge on the fact that the verbal activity is not intrinsic to the subject of study. However, verbal data can be very useful, as long as one is aware of the factors that influence the way they are produced. In a two-person interaction the communication is not made for the investigator's benefit but for the benefit of the other participant. In other words, in a two-person interaction one may capitalise on the fact that communication takes place between people, rather than having it interfere with the object of study. Another advantage of two-person interactions is that, even if subjects are poor at expressing their knowledge, they are likely to persevere in trying to communicate until their partner does understand, while in traditional protocol analysis the investigator is left with the choice of intervening further with requests for clarification or the choice of making inferences from the protocols.

Miyake's original study (Miyake, 1982) involved pairs of subjects. In each pair neither subject was predominantly an expert, so that the subjects' uncertainty forced them to make explicit what they did understand, and to identify the points at which they lacked understanding. Miyake also had her subjects try out the system (in her case, a sewing machine) to provide evidence and counter-evidence for their hypotheses about its mechanism. From this study, she was able to construct a model of the structure of the subjects' knowledge of the topic in terms of its levels and connections. She was able to see subject's schemas or models changing, to see which questions "upset their understanding", and which explanations were effective in giving them new understanding. This method seems to be useful for extracting partial and tacit knowledge from subjects, where analysis of protocols can show both what utterances really seem to advance the hearer's understanding, as well as what problems are actually raised by learners.

The focus of interest in the study reported here was on exploring the potential of Constructive Interaction for human-computer interaction, and in the conditions under which it might be effective. Therefore, the study involved different topics of discussion, and different mixes of participants, in terms of their prior knowledge of the topic (cf. O'Malley, Draper & Riley, 1984).[6]

One example of the sessions conducted was a simple two-person interaction, concerning a tutorial, in which a novice user, with very little prior experience with computers, was introduced to the system for the first time by someone who had considerable experience in using the system (although the subject was not a programmer). The session revealed some interesting sources of confusion for the novice, not unlike findings reported by Lewis and others in similar studies (cf. Lewis & Mack, 1982; Mack, Lewis & Carroll, 1982). However, the most striking feature of the session was the extent to which the tutorial was interrupted by questions concerning what was, to the tutor, irrelevant to the main theme — how to use the message system. Figure 6.6 shows schematically the interaction in terms of topics discussed (shaded areas) and number of utterances (to scale), where utterances refer to turns taken in the discourse. The total length of the scale is roughly equivalent to one hour. This figure shows the proportion of student-initiated interruptions (dashed lines), as opposed to tutor-initiated digressions.

The figure demonstrates a conflict between what the tutor wanted to convey — which was a basic ability to "login" to the machine and read electronic mail — and the questions of the student, which were largely driven by the screen display. Thus, in order to introduce the user to the message system, the tutor had to spend over

6. Eight pairs of subjects participated in this pilot study.

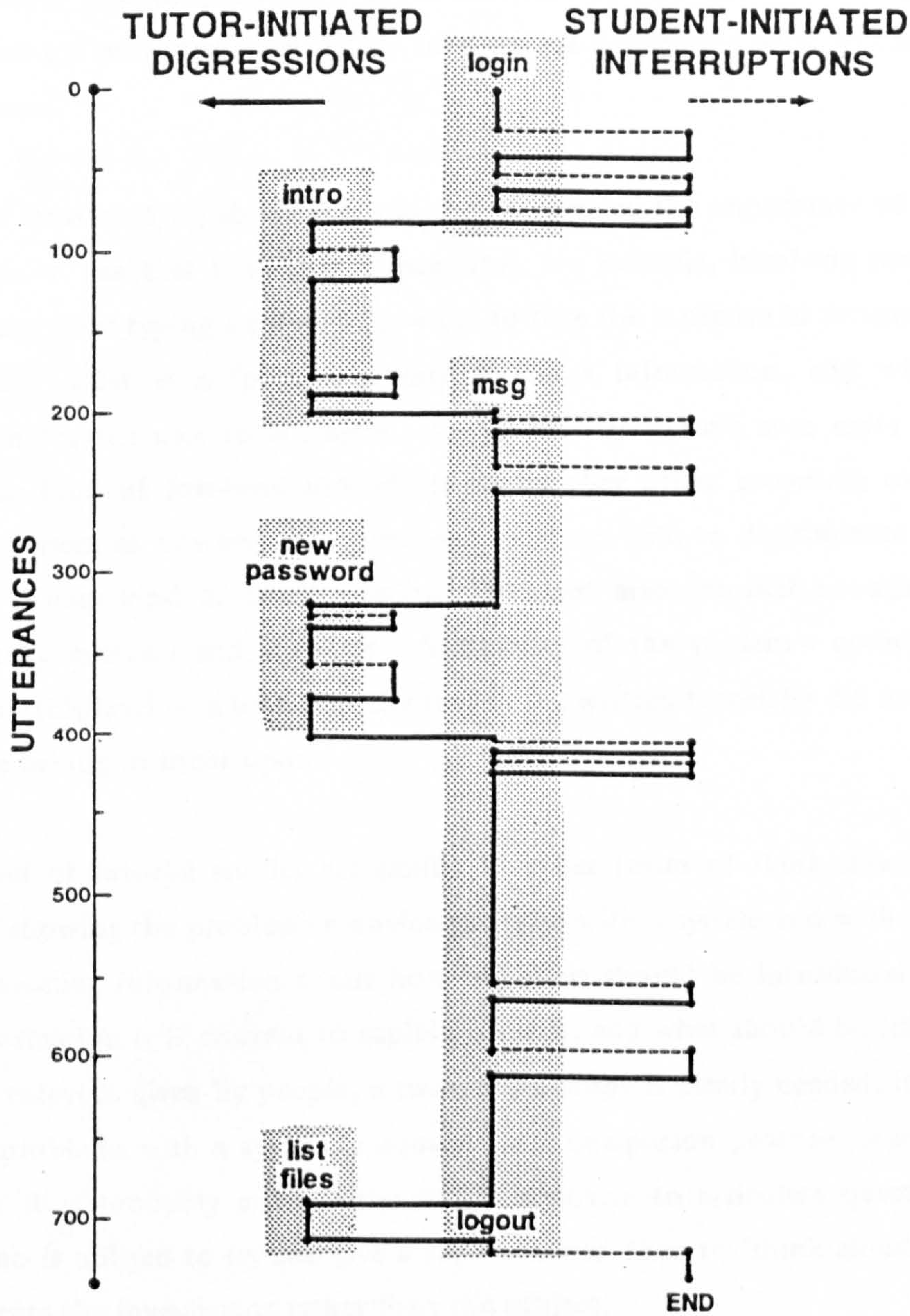

Figure 6.6. Schematic diagram of tutorial session. Shaded areas denote topics discussed; scale indicates number of utterances (i.e., turns taken in the discourse).

half the session explaining various aspects of the system as a result of the user's queries about what was happening on the screen, much of which was in fact unnecessary for learning how to perform the basic task of reading and sending mail. In fact, about 55% of the tutor's utterances were in response to interruptions by the student for an explanation of low-level details concerning what was happening on the screen.

What was most striking about this was that it revealed the importance of low-level procedures to the first time user, procedures, for example, involving pressing the RETURN key after typing a command; when to take the initiative in interacting with the system; what is a "prompt"; what is status information, and what is an instruction for the user to act upon, and so on. Users have even more problems with these kind of low-level procedures where they differ according to context, especially when, as novices, they have not yet been able to discriminate different contexts. These kind of discriminations tend not made explicit enough in most introductory manuals and tutorials. A majority of the student's questions were directed at this level — a level that the tutor (like written tutorials) did not seem to anticipate having to focus upon.

These kind of tutorial studies are similar to other forms of think-aloud protocol study, in showing the problems a novice can have with a system and with a tutorial, and in revealing information about how beginners should be introduced — about what information it is relevant to explain initially, and what should be left out. In studying tutorials given by people, a two-person study is clearly needed; in studying novices' problems with a system, a conventional one-person protocol study does as well, but it is probably much easier to get a novice to articulate questions to a tutor, who is obliged to try and give a useful answer, than to "think aloud" in a way that benefits the investigator rather than the subject.

The kind of session just described may be contrasted with another type conducted in this study, which was felt to be more characteristic of true Constructive Interaction studies. This study was particularly interesting in that it revealed the potential of the method for exploring users' understanding of system concepts. The topic being discussed by the subjects in this study was the UNIX *C-shell* command interpreter, and the rules governing when variable values will get passed to subordinate processes. The two participants both knew the system moderately well, but were not experts. The study revealed that the subjects were seeking different kinds of explanation, based on different kinds of system models. For example, at one point in the session, one of the subjects (A) suggested trying an experiment of explicitly executing an *exec* command from the C-shell. This turned out to be possible because the *C-shell* recognises *exec* as a command. As a control, the subjects typed *lf*, the command for listing files, which listed the files in the current directory, and returned the normal shell prompt. The subjects then typed *exec lf*, which again listed the files, but instead of printing the usual shell prompt, it printed the *login* prompt. In other words, the session had been terminated and they were logged out.

Although this was surprising initially to both subjects, one of them (B) was able to quickly construct an interpretation that fitted his model, so that the experiment served as an illuminating confirmation for him. (What happens is that the *exec* primitive overwrites the calling program with an instance of the new program. Thus in this case, the *C-shell* had overwritten itself with the directory listing program, which had run normally and terminated. The system had detected that there were no more processes associated with that terminal and had prompted for a new login.) This explanation was offered by the first subject (B) to the other, but, although he did not fail to understand it, nor argue directly against it, (A) nevertheless refused to accept it fully, because, in his model of systems, the command interpreter is a part of the operating system that can never die, nor allow itself to be replaced by some other program.

Thus, the acceptance by A of B's explanations was blocked by interference from A's model of how computer systems work, which was imported in the absence of any readily available description in the documentation of how UNIX is structured. This interference was interesting, in that it did not prevent A's apparent "understanding" of B's theory, when described in and of itself: The trouble came in applying it to phenomena that A's model also addressed — that is, logging in, and the lifetime of the command interpreter. In practice A demanded an explanation of the visible events and objects before being able to truly accept a description of the system primitives — although "logically" speaking, an explanation of the visible or surface events is based on these primitives.

This exploratory session revealed a particular gap in the standard documentation, which does contain the actual information about the system primitives, but not in a form that is obviously related to visible system events and objects. What was more interesting were the relationships that were revealed between aspects and stages of understanding in the subjects, which were also brought out in Miyake's original studies, and are relevant to what will succeed as an explanation for users, and hence are relevant to designing documentation.

The first kind of study discussed above — the tutorial session — is not remarkably different from other protocol studies, except that it involves two people rather than one. However, it may be contrasted with the study just described, which characterises much more the technique of Constructive Interaction. The advantage of Constructive Interaction is that the subjects reveal different points of view about a common problem. This difference between them forces them to articulate the rationale behind their hypotheses to each other, and to try and resolve a common ground between them.

The studies reported on here, although exploratory, suggest that Constructive Interaction may be a useful technique to adopt in studying users' conceptualisations of computer systems. In allowing subjects to explore a problem and to develop the solution without prompting or guidance, the investigator is able not only to observe the solution reached by subjects, but also to distinguish the ways of expressing explanations that prove effective for the participants from amongst other less effective attempts. In Constructive Interaction studies one can not only observe what concepts users understand, and what models they have, but one can also observe them changing these ideas and concepts as they move between different levels of understanding the problem, and as they alternately pose and solve problems for their own and the other participants' hypotheses, rather than just observing them react to the situation.

CHAPTER 7

CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCH

In summary, I have argued in this thesis firstly, that it is possible to separate the structural factors contributing to ease of text comprehension and retrieval from those factors pertaining to the content of the material. This argument was supported by the results of a study using scientific texts, where knowledge of the structural form of the text type facilitated comprehension and recall, even when the content of the text was unfamiliar. The results of this study provide support for arguments proposed by the proponents of story grammar theory, and evidence against criticisms of such theories — i.e., that it is the content rather than the form of texts that is responsible for ease of comprehension and recall. The results of this study also provide empirical support for some of the hypotheses made by Kintsch and Van Dijk (1978; Van Dijk & Kintsch, 1983) concerning the role of the superstructural schema in guiding comprehension and recall.

I have also argued that the evidence concerning the facilitative effects of knowledge of the structure in accessing information in conventionalised texts suggests that such structures might be built into and made explicit in less conventionalised texts, in order to improve comprehension and recall, particularly if it contains information unfamiliar to the reader. Furthermore, it has been suggested that readers might be explicitly taught to use the structures provided, in order to obtain access to the information more quickly and easily. However, it has also been pointed out that the structure used to access information depends on the task for which the information is required, and the goals or questions of the reader.

The same argument has been applied to the domain of online computer documentation, where problems of providing appropriate access structures become even more acute than with printed text. The study of the use made of the UNIX online manual demonstrated the need for providing several forms of documentation, depending on the different tasks for which the information is required. The system that was developed as a result of these findings was shown to be successful in supplying a requirement neglected by the existing system. This, together with the evaluation study, demonstrates the importance of paying attention to the needs of users, and to feedback from users themselves, as well as performance or usage data. These studies also suggest that designers of online information systems need to be aware of the importance of structures for accessing information that are flexible with respect to the tasks, goals, intentions, and queries of the user. Finally, it has also been demonstrated that a good deal of information concerning users' conceptualisations, queries, and explanations of computer systems can be gleaned from studies of joint interaction, as well as from studying individuals.

The studies just described are all rather different approaches to the common problem of designing user-oriented documentation that is sensitive not only to the differences in background and experience of users, but also to the fact that at different times in interacting with a system, a given user will need different kinds of information and at different levels, ranging from brief reminders to conceptual explanations. However, whilst modularising the design problem is useful for some purposes, it must be recognised that the problem of providing effective documentation should consider the user-computer-user environment as a whole. The term "information flow" is meant to characterise this latter type of emphasis. Users get their information from a variety of sources, including those not considered much by past research efforts: for example, other users, incidental information, and so on. These remain as gaps in our understanding of the information flow in an interactive computing environment.

There are several issues concerning the design and development of help systems — intelligent or otherwise — that may benefit from investigations of "natural explanation" such as the one discussed in Chapter 6. Firstly, current methods of knowledge elicitation — an important issue for the design of intelligent systems — have several inherent problems that may be overcome by this technique. These problems include interpretation by the investigator, incompleteness and inaccuracy of single-subject reports, interference with tasks, lack of ecological validity, and so forth (cf. Wielinga & Breuker, 1984). In Constructive Interaction studies not only is the pragmatic aspect of communication taken into account, but the technique actually capitalises on it.

Even if subjects are poor at expressing their knowledge they are likely to persevere until the other person does understand, while in one-person studies the investigator is left with the choice of intervening for clarificaton, or of making inferences from the protocols. Secondly, in two-person dialogues it is less likely that important information will be omitted, since each participant is likely to object if the other's explanation is inconsistent. Thirdly, although subjects may have a comparable level of expertise, they are likely to have different points of view about a common problem. This difference between them can motivate the argument and force them to articulate the rationale behind their hypotheses to each other, and to try and resolve a common ground between them. The fact that each person has different points of view and experiences may lead to more complete data. As Goguen, Weiner and Linde (1983) note, if a person is trying to convince another of something his or her task is to provide enough information to allow the other person to follow the argument as a whole. Furthermore, using this technique, the investigator is able not only to observe users' knowledge in terms of what is known, but also to observe the ways of expressing explanations that prove effective.

As Wielinga and Breuker (1984) have pointed out, the purpose of interpreting protocols is to establish a mapping between verbal data and knowledge structures. They have also rightly pointed out the need for research on several epistemological and representational levels, intervening between the linguistic level of discourse structure and the level at which knowledge obtained from the data may be implemented. Future studies of the kind just described should examine two-person dialogues within a similar kind of interpretation framework as that suggested by Wielinga and Breuker, but they should also include an element involving the meta-communicative processes and strategies that direct and control explanation-giving. This level is required for an analysis of the ways in which participants gauge the acceptability of their explanations to each other, the ways of expressing explanations that prove effective, and strategies for making repairs when there is a breakdown in mutual understanding.

Future studies along the lines suggested here can provide valuable contributions toward the development of a cognitive theory of natural explanation, examining such issues as how people model other people's knowledge; how they know what kind of explanation to give, and how to convey it; what are the processes of communication that enable repair when there is a breakdown in mutual understanding; and how people recognise the acceptability of an explanation, thereby contributing to the development of guidelines for the design of online help systems.

As was pointed out in the beginning of this chapter, users tend not to use manuals or other forms of documentation or help but prefer to consult the "local expert" or other users. One of the advantages of going to another person for help is that the user can actually demonstrate the problems in many cases, by showing the other person the error message that just appeared on the screen, or by giving them a printout of the code, and so on. The expert is able to use these concrete examples of the problem, together with what they understand about what the user is trying to achieve, to "diagnose" the problem and, if they cannot provide the solution themselves, suggest where they might get help. Furthermore, when experts can provide advice, they do not always respond directly to the question they were asked (Pollack, 1985). They may provide a plan to achieve another action, or suggest a plan that achieves something different to what is believed the user intended, either because there is no way to do what was intended, or because there is a better way to achieve something similar.

However, the consultation is often more than just a diagnostic process. Successful interactions between users usually involve a process of co-operative problem solving, where the expert tries to assess what the user is failing to understand, and how much is already known, and where both participants are actively involved in formulating what the problem is and what solution is appropriate (Kidd & Cooper, 1983; Kidd, 1984). In natural consultation not only is there a good deal of dialogue devoted to diagnosing the problem, and specifying and refining the user's goals in order to formulate a query, but there is also a negotiative aspect to the acceptability of the solution or explanation (cf. Weiner, 1980; Kidd, 1984). The expert's goal of convincing the user requires some indication that the explanation is acceptable. In ensuring that the explanation is properly adjusted to the knowledge of the user, the expert may explicitly request information about the hearer's degree of knowledge, or even test the hearer's knowledge. In making sure the explanation is acceptable the expert may justify it by giving reasons or by eliminating possible alternative solutions, or by demonstrating with concrete examples (Goguen, Weiner & Linde, 1983).

There are times when the "local expert" is not the most appropriate person to go to for help. Lang et al. (Lang, Lang & Auld, 1981; Lang, Auld & Lang, 1982) studied sources of information for users in a university environment and found that they were at least as likely to ask another user or colleague as a member of the computer centre staff. An important characteristic of the Constructive Interaction study described in Chapter 6 was the way participants would actively generate and test hypotheses by experimentation. Experimenting with the system helps to confirm hypotheses that participants may have as well as eliminating possible alternatives.

I have argued that successful information access requires multiple mechanisms, routes and structures. However, studies of users consulting colleagues and local experts (Lang et al., 1981, 1982; Coombs & Alty, 1984) suggest that users often need help in actually formulating questions to begin with, not simply translating them into the appropriate specification language.

There are two common types of problems that arise in formulating questions, both of which are often specifically addressed in existing help systems: first, knowing the name of the command to type in order to get help; second, knowing the name of the arguments to give. For example, in UNIX, in order to ask the question *"how do I search for a line in a file?"* I have to formulate the questions as *"man grep"*. Even if I know that the command for getting help is *man*, I still may not know which command I should ask about. A simple way of handling this problem is with keyword search facilities and menus. The use of menus obviates the need to know how to ask for help by showing the user what information is available. However, menus have their own set of difficulties, such as a complex hierarchical structure.

However, the problem is more complicated than needing to know what command to type to find information — that is, asking the right question goes beyond the problem of using the right language or terminology. The user often needs help in actually formulating the question, that is, in specifying it to a detailed enough level.

Moreover, this process often involves several steps, not just one. It can be multi-step in at least two senses: going from the general to the specific, where questions are internally generated, and diagnosing problems, where questions are generated externally.

There may be several stages to go through from the first general question to the kind of specific question that can be asked of the system or documentation. For example, users have some general intention, such as figuring out how to format a letter for printing. In this case they have some idea in mind of the eventual product, and they know how to create the actual text (i.e., how to edit a file), but not how to specify how they want it formatted. In order to get help in realising their intention they have to translate it into specific questions so that they can make use of the help facilities to find the information they need. The translation process may involve a one-to-one or a one-to-many mapping — that is, one question may map straight onto another (e.g., *"how do I search for a line in a file?"* translates into *man grep*), or several questions may be involved (e.g., *"how do I print a file on the laser printer?"* translates into *man tbl, man eqn, man ltroff*, and so on).

Norman (in press) points out that there are several stages in going from very general intentions to the level of actually executing commands on the system. The complexity involved in the planning process has also been highlighted by several other studies (e.g., Moran's "external to internal task mapping" model, 1983; "planning nets", Riley & O'Malley, 1984). An implicit corollary of these analyses is that users also need to be supported in formulating and asking questions of the help system that correspond to each of these stages. In other words, each step potentially involves one or more questions to be formulated and asked. In fact observation of the questions generated naturally at each stage might serve to validate these analyses. Such support is needed particularly where questions are generated "internally" — that is, where they are generated from some plan that the user has not yet made explicit.

Although there may be a one-to-one mapping between the initial diagnosis of the problem (e.g., *"I need to make this paper look better"*) and the intention (*"I have to change .pp to .sp"*), several steps may need to be taken to determine the specific problem. For example, the user has used the right formatting commands in the file, at least as far as the system is concerned, but then finds that the formatted version of the letter is not really what was wanted. There may be problems in figuring out exactly what is wrong, when all that's known is that the letter doesn't "look right". Since there was no error as far as the system is concerned (the formatting commands are "correct"), the user somehow has to convey the fact that it wasn't what was wanted — which is difficult to do on most systems. For example, if I went to another person for help, I could point to a printout of the letter and say "this doesn't look right", or "I don't want so much space here on the left side". Someone who knows that there are several alternative formatting commands for obtaining different forms of paragraphs will recognise from the printout that, for example, blocking them rather than indenting them would make it "look better".

There are two points I want to draw from this example: First, there is a concrete result in the form of a visual representation to focus on; second, there are some objective or external criteria for solving the problem. In this case the user doesn't have to specify in any great detail what is wanted, because there are some objective standards for "neat" as opposed to "messy" letters. If the user's question is generated from some external representation, the designer can use "answers first, then questions" (cf. Owen, in press) as a technique, within the basic question-answering paradigm, in order to circumvent the need for the user to know how to formulate questions. This is the approach taken by the RABBIT database system (cf. Tou, Williams, Fikes, Henderson, & Malone, 1982), where the interaction is arranged around presenting examples, which may then be criticised until the right information is found.

In summary, questions may be driven by "internal" factors, such as goals and plans. There is a direct parallel here with the kind of analysis given by Moran's "external to internal task mapping" model (Moran, 1983) — the formulation of specific intentions from general goals and plans is a complex process, involving several stages. Support is needed for questions that arise at each of these stages. One approach to providing help for "internally-generated" questions, especially when users don't know what questions to ask, is to present an overview of the information, but as I have argued earlier, such an approach is undermined by the fact that users' tasks cut across the boundaries formed by any particular method of organising the information.

Questions may also be driven by external objects or events in interacting with the system. Making things visible helps the user identify problems, but it can also help the system identify how it can help the user. The use of examples in the RABBIT system, capitalises on this. What is needed is the ability for both the user and the system to share the same representation of the problem. For example, it would be useful if the user could point to an area of a soft-formatted version of their file, say something about it, and have the system know that what they are pointing to refers to an area in the input file, so that it can highlight where the problem is located.

The reason why users often go to other people for help rather than to manuals or help systems is not simply because they can get answers to their questions. For one thing, the answer that's given is often not a solution that the user can immediately implement, but rather some advice on where to find information from the documentation or help system. For another, the questions users ask are often initially stated at a very general level, so a good deal of dialogue has to go on in order to get to a specific enough level. Furthermore, each stage in this dialogue can lead to many different questions being generated.

The translation that is necessary in order to move from the first general question to the final answer puts additional loads on an already complex planning process. It would be disruptive if the user had to go to a different person for each step in the process of getting a solution to a problem, or for each different question. The apparent preference on the part of users for getting help from other people may be due to a need to have all the steps in the translation process supported, preferably from a single source. Thus, although I argue that several different methods of getting help should be made available to the user, it is important that all of these systems form a well-integrated whole. There should no gaps in the system. That is, the help system ought to be able to cover all the possible types of query, as well as being able to provide the needed information. The user should also be able to go from each subsystem to the next in a smooth and uninterrupted fashion, so that the "chain of dialogue" is maintained.

There is another aspect to the point about the importance of maintaining the chain of dialogue in getting help: One thing that is not usually supported by help systems is the ability, once some piece of information has been found, to refind it at a later time, rather than go through the same process of problem-solving over again. This is especially important when you come across information serendipitously, e.g., by browsing through a manual. You should be able to put some sort of bookmark in at that place. People tend to do this for themselves anyway, in effect creating their own documentation: e.g., paper clips in their manuals, notes stuck on their terminals, little black books of useful information, and so on.

There are two aspects to the issue of self-documentation. The first is the need for dynamic structures — users should be able to (re-)structure the information to suit their own unique purposes. When users make notes from documentation, they filter out irrelevant items, mark important things to remember, and put the information in a unique perspective — one that is unrealistic to expect the designer to anticipate. The second aspect concerns the need for refinding information — users should be able to refind useful information without having to go through the initial process of constructing the query again from scratch. With an online help system there should be some way of recording the use made of help that might allow the system to retrieve information that was previously accessed. (Although the kind of "self-documentation" I discuss here concerns a single user being able to personalise the help system, the same point applies to whole groups of users — cf. Owen, in press.)

One possible method of implementing these ideas is through the use of "hypertext". Hypertext is a term coined by Nelson (1967, 1981) to capture the notion of non-linear or non-sequential text that allows for interactive branching and dynamic display of information. In a hypertext system a user can construct arbitrary links from any point in a document to other points within the same document or in another document. By using the resulting sets of links as "access structures", the user can retrieve information in a dynamically organised fashion. For example, a simple form of hypertext for a book might consist of a table of contents with links to files containing each chapter, each of which contains links to sections. As a reader, you could combine these links in any way you choose to construct your own organisation of the text. Links can also replace footnotes to point to the actual material referenced (cf. Meyrowitz & Van Dam, 1982).

With such a system, users can create links to new files in which they can comment on the information, so that the next time they access the information they also get back their own annotations. Such a system could also maintain histories or backtraces of the interaction, via sets of links connecting the various modules of information. Reactivating these links by running a backtrace would automatically retrieve the modules pointed to by the links. These links would thus provide filters through which to view the information.

The notion of hypertext has been around for some time: Some systems are still in the process of being developed (e.g., Nelson's XANADU system, Nelson, 1981); other systems have been instantiated (e.g., NLS, Englebart & English, 1968; FRESS, Van Dam & Rice, 1971; DATALAND, Negroponte, 1979; DYNABOOK, Weyer, 1982; Feiner, Nagy & van Dam, 1982). For example, Price (1982) describes a system called THUMB (to capture the notion of leafing through a document) that supports the retrieval and maintenance of online documentation, incorporating hypertext notions such as linking, cross-referencing and structuring text passages. Documents are represented as text passages or modules that can be accessed from several perspectives. Users can "keep their place" by use of "checkpoint" and "return" commands. Weyer (1982) also notes the importance of being able to leave annotations in the text, or highlight it, or use some kind of bookmark, as a means of managing and structuring information. He argues that this structure should become part of the accessible information: "Not only should a dynamic book store a history of where the reader as already looked, but it should be capable of returning the reader to previous choice points along an exploratory path or suggesting places where he has not looked. So the patterns of search are themselves information, and should be viewable and modifiable to construct new patterns or descriptions" (Weyer, 1982).

The approach of viewing some of the main problems with help systems as problems concerning access to information does not imply that one has to design the organisational structure of the information a priori. "Task-based" systems are able to maintain the dynamic features characteristic of help that is generated "on the fly", by modularising the information units to a level that allows multiple perspectives and access routes to information, since the modules can be combined in a flexible manner (cf. COUSIN, Hayes & Szekeley, 1983; O'Malley et al., 1983; Smolensky et al., 1984).

An example of a hypertext system that supports user-created documentation is SYNVIEW (Lowe, 1984), which represents the structure of topics in argument form, and makes explicit the strength of evidence for items via a system of weighting opinions from several users. The system can be used to support several applications, including the maintenance and updating of documentation, user-created documentation, and co-authoring or conferencing systems. If a user types a keyword or phrase the system first presents an overview of the topic in outline form, where topics are ranked in decreasing order of importance — i.e., what is important to know for a general understanding of the lead topic. Similar suggestions for the use of hypertext-like systems for authoring and co-authoring have been suggested by Cypher (in press) and Brown (in press). Future research should, in a similar way, exploit more fully the power and flexibility of computer-based systems, so that they become not only useful and usable sources of information, but also so that they may become "idea amplifiers" (Brown, 1984), that extend the potential and creativity of their users.

References

Aaronson, D., & Scarborough, H.S. (1977). Performance theories for sentence coding: Some quantitative models. *Journal of Verbal Learning and Verbal Behavior*, *16*, 277-303.

Abelson, R.P. (1975). Concepts for representing mundane reality in plans. In D.G. Bobrow & A.M. Collins (Eds.), *Representation and Understanding: Studies in Cognitive Science*. New York: Academic Press.

Anderson, J.R. (1976). *Language, Memory, and Thought*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Anderson, J.R. (1981). Effects of prior knowledge on memory for new information. *Memory & Cognition*, *9*, 237-246.

Anderson, J.R. (1983). *The Architecture of Cognition*. Cambridge, Ma.: Harvard University Press.

Anderson, J.R., & Bower, G.H. (1973). *Human Associative Memory*. Washington: Winston.

Anderson, J.R., & Reder, L.M. (1979). An elaborative processing explanation of depth processing. In L.S. Cermak & F.I.M. Craik (Eds.), *Levels of Processing in Human Memory*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Anderson, R.C., & Pichert, J.W. (1978). Recall of previously unrecallable information following a shift in perspective. *Journal of Verbal Learning and Verbal Behavior*, *17*, 1-12.

Anderson, R.C., Reynolds, R.E., Schallert, D.L., & Goetz, E.T. (1977). Frameworks for comprehending discourse. *American Educational Research Journal*, *14*, 367-381.

Anderson, R.C., Spiro, R.J., & Anderson, M.C. (1978). Schemata as scaffolding for the representation of information in connected discourse. *American Educational Research Journal*, *15*, 433-440.

Aristotle, (1960). *Rhetoric* (L. Cooper, trans.). New York: Appleton-Century-Crofts.

Armbruster, B.B., & Anderson, T.H. (1980, February). *The effect of mapping on the free recall of expository text* (Technical Report No. 160). University of Illinois at Urbana-Champaign: Center For the Study of Reading.

Armbruster, B.B., & Anderson, T.H. (1982, October). *Idea-mapping: The technique and its use in the classroom, or, simulating the "ups" and "downs" of reading comprehension* (Reading Education Report No. 36). University of Illinois at Urbana-Champaign: Center For the Study of Reading.

Ausubel, D.P. (1960). The use of advance organizers in the learning and retention of meaningful verbal material. *Journal of Educational Psychology, 51*, 267-272.

Ausubel, D.P. (1968). *Educational Psychology: A Cognitive View.* New York: Holt, Rinehart & Winston.

Ballmer, T.T. (1981). Words, sentences, texts, and all that. *Text, 1*, 163-189.

Bannon, L.J. (In press). Helping users help one another. In D.A. Norman & S.W. Draper (Eds.), *User Centered System Design: New Perspectives on Human-Computer Interaction.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Bannon, L.J., & O'Malley, C.E. (1984). Problems in evaluation of human-computer interfaces: A case study. In B. Shackel (Ed.), *Interact '84: First Conference on Human-Computer Interaction.* Amsterdam: North-Holland.

Bartlett, F.C. (1932). *Remembering: A Study in Experimental and Social Memory.* New York: MacMillan.

Bellezza, F.S., & Bower, G.H. (1982). Remembering script-based text. *Poetics, 11*, 1-12.

Black, J.B., & Bern, H. (1981). Causal coherence and memory for events in narratives. *Journal of Verbal Learning and Verbal Behavior, 20*, 267-275.

Black, J.B., & Bower, G.H. (1979). Episodes as chunks in narrative memory. *Journal of Verbal Learning and Verbal Behavior, 18*, 309-318.

Black, J.B., & Bower, G.H. (1980). Story understanding as problem solving. *Poetics, 9*, 223-250.

Black, J.B., & Wilensky, R. (1979). An evaluation of story grammars. *Cognitive Science, 3*, 213-230.

Bock, M. (1980). Some effects of titles on building and recalling text structures. *Discourse Processes, 3*, 301-311.

Bork, A. (1983). A preliminary taxonomy of ways of displaying text on screens. *Information Design Journal, 3*, 206-214.

Bradshaw, G.H., & Anderson, J.R. (1982). Elaborative encoding as an explanation of levels of processing. *Journal of Verbal Learning and Verbal Behavior, 21*, 165-174.

Bransford, J.D., Barclay, J.R., & Franks, J.J. (1972). Sentence memory: A constructive versus interpretive approach. *Cognitive Psychology, 3*, 193-209.

Bransford, J.D., & Franks, J.J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology, 2*, 331-350.

Bransford, J., & Johnson, M.K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior, 11*, 717-726.

Brewer, W.F. (1980). Literary theory, rhetoric, and stylistics: Implications for psychology. In R.J. Spiro, B.C. Bruce & W.F. Brewer (Eds.), *Theoretical Issues in Reading Comprehension*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Brewer, W.F., & Lichtenstein, E.H. (1981). Event schemas, story schemas, and story grammars. In A.D. Baddeley & J.D. Long (Eds.), *Attention and Performance* (IV). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Britton, B.K., Meyer, B.J.F., Hodge, M.H., & Glynn, S.M. (1980). Effects of the organization of text on memory: Tests of retrieval and response criterion hypotheses. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 620-629.

Britton, B.K., Meyer, B.J.F., Simpson, R., Holdredge, T.S., & Curry, C. (1979). Effects of the organization of text on memory: Tests of two implications of a selective attention hypothesis. *Journal of Experimental Psychology: Human Learning and Memory, 5*, 496-506.

Brown, A.L., & Day, J.D. (1983). Macrorules for summarizing texts: The development of expertise. *Journal of Verbal Learning and Verbal Behavior*, *22*, 1-14.

Brown, A.L., & Smiley, S.S. (1977). Rating the importance of structural units of prose passages: A problem of metacognitive development. *Child Development*, *48*, 1-8.

Brown, J.S. (1984, July). *Idea amplifiers: New kinds of electronic learning environments* (Unpublished paper). Palo Alto, Ca.: Intelligent Systems Laboratory, Xerox PARC.

Brown, J.S. (In press). From cognitive to social ergonomics and beyond. In D.A. Norman & S.W. Draper (Eds.), *User Centered System Design: New Perspectives on Human-Computer Interaction*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Bruce, B., & Newman, D. (1978). Interacting plans. *Cognitive Science*, *2*, 195-233.

Caramazza, A., Grober, E., Garvey, C., & Yates, J. (1977). Comprehension of anaphoric pronouns. *Journal of Verbal Learning and Verbal Behavior*, *16*, 601-609.

Carpenter, P.A., & Just, M.A. (1977). Integrative processes in comprehension. In D. LaBerge & S.J. Samuels (Eds.), *Basic Processes in Reading: Perception and Comprehension*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Cermak, L.S., & Craik, F.I.M. (Eds.), *Levels of Processing in Human Memory*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Chaffin, R. (1979). Knowledge of language and knowledge about the world: A reaction time study of invited and necessary inferences. *Cognitive Science*, *3*, 311-328.

Chang, F.R. (1980). Active memory processes in visual sentence comprehension: Clause effects and pronominal reference. *Memory & Cognition*, *8*, 58-64.

Charolles, M. (1983). Coherence as a principle in the interpretation of discourse. *Text*, *3*, 71-97.

Cherry, L.L. (1982). Writing tools. *IEEE Transactions on Communication*, *30*, 100-105.

Chiesi, H.L., Spilich, G.J., & Voss, J.F. (1979). Acquisition of domain-related information in relation to high and low domain knowledge. *Journal of Verbal*

Learning and Verbal Behavior, 18, 257-273.

Chomsky, N. (1970). Remarks on nominalizations. In R. Jacobs & P. Rosenbaum (Eds.), *Readings in English Transformational Grammar*. Waltham, Ma.: Blaisdell.

Cicourel, A.V. (1975). *Cognitive Sociology*. New York: Academic Press.

Cirilo, R.K. (1981). Referential coherence and text structure in story comprehension. *Journal of Verbal Learning and Verbal Behavior, 20*, 358-367.

Cirilo, R.K., & Foss, D.J. (1980). Text structure and reading times for sentences. *Journal of Verbal Learning and Verbal Behavior, 19*, 96-109.

Clark, H.H., & Clark, E.V. (1968). Semantic distinctions and memory for complex sentences. *Quarterly Journal of Experimental Psychology, 20*, 129-138.

Clark, H.H., & Haviland, S.E. (1977). Comprehension and the Given-New contract. In R.O. Freedle (Ed.), *Discourse Production and Comprehension.* Norwood, N.J.: Ablex.

Clement, D. (1984). *A model of comprehending computer documentation* (Unpublished PhD thesis). University of California, Berkeley.

Cohen, R. (1980). Understanding arguments. *Proceedings of the Conference of the Canadian Society for Computational Studies of Intelligence.*

Coke, E.V., & Koether, M.E. (1983). A study of the match between the stylistic difficulty of technical documents and the reading skills of technical personnel. *The Bell System Technical Journal, 62*, 1849-1863.

Colby, B.N. (1973). A partial grammar of Eskimo folktales. *American Anthropologist, 75*, 645-622.

Collins, A., Brown, J.S., & Larkin, K.M. (1980). Inference in text understanding. In R.J. Spiro, B.C. Bruce & W.F. Brewer (Eds.), *Theoretical Issues in Reading Comprehension.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Coombs, M., & Alty, J. (1984). Expert systems: An alternative paradigm. *International Journal of Man-Machine Studies, 20*, 21-43.

Craik, F.I.M., & Lockhart, R.S. (1972). Levels of processing: A framework for memory research. *Journal of Experimental Psychology: General, 104*, 268-294.

Cypher, A. (In press). The structure of users' activities. In D.A. Norman & S.W. Draper (Eds.), *User Centered System Design: New Perspectives on Human-Computer Interaction.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

De Beaugrande, R. (1980a). *Text, Discourse, and Process.* London: Longman.

De Beaugrande, R. (1980b). Text and discourse in European research. *Discourse Processes, 3*, 287-300.

De Beaugrande, R. (1981a). Linguistic theory and metatheory for a science of texts. *Text, 1*, 113-161.

De Beaugrande, R., & Dressler, W. (1981). *Introduction to Text Linguistics.* London: Longman.

De Kleer, J., & Brown, J.S. (1983). Assumptions and ambiguities in mechanistic mental models. In D. Gentner & A.L. Stevens (Eds.), *Mental Models.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Dixon, P. (1982). Plans and written directions for complex tasks. *Journal of Verbal Learning and Verbal Behavior, 21*, 70-84.

Dooling, D.J., & Lachman, R. (1971). Effects of comprehension and retention of prose. *Journal of Experimental Psychology, 88*, 216-222.

Dooling, D.J., & Mullet, R.L. (1973). Locus of thematic effects in retention of prose. *Journal of Experimental Psychology, 99*, 404-406.

Draper, S.W. (1984). The nature of expertise in Unix. In B. Shackel (Ed.), *Interact '84: First IFIP Conference on Human-Computer Interaction.* Amsterdam: North-Holland.

Dressler, W.U. (1977). (Ed.) *Current Trends in Textlinguistics.* New York: De Gruyter.

Dretske, F.I. (1981). *Knowledge and the Flow of Information.* Boston, Ma.: MIT Press/Bradford Books.

Dretske, F.I. (1983). Precis of *Knowledge and the Flow of Information*. *The Behavioral and Brain Sciences*, *6*, 55-90.

Duffy, T.M., Curran, T.E., & Sass, D. (1983). Document design for technical job tasks: An evaluation. *Human Factors*, *25*, 143-160.

Ehrlich, K. (1980). Comprehension of pronouns. *Quarterly Journal of Experimental Psychology*, *32*, 247-255.

Englebart, D.C., & English, W.K. (1968). A research center for augmenting human intellect. *Proceedings of the AFIPS Fall Joint Computer Conference*, *33*, 395-410.

Enkvist, N.E. (1981). Experiential iconicism in text strategy. *Text*, *1*, 77-111.

Faigley, L., & Meyer, P. (1983). Rhetorical theory and readers' classification of text types. *Text*, *3*, 305-325.

Fass, W., & Schumacher, G.M. (1981). Schema theory and prose retention: Boundary conditions for encoding and retrieval effects. *Discourse Processes*, *4*, 17-26.

Feiner, S., Nagy, S., & Van Dam, A. (1982). An experimental system for creating and presenting interactive graphical documents. *ACM Transactions on Graphics*, *1*, 59-77.

Fikes, R. (1982). A commitment-based framework for describing informal cooperative work. *Cognitive Science*, *6*, 331-347.

Fillmore, C. (1968). The case for case. In Bach & Harms (Eds.), *Universals in Linguistic Theory*. New York: Holt, Rinehart & Winston.

Fischer, G., Lemke, A., & Schwab, T. (1985). Knowledge-based help systems. In L. Borman & W. Curtis (Eds.), *Proceedings of the CHI'85 Conference on Human Factors in Computing Systems*. New York: ACM.

Flesch, R.F. (1948). A new readability yardstick. *Journal of Applied Psychology*, *32*, 221-233.

Fletcher, C.R. (1981). Short-term memory processes in text comprehension. *Journal of Verbal Learning and Verbal Behavior*, *20*, 564-574.

Fletcher, C.R. (1984). Markedness and topic continuity in discourse processing. *Journal of Verbal Learning and Verbal Behavior*, *23*, 487-493.

Fodor, J.A., Bever, T.G., & Garrett, M.K. (1974). In *The Psychology of Language: An Introduction to Psycholinguistics and Generative Grammar*. New York: McGraw-Hill.

Foster, J.J. (1979). The use of visual cues in text. In P.A. Kolers, M.E. Wrolstad & H. Bouma (Eds.), *Processing Visible Language* (Vol. 1, pp. 189-203). New York: Plenum Press.

Frase, L.T. (1969). Paragraph organization of written materials: The influence of conceptual clustering upon the level and organization of recall. *Journal of Educational Psychology*, *60*, 394-401.

Frase, L.T., & Schwartz, B.J. (1979). Typographical cues that facilitate comprehension. *Journal of Educational Psychology*, *71*, 197-206.

Frederiksen, C.H. (1975). Acquisition of semantic information from discourse: Effects of repeated exposure. *Journal of Verbal Learning and Verbal Behavior*, *14*, 158-169.

Fredericksen, J.R. (1981). Understanding anaphora: Rules used by readers in assigning pronominal referents. *Discourse Processes*, *4*, 323-347.

Frisch, A.M., & Perlis, D. (1981). A re-evaluation of story grammars. *Cognitive Science*, *5*, 79-86.

Furnas, G. (1983, October). *Small views of large structures*. Paper presented at the Workshop on Intelligent User Interfaces, Jackson, New Hampshire.

Garnham, A. (1982, March). *Testing psychological theories about inference making* (Unpublished paper). Brighton: University of Sussex.

Garnham, A., Oakhill, J., & Johnson-Laird, P.N. (1982). Referential continuity and the coherence of discourse. *Cognition*, *11*, 29-46.

Garrod, S., & Sanford, A. (1977). Interpreting anaphoric relations: The integration of semantic information while reading. *Journal of Verbal Learning and Verbal Behavior*, *16*, 77-90.

Garrod, S.C., & Sanford, A.J. (1982). Bridging inferences and the extended domain of reference. In J. Long & A. Baddeley (Eds.), *Attention and Performance* (IX). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Gentner, D. (1981). Integrating verb meanings into context. *Discourse Processes, 4*, 349-375.

Gentner, D. (1982). Structure-mapping: A theoretical framework for analogy and similarity. *Proceedings of the 4th Annual Conference of the Cognitive Science Society.*

Geva, E. (1981, July). *Facilitating reading comprehension through flowcharting* (Technical Report No. 211). University of Illinois at Urbana-Champaign: Center For the Study of Reading.

Giora, R. (1983). Segmentation and segment cohesion: On the thematic organisation of the text. *Text, 3*, 155-181.

Givón, T. (1983). (Ed.) *Topic Continuity in Discourse: A Quantitative Cross-Linguistic Study.* Amsterdam: J. Benjamin's.

Glanzer, M., Dorfman, D., & Kaplan, B. (1981). Short-term storage in processing text. *Journal of Verbal Learning and Verbal Behavior, 20*, 656-670.

Glucksberg, S., & McCloskey, M. (1981). Decisions about ignorance: Knowing that you do not know. *Journal of Experimental Psychology: Human Learning and Memory, 7*, 311-325.

Glynn, S.M., & DiVesta, F.J. (1979). Control of prose processing via instructional and typographic cues. *Journal of Educational Psychology, 71*, 595-603.

Goetz, E.T. (1979). Inferring from text: Some factors influencing which inferences will be made. *Discourse Processes, 2*, 179-195.

Goetz, E.T., & Armbruster, B.B. (1980). Psychological correlates of text structure. In R.J. Spiro, B.C. Bruce & W.F. Brewer (Eds.), *Theoretical Issues in Reading Comprehension.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Goguen, J.A., Weiner, J.L., & Linde, C. (1983). Reasoning and natural explanation. *International Journal of Man-Machine Studies, 19*, 521-559.

Goldman, M., Florez, C., & Fuller, G.L. (1981). Factors affecting courteous behavior. *Journal of Social Psychology, 115*, 169-174.

Graesser, A.C. (1978). How to catch a fish: The memory and representation of common procedures. *Discourse Processes, 1*, 72-89.

Graesser, A.C., & Goodman, S.M. (1985). Implicit knowledge, question answering, and the representation of expository text. In B.K. Britton & J.B. Black (Eds.), *Understanding Expository Text: A Theoretical and Practical Handbook for Analyzing Explanatory Text* (pp. 109-171). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Graesser, A.C., Hauft-Smith, K., Cohen, A.D., & Pyles, L.D. (1980). Advanced outlines, familiarity, and text genre on retention of prose. *Journal of Experimental Education, 40*, 281-190.

Graesser, A.C., Higginbotham, M.W., Robertson, S.P., & Smith, W.R. (1978). A natural inquiry into the National Enquirer: Self-induced versus task-induced reading comprehension. *Discourse Processes, 1*, 355-372.

Graesser, A.C., Hoffman, N.L., & Clark, L.F. (1980). Structural components of reading time. *Journal of Verbal Learning and Verbal Behavior, 19*, 135-151.

Graesser, A.C., & Murachver, T. (In press). Symbolic procedures of question answering. In A.C. Graesser & J.B. Black (Eds.), *The psychology of questions*. Hillsdale, N. J.: Lawrence Erlbaum Associates.

Graesser, A.C., Robertson, S.P., & Anderson, P.A. (1981). Incorporating inferences in narrative representations: A study of how and why. *Cognitive Psychology, 13*, 1-26.

Granger, R.H., & Holbrook, J.K. (1983). Perseverers, Recencies and Deferrers: New experimental evidence for multiple inference strategies in understanding. *Proceedings of the 5th Annual Conference of the Cognitive Science Society.*

Green, T.R.G., & Payne, S.J. (1982). The woolly jumper: Typographic problems of concurrency in information display. *Visible Language, 14*, 391-403.

Greenspan, S.L., & Segal, E.M. (1984). *Reference and comprehension: A topic-comment analysis of sentence-picture verification* (Report No. 8403). University of California, San Diego: Institute for Cognitive Science.

Greimas, A.J. (1966). *Semantique Structurale* [Structural Semantics]. Paris: Larousse.

Grimes, J.E. (1975). *The Thread of Discourse*. The Hague: Mouton.

Grimes, J.E. (1982). Reference spaces in text. In S. Allen (Ed.), *Text Processing*. Stockholm: Almquist & Wiksell International.

Grosz, B.J. (1979). *Focusing and description in natural language dialogues* (Technical Note No. 185). Menlo Park, Ca.: SRI International.

Gruber, E.H., Beardsley, W., & Caramazza, A. (1978). Parallel function strategy in pronoun assignment. *Cognition, 6*, 117-133.

Guindon, R., & Kintsch, W. (1984). Priming macropropositions: Evidence for the primacy of macropropositions in the memory for text. *Journal of Verbal Learning and Verbal Behavior, 23*, 508-518.

Haberlandt, K. (1980). Story grammar and reading time of story constituents. *Poetics, 9*, 99-118.

Haberlandt, K., Berian, C., & Sandson, J. (1980). The episode schema in story processing. *Journal of Verbal Learning and Verbal Behavior, 19*, 635-650.

Haberlandt, K., & Bingham, G. (1978). Verbs contribute to the coherence of brief narratives: Reading related and unrelated sentence triples. *Journal of Verbal Learning and Verbal Behavior, 17*, 419-425.

Halliday, M.A.K. (1967). Notes on transitivity and theme in English, Part 2. *Journal of Linguistics, 3*, 199-244.

Halliday, M.A.K. (1982). How is a text like a clause?. In S. Allen (Ed.), *Text Processing*. Stockholm: Almquist & Wiksell International.

Halliday, M.A.K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.

Halloran, S.M. (1978). Technical writing and the rhetoric of science. *Journal of Technical Writing and Communication, 8*, 77-88.

Harris, R.J. (1981). Inferences in information processing. In G.H. Bower (Ed.), *The Psychology of Learning and Motivation* (pp. 81-128). New York: Academic

Press.

Harris, Z. (1952). Discourse analysis. *Language*, *28*, 1-30.

Harris, Z. (1970). *Papers in Structural and Transformational Linguistics*. Dordrecht: D. Reidel.

Hartley, J. (1980). Spatial cues in text. *Visible Language*, *14*, 62-69.

Hartley, J. (1981). Eighty ways of improving instructional text. *IEEE Transactions on Professional Communication*, *24*, 17-27.

Hartley, J. (1982). Information mapping: A critique. *Information Design Journal*, *3*, 51-58.

Hartley, J., Bartlett, S., & Branthwaite, A. (1980). Underlining can make a difference—sometimes. *Journal of Educational Research*, *73*, 218-224.

Hartley, J., & Burnhill, P. (1976). *Textbook Design: A Practical Guide*. Paris: UNESCO.

Hartley, J., Burnhill, P., & Davies, L. (1978). The effects of line length and paragraph denotation on the retrieval of information from prose text. *Visible Language*, *12*, 183-194.

Hartley, J., & Davies, I.K. (1976). Pre-instructional strategies: The role of pre-tests, behavioural objectives and advance organisers. *Review of Research in Education*, *46*, 239-265.

Hartley, J., Kenely, J., Owen, G., & Trueman, M. (1980). The effects of headings on children's recall of prose text. *British Journal of Educational Psychology*, *50*, 304-307.

Hartley, J., & Trueman, M. (1981). The effects of changes in layout and changes in wording on preferences for instructional text. *Visible Language*, *15*, 13-31.

Harweg, R. (1968). *Pronomina und Textkonstitution* [Pronouns and Text Constitution]. Munich: Wilhelm Fink Verlag.

Harweg, R. (1977). Substitutional textlinguistics. In W. Dressler (Ed.), *Current*

Trends in Textlinguistics. New York: Walter de Gruyter.

Harweg, R. (1980). Beginning a text. *Discourse Processes, 3,* 313-326.

Haviland, S.E., & Clark, H.H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior, 13,* 512-521.

Hayes, P.J., & Szekeley, P.A. (1983). Graceful interaction through the COUSIN command interface. *International Journal of Man-Machine Studies, 19,* 285-306.

Hayes-Roth, B., & Thorndyke, P.W. (1979). Integration of knowledge from text. *Journal of Verbal Learning and Verbal Behavior, 18,* 91-108.

Hershberger, W.A., & Terry, D.F. (1965). Typographic cueing in conventional and programmed texts. *Journal of Applied Psychology, 49,* 55-60.

Hobbs, J.R. (1979). Coherence and coreference. *Cognitive Science, 3,* 67-90.

Hobbs, J.R. (1982). Towards an understanding of coherence in discourse. In W.G. Lehnert & M.H. Ringle (Eds.), *Strategies for Natural Language Processing.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Hoey, M. (1983). *On the Surface of Discourse.* London: Allen & Unwin.

Horn, R.E. (1975, January). Information mapping. *Datamation,* 85-87.

Horn, R.E. (1982). Structured writing and text design. In D.H. Jonassen (Ed.), *The Technology of Text: Principles for Structuring, Designing, and Displaying Text* (pp. 341-367). Englewood Cliffs, N.J.: Educational Technology Publications.

Hornby, P.A. (1974). Surface structure and presupposition. *Journal of Verbal Learning and Verbal Behavior, 13,* 530-538.

Iser, W. (1980). Texts and readers. *Discourse Processes, 3,* 327-343.

Jarvella, R.J. (1971). Syntactic processing of connected speech. *Journal of Verbal Learning and Verbal Behavior, 10,* 409-416.

Jarvella, R.J. (1979). Immediate memory and discourse processing. In G.H. Bower (Ed.), *The Psychology of Learning and Motivation* (Vol. 13). New York: Academic Press.

Johnson, N.S., & Mandler, J.M. (1980). A tale of two structures: Underlying and surface forms in stories. *Poetics, 9,* 51-86.

Johnson, R.E. (1970). Recall of prose as a function of the structural importance of the linguistic units. *Journal of Verbal Learning and Verbal Behavior, 9,* 12-20.

Johnson, W., & Kieras, D.E. (1983). Representation-saving effects of prior knowledge in memory for simple technical prose. *Memory & Cognition, 11,* 456-466.

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

Johnson-Laird, P.N., & Wason, P.C. (1977). A theoretical analysis of insight into a memory task. In P.N. Johnson-Laird & P.C. Wason (Eds.), *Thinking: Readings in Cognitive Science.* Cambridge: Cambridge University Press.

Jonassen, D.H. (1981). Information mapping: A description, rationale and comparison with programmed instruction. *Visible Language, 15,* 55-66.

Jonassen, D.H., & Falk, L.M. (1980). Mapping and programming textual materials. *Programmed Learning and Educational Technology, 17,* 20-26.

Jordan, M.P. (1980). Short texts to explain problem-solution structures — and vice versa. *Instructional Science, 9,* 221-252.

Just, M.A., & Carpenter, P.A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87,* 329-354.

Katz, J., & Fodor, J. (1963). The structure of semantic theory. *Language, 39,* 170-210.

Keenan, J.M., Baillet, S.D., & Brown, P. (1984). The effects of causal cohesion on comprehension and memory. *Journal of Verbal Learning and Verbal Behavior, 23,* 115-126.

Kemper, S. (1982). Filling in the missing links. *Journal of Verbal Learning and Verbal Behavior, 21,* 99-107.

Kidd, A.L. (1984, November). *Dialogue and explanation in expert systems.* Position paper for the Joint MMI/IKBS Alvey Workshop, Great Britain.

Kidd, A.L., & Cooper, M.B. (1983, December). Man-machine interface for an expert system. *Proceedings of Expert Systems '83.*

Kieras, D.E. (1974). *Analysis of the effects of word properties and limited reading time in a sentence* Unpublished PhD thesis. University of Michigan.

Kieras, D.E. (1978). Good and bad structure in simple paragraphs: Effects on apparent theme, reading time, and recall. *Journal of Verbal Learning and Verbal Behavior, 17*, 13-28.

Kieras, D.E. (1979). *Modelling reading times in different reading tasks with a simulation model of comprehension* (Technical Report No. 2). Tucson, Az.: University of Arizona.

Kieras, D.E. (1980). Initial mention as a signal for thematic content in technical passages. *Memory & Cognition, 8*, 345-353.

Kieras, D.E. (1981a). The role of major referents and sentence topics in the construction of passage macrostructure. *Discourse Processes, 4*, 1-15.

Kieras, D.E. (1981b). Topicalization effects in cued recall of technical prose. *Memory & Cognition, 9*, 541-549.

Kieras, D.E. (1982a). A model of reader strategy for abstracting main ideas from simple technical prose. *Text, 2*, 47-81.

Kieras, D.E. (1982b, October). *What people know about electronic devices: A descriptive study* (Technical Report No. 12). Tucson, Az.: University of Arizona.

Kieras, D.E. (1985). Thematic processes in the comprehension of technical prose. In B.K. Britton & J.B. Black (Eds.), *Understanding Expository Text: A Theoretical and Practical Handbook for Analyzing Explanatory Text* (pp. 89-107). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Kieras, D.E., & Bovair, S. (1981). *Strategies for abstracting main idea from simple technical prose* (Technical Report No. 10). Tucson, Az.: University of Arizona.

Kieras, D.E., & Bovair, S. (1984). The role of a mental model in learning to operate a device. *Cognitive Science*, *8*, 255-273.

Kieras, D.E., Tibbits, M., & Bovair, S. (1984, February). *How experts and nonexperts operate electronic equipmend from instructions* (Technical Report No. 14). Tucson, Az.: University of Arizona.

Kintsch, W. (1974). *The Representation of Meaning in Memory.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Kintsch, W. (1980). Learning from text, levels of comprehension, or: Why anyone would read a story anyway. *Poetics*, *9*, 87-98.

Kintsch, W., & Keenan, J.M. (1973). Recall of propositions as a function of their position in the hierarchical structure. *Cognitive Psychology*, *5*, 257-274.

Kintsch, W., & Kozminsky, E. (1977). Summarizing stories after reading and listening. *Journal of Educational Psychology*, *69*, 491-499.

Kintsch, W., Kozminsky, E., Streby, W.J., McKoon, F., & Keenan, J.M. (1975). Comprehension and recall of text as a function of content variables. *Journal of Verbal Learning and Verbal Behavior*, *14*, 196-214.

Kintsch, W., Mandel, T.S., & Kozminsky, E. (1977). Summarizing scrambled stories. *Memory & Cognition*, *5*, 547-552.

Kintsch, W., & Van Dijk, T.A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*, 363-394.

Kintsch, W., & Vipond, D. (1979). Reading comprehension and readability in educational practise and psychological theory. In L.G. Nilsson (Ed.), *Perspectives on Memory Research.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Kintsch, W., & Young, S.R. (1984). Selective recall of decision-relevant information from texts. *Memory & Cognition*, *12*, 112-117.

Klare, G.R., Shuford, E.H., & Nichols, W.H. (1958). The relation of format organization to learning. *Educational Research Bulletin*, *37*, 39-45.

Kommers, P. (1984). Web teaching as a design consideration for the adaptive presentation of textual information. In G.C. van der Veer, M.J. Tauber, T.R.G.

Green, & P. Gorny (Eds.), *Readings on Cognitive Ergonomics, Mind and Computers: Proceedings of the 2nd European Conference, September 1984, Gmunden, Austria* (pp. 161-169). Berlin: Springer Verlag.

Kozminsky, E. (1977). Altering comprehension: The effect of biasing titles on text comprehension. *Memory & Cognition, 5*, 482-490.

Kucer, S.B. (1983). Text coherence from a transactional perspective. In J.A. Niles & L.A. Harris (Eds.), *Searches for Meaning in Reading and Language Processing..*

Labov, W., & Waletsky, Y. (1967). Narrative analysis: Oral versions of personal experience. In J. Helm (Ed.), *Essays on the Verbal and Visual Arts.* Seattle: Washington University Press.

Lakoff, G. (1972). Structural complexity in fairy tales. *The Study of Man, 1*, 128-150.

Lang, T., Lang, K.N., & Auld, R. (1981). Support for users of operating systems and applications software. *International Journal of Man-Machine Studies, 14*, 269-282.

Lang, K.N., Auld, R., & Lang, T. (1982). The goals and methods of computer users. *International Journal of Man-Machine Studies, 17*, 375-399.

Langleben, M. (1981). Latent coherence, contextual meanings, and the interpretation of a text. *Text, 1*, 279-313.

Lehnert, W.G. (1978). *The Process of Question-Answering.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Lehnert, W.G. (1981). Plot units and narrative summarization. *Cognitive Science, 4*, 293-331.

Lehnert, W.G., Black, J.B., & Reiser, B.J. (1981). Summarizing narratives. *Proceedings of the 7th International Joint Conference on Artificial Intelligence.*

Lesgold, A.M., & Perfetti, C.A. (1978). Interactive processes in reading comprehension. *Discourse Processes, 1*, 323-336.

Lesgold, A.M., Roth, S.F., & Curtis, M.E. (1979). Foregrounding effects in discourse comprehension. *Journal of Verbal Learning and Verbal Behavior, 18*, 291-308.

Levelt, W.J.M., & Kelter, S. (1982). Surface form and memory in question answering. *Cognitive Psychology*, *14*, 78-106.

Lewis, C., & Mack, R. (1982). *The role of abduction in learning to use a computer system* (Technical Report No. RC 9433 (#41620)). New York: IBM Thomas Watson Research Center.

Lewis, C., & Norman, D.A. (In press). Designing for error. In D.A. Norman & S.W. Draper (Eds.), *User Centered System Design: New Perspectives on Human-Computer Interaction.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Lichtenstein, E.H., & Brewer, W.F. (1980). Memory for goal-directed events. *Cognitive Psychology*, *12*, 412-445.

Longacre, R. (1972). *Hierarchy and Universality of Discourse Constituents in the New Guinea Languages.* Washington, DC: Georgetown University Press.

Lowe, D. (1984). The representation of debate as a basis for information storage and retrieval. *Proceedings of the AFIPS National Computer Conference*, *53*.

Mack, R., Lewis, C., & Carroll, J. (1982). *Learning to use word processors: Problems and prospects* (Technical Report No. RC 9712 (#42887)). New York: IBM Thomas Watson Research Center.

Mandler, G.M., & Dean, P.J. (1969). Seriation: Development of serial order in free recall. *Journal of Experimental Psychology*, *81*, 207-215.

Mandler, J.M. (1978). A code in the node: The use of a story schema in retrieval. *Discourse Processes*, *1*, 14-35.

Mandler, J.M. (1982a). Some uses and abuses of a story grammar. *Discourse Processes*, *5*, 305-318.

Mandler, J.M. (1982b). An analysis of story grammars. In F. Klix, J. Hoffman & E. Van Der Meer (Eds.), *Cognitive Research in Psychology.* Amsterdam: North-Holland.

Mandler, J.M. (1983, August). *Stories: The function of structure.* Invited address to the American Psychological Association, Anaheim, California.

Mandler, J.M. (1984). *Stories, Scripts and Scenes: Aspects of Schema Theory.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Mandler, J.M., & DeForest, M. (1979). Is there more than one way to recall a story?. *Child Development, 50*, 886-889.

Mandler, J.M., & Goodman, M.S. (1982). On the psychological validity of story structure. *Journal of Verbal Learning and Verbal Behavior, 21*, 507-523.

Mandler, J.M., & Johnson, N.S. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology, 9*, 111-191.

Mandler, J.M., & Johnson, N.S. (1980). On throwing out the baby with the bathwater: A reply to Black and Wilensky's evaluation of story grammars. *Cognitive Science, 4*, 305-312.

Manelis, L. (1983). Effects of prose structures on memory. *Discourse Processes, 6*, 403-410.

Manelis, L., & Yekovich, F.R. (1976). Repetitions of propositional arguments in sentences. *Journal of Verbal Learning and Verbal Behavior, 15*, 301-312.

Mayer, R.E. (1979). Can advance organizers influence meaningful learning?. *Review of Educational Research, 49*, 371-383.

McCloskey, M., & Glucksberg, S. (1979). Decision processes in verifying category membership statements: Implications for models of semantic memory. *Cognitive Psychology, 11*, 1-37.

McKoon, G. (1977). Organization of information in text memory. *Journal of Verbal Learning and Verbal Behavior, 16*, 247-260.

McKoon, G., & Ratcliff, R. (1980). The comprehension processes and memory structures involved in anaphoric reference. *Journal of Verbal Learning and Verbal Behavior, 19*, 668-682.

Meyer, B.J.F. (1975). *The Organization of Prose and its Effects on Memory.* Amsterdam: North-Holland.

Meyer, B.J.F. (1977). What is remembered from prose: A function of passage structure. In R.O. Freedle (Ed.), *Discourse Production and Comprehension:*

Advances in Research and Theory. Norwood, N.J.: Ablex.

Meyer, B.J.F. (1984). Text dimensions and cognitive processing. In H. Mandl, N. Stein & T. Trabasso (Eds.), *Learning From Texts.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Meyer, B.J.F. (1985). Prose analysis: Purposes, procedures, and problems (Part 1). In B.K. Britton & J.B. Black (Eds.), *Understanding Expository Text: A Theoretical and Practical Handbook for Analyzing Explanatory Text* (pp. 11-64). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Meyer, B.J.F., Brandt, D.M., & Bluth, G.J. (1980). Use of the top-level sdructure in a text: Key for reading comprehension of ninth-grade students. *Reading Research Quarterly, 16*, 72-103.

Meyer, B.J.F., & Freedle, R.O. (1984). The effects of different discourse types on recall. *American Educational Research Journal, 21*, 121-143.

Meyer, B.J.F., & McConkie, G.W. (1973). What is recalled after hearing a passage?. *Journal of Educational Psychology, 65*, 109-117.

Meyer, B.J.F., & Rice, G.E. (1982). The interaction of reader strategies and the organisation of text. *Text, 2*, 155-192.

Meyrowitz, N., & Van Dam, A. (1982). Interactive editing systems. In J. Nievergelt, G. Coray, J.D. Nicoud & A.C. Shaw (Eds.), *Document Preparation Systems* (pp. 21-132). Amsterdam: North-Holland.

Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits of our capacity for processing information. *Psychological Review, 63*, 81-97.

Miller, J.R. (1981, November). *Effects of semantic and syntactic focus on text production and comprehension.* Paper presented at the 22nd annual meeting of the Psychonomic Society.

Miller, J.R. (1985). A knowledge-based model of prose comprehension: Applications to expository text. In B.K. Britton & J.B. Black (Eds.), *Understanding Expository Text: A Theoretical and Practical Handbook for Analyzing Explanatory Text.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Miller, J.R., & Kintsch, W. (1980). Readability and recall of short prose passages: a theoretical analysis. *Journal of Experimental Psychology: Human Learning and*

Memory, *6*, 335-354.

Miller, J.R., & Kintsch, W. (1981). Knowledge-based aspects of prose comprehension and readability. *Text*, *1*, 215-231.

Mitchell, D.C., & Green, D.W. (1978). The effects of context and content on immediate processing in reading. *Quarterly Journal of Experimental Psychology*, *30*, 609-636.

Miyake, N. (1982). *Constructive Interaction* (CHIP Report No. 113). Center for Human Information Processing, University of California, San Diego.

Miyake, N., & Norman, D.A. (1979). To ask a question one must know enough to know what is not known. *Journal of Verbal Learning and Verbal Behavior*, *18*, 357-364.

Moeser, S.D. (1976). Inferential reasoning in episodic memory. *Journal of Verbal Learning and Verbal Behavior*, *15*, 193-212.

Moran, T.P. (1983). Getting into a system: External-internal task mapping analysis. In A. Janda (Ed.), *Proceedings of the CHI'83 Conference on Human Factors in Computing Systems* (pp. 45-49). New York: ACM.

Morgan, J.L., & Sellner, M.B. (1980). Discourse and linguistic theory. In R.J. Spiro, B.C. Bruce & W.F. Brewer (Eds.), *Theoretical Issues in Reading Comprehension* (pp. 165-200). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Morris, C. (1946). *Signs, Language, and Behavior*. New York: George Braziller.

Murphy, M.D. (1979). Measurement of category clustering in free recall. In C.R. Puff (Ed.), *Memory Organization and Structure* (pp. 51-83). New York: Academic Press.

Negroponte, N. (1979). Books without pages. *IEEE International Conference on Communications*, *IV*, 1-8.

Nelson, T.H. (1967). Getting it out of our system. In G. Schecter (Ed.), *Information Retrieval: A Critical Review* (pp. 191-210). Washington, DC: Thompson Book Company.

Nelson, T.H. (1981). *Literary machines*. Pennsylvania: Swarthmore.

Norman, D.A. (1984). Four stages of user activities. In B. Shackel (Ed.), *Interact '84: Proceedings of the First IFIPS Conference on Human-Computer Interaction.* Amsterdam: North-Holland.

Norman, D.A. (In press). Cognitive engineering. In D.A. Norman & S.W. Draper (Eds.), *User Centered System Design: New Perspectives on Human-Computer Interaction.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Norman, D.A., & Bobrow, D.G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology, 7,* 44-64.

Norman, D.A., & Draper, S.W. (In press). (Eds.) *User Centered System Design: New Perspectives on Human-Computer Interaction.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Olson, G.M., Mack, R.L., & Duffy, S.A. (1981). Cognitive aspects of genre. *Poetics, 10,* 283-315.

O'Malley, C.E., Draper, S.W., & Riley, M.S. (1984). Constructive interaction: A method for studying user-computer-user interaction. In B. Shackel (Ed.), *Interact '84: First IFIP Conference on Human-Computer Interaction.* Amsterdam: North-Holland.

O'Malley, C.E., Smolensky, P., Bannon, L., Conway, E., Graham, J., Sokolov, J., & Monty, M. (1983). A proposal for user centered system documentation. In A. Janda (Ed.), *Proceedings of the CHI '83 Conference on Human Factors in Computing Systems* (pp. 282-285). New York: ACM.

Omanson, R.C. (1982). An analysis of narratives: Identifying central, supportive, and distracting content. *Discourse Processes, 5,* 195-224.

Omanson, R.C., Warren, W.H., & Trabasso, T. (1978). Goals, inferential comprehension, and recall of stories by children. *Discourse Processes, 1,* 355-372.

Owen, D. (In press). Answers first, then questions. In D.A. Norman & S.W. Draper (Eds.), *User Centered System Design: New Perspectives on Human-Computer Interaction.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Perfetti, C.A., & Goldman, S.R. (1974). Thematization and sentence retrieval. *Journal of Verbal Learning and Verbal Behavior, 13,* 70-79.

Perfetti, C.A., & Goldman, S.R. (1975). Discourse functions of thematization and topicalization. *Journal of Psycholinguistic Research, 4*, 257-271.

Petöfi, J. (1971). *Transformationsgrammatik und eine ko-textuelle Texttheoire* [Transformational Grammars and a Co-Textual Text Theory]. Frankfurt: Athenaeum.

Petöfi, J.S., & Reiser, H. (1973). (Eds.) *Studies in Text Grammar*. Dordrecht: Reidel.

Pichert, J.W., & Anderson, R.C. (1977). Taking different perspectives on a story. *Journal of Educational Psychology, 69*, 309-315.

Pike, K. (1967). *Language in Relation to a Unified Theory of the Structure of Human Behavior*. The Hague: Mouton.

Pollack, M.E. (1985). Information sought and information provided: An empirical study of user/expert dialogues. In L. Borman & W. Curtis (Eds.), *Proceedings of the CHI'85 Conference on Human Factors in Computing Systems*. New York: ACM

Pollard-Gott, L., McCloskey, M., & Todres, A.K. (1979). Subjective story structure. *Discourse Processes, 2*, 251-281.

Potts, G.R. (1977). Integrating new and old information. *Journal of Verbal Learning and Verbal Behavior, 16*, 305-320.

Potts, G.R., & Peterson, S.B. (1985). Incorporation versus compartmentalization in memory for discourse. *Journal of Memory and Language, 24*, 107-118.

Poulsen, D., Kintsch, E., Kintsch, W., & Premack, D. (1979). Children's comprehension and memory for stories. *Journal of Experimental Child Psychology, 28*, 379-403.

Price, L.A. (1982). THUMB: An interactive tool for accessing and maintaining text. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-12*(2), 155-161.

Prince, G. (1973). *A Grammar of Stories*. The Hague: Mouton.

Propp, V. (1968). *Morphology of the Folktale*. Austin: University of Texas Press.

Ratcliff, R., & McKoon, G. (1978). Priming in item recognition: Evidence for the propositional structure of sentences. *Journal of Verbal Learning and Verbal Behavior*, *17*, 403-418.

Reder, L.M. (1980). The role of elaboration in the comprehension and retention of prose: A critical review. *Review of Educational Research*, *50*, 5-53.

Reder, L.M. (1982a). Elaborations: When do they help and when do they hurt? *Text*, *2*, 211-224.

Reder, L.M. (1982b). Plausibility judgements versus fact retrieval: Alternate strategies for sentence verification. *Psychological Review*, *89*, 250-280.

Reder, L.M., & Anderson, J.R. (1980). A comparison of texts and their summaries: Memorial consequences. *Journal of Verbal Learning and Verbal Behavior*, *19*, 121-134.

Reichman, R. (1978). Conversational coherency. *Cognitive Science*, *2*, 283-327.

Reiser, B.J., & Black, J.B. (1982). Processing and structural models of comprehension. *Text*, *2*, 225-252.

Rich, E.A. (1982). Programs as data for their help systems. *Proceedings of the AFIPS National Computer Conference*, *51*, 481-485.

Riley, M.S., & O'Malley, C.E. (1984). Planning nets: A framework for analyzing user-computer interactions. In B. Shackel (Ed.), *Interact '84: Proceedings of the First IFIPS Conference on Human-Computer Interaction.* Amsterdam: North-Holland.

Robertson, S., & Black, J. (1983). Planning units in text-editing behavior. In A. Janda (Ed.), *Proceedings of the CHI'83 Conference on Human Factors in Computing Systems* (pp. 217-221). New York: ACM.

Rothkopf, E.Z. (1977). Open discussion on the contributions by Meyer and Carter. In R.C. Anderson, R.J. Spiro & W.E. Montague (Eds.), *Schooling and the Acquisition of Knowledge*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Rothkopf, E.Z. (1982). Adjunct aids and the control of mathemagenic activities during purposeful reading. In W. Otto & S. White (Eds.), *Reading Expository Material* (pp. 109-138). New York: Academic Press.

Rothkopf, E.Z., Fisher, D.G., & Billington, M.J. (1982). Effects of spatial context during acquisition on the recall of attributive information. *Journal of Experimental Psychology: Learning, Memory and Cognition, 8*, 126-138.

Rumelhart, D.E. (1975). Notes on a schema for stories. In D.G. Bobrow & A. Collins (Eds.), *Representation and Understanding.* New York: Academic Press.

Rumelhart, D.E. (1977). Understanding and summarizing brief stories. In D. LaBerge & S.J. Samuels (Eds.), *Basic Processes in Reading: Perception and Comprehension.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Rumelhart, D.E. (1980). A reply to Black and Wilensky. *Cognitive Science, 4*, 313-316.

Rumelhart, D.E., Lindsay, P., & Norman, D.A. (1972). A process model for long-term memory. In E. Tulving & W. Donaldson (Eds.), *Organisation of Memory.* New York: Academic Press.

Rumelhart, D.E., & Norman, D.A. (1978). Accretion, tuning and restructuring: Three modes of learning. In J.W. Cotton & R.L. Klatsky (Eds.), *Semantic Factors in Cognition.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Schallert, D.L. (1976). Improving memory for prose: The relationship between depth of processing and context. *Journal of Verbal Learning and Verbal Behavior, 15*, 621-632.

Schank, R.C. (1975). The structure of episodes in memory. In D.G. Bobrow & A. Collins (Eds.), *Representation and Understanding.* New York: Academic Press.

Schank, R.C., & Abelson, R. (1977). *Scripts, Plans, Goals and Understanding.* Hillsdale N.J.: Lawrence Erlbaum Associates.

Scharer, L.L. (1983, July). User training: Less is more. *Datamation*, 175-182.

Schwarz, M.N.K., & Flammer, A. (1981). Text structure and title: Effects on comprehension and recall. *Journal of Verbal Learning and Verbal Behavior, 20*, 61-66.

Shebilske, W.L., & Reid, L.S. (1979). Reading eye movements, macro-structure and comprehension processes. In P.A. Kolers, M.E. Wrolstad & H. Bouma (Eds.), *Processing Visible Language* (Vol. 1, pp. 97-110). New York: Plenum.

Shebilske, W.L., & Rotondo, J.A. (1981). Typographical and spatial cues that facilitate learning from textbooks. *Visible Language, 15*, 41-54.

Sidner, C.L. (1983). Focusing and discourse. *Discourse Processes, 6*, 107-130.

Singer, M. (1979). Processes of inference during sentence encoding. *Memory & Cognition, 7*, 192-200.

Singer, M. (1980). The role of case-filling inferences in the coherence of brief passages. *Discourse Processes, 3*, 185-201.

Smith, E.E., & Goodman, L. (1982, July). *Understanding instructions: The role of explanatory material* (Technical Report No. 5088). Bolt, Beranek & Newman Inc..

Smith, E.E., Spoehr, K.T., & Collins, A. (1982, October). *Basic processes and individual differences in understanding and using instructions* (Proposal No. P82-ISD-45A). Bolt, Beranek & Newman Inc..

Smith, R.N., & Frawley, W.J. (1983). Conjunctive cohesion in four English genres. *Text, 3*, 347-374.

Smolensky, P., Monty, M.L., & Conway, E. (1984). Formalizing task descriptions for command specification and documentation. In B. Shackel (Ed.), *Interact '84: First IFIPS Conference on Human-Computer Interaction.* Amsterdam: North-Holland.

Spilich, G.J., Vesonder, G.T., Chiesi, H.L., & Voss, J.F. (1979). Text processing of domain-related information for individuals with high and low domain knowledge. *Journal of Verbal Learning and Verbal Behavior, 18*, 275-290.

Spiro, R.J. (1977). Remembering information from text: The "State of Schema" approach. In R.C. Anderson, R.J. Spiro & W.E. Montague (Eds.), *Schooling and the Acquisition of Knowledge*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Spiro, R.J. (1980). Accommodative reconstruction in prose recall. *Journal of Verbal Learning and Verbal Behavior, 19*, 84-95.

Spiro, R.J., & Esposito, J. (1981). Superficial processes of explicit inferences in text. *Discourse Processes, 4*, 313-322.

Spoehr, K.T., Morris, M.E., & Smith, E.E. (1983, November). *Comprehension of instructions for operating devices.* Paper presented at the 24th meeting of the Psychonomic Society, San Diego, California.

Steffenson, M.S., Jogdeo, C., & Anderson, R.C. (1978). *A cross-cultural perspective on reading comprehension* (Technical Report No. 97). University of Illinois at Urbana-Champaign: Center for the Study of Reading.

Stein, N.L. (1982). What's in a story: Interpreting the interpretations of story grammars. *Discourse Processes, 5*, 319-335.

Stein, N.L., & Glenn, C.G. (1979). An analysis of story comprehension in elementary school children. In R. Freedle (Ed.), *New Directions in Discourse Processing.* Norwood, N.J.: Ablex.

Stein, N.L., & Nezworski, T. (1978). The effects of organisation and instructional set on story memory. *Discourse Processes, 1*, 177-194.

Stein, N.L., & Policastro, M. (1984). The concept of a story: A comparison between children's and teachers' viewpoints. In H. Mandl, N.L. Stein & T. Trabasso (Eds.), *Learning and Comprehension of Text.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Sticht, T. (1972). Learning by listening. In J.B. Carroll & R.O. Freedle (Eds.), *Language Comprehension and the Acquisition of Knowledge.* New York: Winston.

Sticht, T.G. (1977). Comprehending reading at work. In M.A. Just & P.A. Carpenter (Eds.), *Cognitive processes in comprehension.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Swartout, W.R. (1983). XPLAIN: A system for creating and explaining expert consulting programs. *Artificial Intelligence, 21*, 285-325.

Tanenhaus, M.K., & Seidenberg, M.S. (1981). Discourse context and sentence perception. *Discourse Processes, 4*, 197-220.

Taylor, H.G. (1982). Age differences in peripheral letter perception. *Journal of Experimental Psychology: Human Perception and Performance, 8*, 106-112.

Thorndyke, P.W. (1976). The role of inferences in discourse comprehension. *Journal of Verbal Learning and Verbal Behavior, 15*, 437-446.

Thorndyke, P.W. (1977). Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, *9*, 77-110.

Thorndyke, P.W., & Yekovich, F.R. (1980). A critique of schema-based theories of human story memory. *Poetics*, *9*, 23-49.

Tierney, R.J., & Mosenthal, J.H. (1981). *The cohesion concept's relationship to the coherence of text* (Technical Report No. 221). University of Illinois at Urbana-Champaign: Center for the Study of Reading.

Todorov, T. *Textes des Formalistes Russes*. Paris: Seuil.

Tou, F.N., Williams, M.D., Fikes, R., Henderson, A., & Malone, T. (1982). RABBIT: An intelligent database assistant. *Proceedings of the Conference of the American Association for Artificial Intelligence*, 314-318.

Toulmin, S.E. (1958). *The Uses of Argument*. Cambridge: Cambridge University Press.

Trabasso, T., Secco, T., & Van Den Broek, P. (1984). Causal cohesion and story coherence. In H. Mandl, N.L. Stein, & T. Trabasso (Eds.), *Learning and Comprehension of Text*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Trabasso, T., & Van Den Broek, P. (1984). *Causal thinking and story understanding* (Unpublished paper). University of Chicago.

Tulving, E., & Thomson, D.M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*, 352-373.

Van Dam, A., & Rice, D.E. (1971). On-line text editing: A survey. *Computing Surveys*, *3*, 93-114.

Van Dijk, T.A. (1972). *Some Aspects of Text Grammars*. The Hague: Mouton.

Van Dijk, T.A. (1975). *Recalling and summarizing complex discourse* (Unpublished paper). University of Amsterdam.

Van Dijk, T.A. (1977a). *Text and Context: Explorations in the Semantics and Pragmatics of Discourse*. London: Longman.

Van Dijk, T.A. (1977b). Perspective paper: Complex semantic information processing. In D.E. Walker, H. Karlgren & M. Kay (Eds.), *Natural Language in Information Science* (pp. 127-163). Stockholm, Sweden: Skriptor.

Van Dijk, T.A. (1977c). Semantic macrostructures and knowledge frames in discourse comprehension. In M.A. Just & P.A. Carpenter (Eds.), *Cognitive Processes in Comprehension.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Van Dijk, T.A. (1979). Relevance assignment in discourse comprehension. *Discourse Processes, 2*, 113-126.

Van Dijk, T.A. (1980a). *Macrostructures.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Van Dijk, T.A. (1980b). Story comprehension: An introduction. *Poetics, 9*, 1-21.

Van Dijk, T.A. (1982). Relevance in text and context: Discussion of Joseph E. Grimes' preprint "Context structure patterns". In S. Allen (Ed.), *Text Processing.* Stockholm: Almquist & Wiksell International.

Van Dijk, T.A., & Kintsch, W. (1983). *Strategies of Discourse Comprehension.* New York: Academic Press.

Van Dijk, T.A., & Petöfi, J.S. (1977). (Eds.) *Grammars and Description.* New York: De Gruyter.

Vesonder, G.T. (1979). *The role of knowledge in processing experimental reports* (Unpublished PhD thesis). University of Pittsburgh.

Vipond, D. (1980). Micro- and macroprocesses in text comprehension. *Journal of Verbal Learning and Verbal Behavior, 19*, 276-296.

Voss, J.F. (1984). On learning and learning from text. In H. Mandl, N.L. Stein & T. Trabasso (Eds.), *Learning and Comprehension of Text.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Voss, J.F., & Bisanz, G. (1985). Knowledge and the processing of narrative and expository texts. In B.K. Britton & J.B. Black (Eds.), *Understanding Expository Text: A Theoretical and Practical Handbook for Analyzing Explanatory Text* (pp. 173-198). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Voss, J.F., Vesonder, G.T., & Spilich, G.J. (1980). Text generation and recall by high-knowledge and low-knowledge individuals. *Journal of Verbal Learning and Verbal Behavior*, *19*, 651-667.

Walker, C.H., & Meyer, B.J.F. (1980a). Integrating different types of information in text. *Journal of Verbal Learning and Verbal Behavior*, *19*, 263-275.

Walker, C.H., & Meyer, B.J.F. (1980b). Integrating information from text: An evaluation of current theories. *Review of Educational Research*, *50*, 421-437.

Waller, R.H.W. (1979). Typographic access structures for educational texts. In P.A. Kolers, M.E. Wrolstad, & H. Bouma (Eds.), *Processing Visible Language* (Vol. 1). Plenum Press.

Waller, R. (1982). Text as diagram: Using typography to improve access and understanding. In D.H. Jonassen (Ed.), *The Technology of Text: Principles for Structuring, Designing, and Displaying Text* (pp. 137-166). Englewood Cliffs, N.J.: Educational Technology Publications.

Warren, W.H., Nicholas, D.W., & Trabasso, T. (1979). Event chains and inferences in understanding narratives. In R.O. Freedle (Ed.), *New Directions in Discourse Processing* (Vol. 2). Norwood, N.J.: Ablex.

Weaver, P.A., & Dickinson, D.K. (1982). Scratching below the surface structure: Exploring the usefulness of story grammars. *Discourse Processes*, *5*, 225-243.

Webber, B.L. (1980). Syntax beyond the sentence: Anaphora. In R.J. Spiro, B.C. Bruce & W.F. Brewer (Eds.), *Theoretical Issues in Reading Comprehension.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Weyer, S.A. (1982). The design of a dynamic book for information search. *International Journal of Man-Machine Studies*, *17*, 87-107.

Wielinga, B.A., & Breuker, J.A. (1984). Interpretations of verbal data for knowledge acquisition. In T. O'Shea (Ed.), *ECAI-84: Advances in Artificial Intelligence.* Amsterdam: North-Holland.

Weiner, J.L. (1980). BLAH, a system which explains its reasoning. *Artificial Intelligence*, *15*, 19-48.

Wilensky, R. (1983). Story grammars versus story points. *The Behavioral and Brain*

Sciences, 6, 579-623.

Woods, D.D. (1984). Visual momentum: A concept to improve the cognitive coupling of person and computer. *International Journal of Man-Machine Studies, 21*, 229-244.

Wright, P. (1977). Presenting technical information: A survey of research findings. *Instructional Science, 6*, 93-134.

Wright, P. (1978). Feeding the information eaters: Suggestions for informing pure and applied research. *Instructional Science, 7*, 249-312.

Wright, P. (1979). Usability: The criterion for designing written information. In P.A. Kolers, M.E. Wrolstad & H. Bouma (Eds.), *Processing Visible Language* (Vol. 2, pp. 183-205). New York: Plenum Press.

Wright, P. (1983). Manual dexterity: A user-oriented approach to creating computer docmentation. In A. Janda (Ed.), *Proceedings of the CHI'83 Conference on Human Factors in Computing Systems* (pp. 11-18). New York: ACM.

Yekovich, F.R., & Thorndyke, P.W. (1981). An evaluation of alternative functional models of narrative schema. *Journal of Verbal Learning and Verbal Behavior, 20*, 454-469.

Yekovich, F.R., & Walker, C.H. (1978). Identifying and using referents in sentence comprehension. *Journal of Verbal Learning and Verbal Behavior, 17*, 265-277.

Yekovich, F.R., Walker, C.H., & Blackman, H.S. (1979). The role of presupposed and focal information in integrating sentences. *Journal of Verbal Learning and Verbal Behavior, 18*, 535-548.

APPENDIX A

ORIGINAL MATERIAL

Text1

(Title) *Factors affecting courteous behaviour.*[1]

Altruistic actions, generally referred to as prosocial behaviour, involve activities where one person aids or assists another, but receives no obvious reward for his help. Studies in this area have investigated helping behaviour, such as aiding a seizure victim, protecting property from theft, or offering assistance to a motorist with a flat tire. There are, however, behaviours which aid or help others but are less dramatic, where the cost of carrying them out is considerably less, which involve less time, and where not engaging in the behaviour has less consequence for another individual. Examples include informing someone that they have forgotten to turn off the lights of their car. *These behaviours would be regarded as accommodating, courteous, or thoughtful and an individual doing these acts would be seen as well-mannered, polite, and considerate.*

Researchers investigating prosocial behaviour have found that it is possible for a person to influence the likelihood of receiving help. Baer, Goldman, and Junhnke have shown that if person A had previously had a brief verbal contact with B, person B would be more apt to help A than if this initial contact had not occurred. Boice and Goldman reported that a victim would obtain more help if he used a victim-oriented rather than a target-oriented request.

The norm of reciprocity theory states that people should help those who have helped them. If a favour has been extended to an individual, he will be motivated to return that favour. Conversely, if an individual has not been pleasant to another, the latter will not go out of his way to be pleasant to the former. Support for the norm of reciprocity has been presented by several studies. *In addition, equity theory states that interacting individuals seek to balance their rewards and costs.* The basic proposition states that a person will attempt to maintain proportionality between inputs and outcomes.

On first glance it would appear that both the norm of reciprocity and equity theory would make similar behavioural predictions. For example, if A has helped B, equity theory should suggest that B has been rewarded at no cost to himself. B would thus be motivated to re-establish equity by helping A at the first opportunity. The norm of reciprocity would also suggest that if A has helped B, B would also be likely to reciprocate by helping A. Thus, both theories would imply that A helping B would result in B helping A.

It is possible, however, to set up a situation in which the two theories would make alternate predictions. Levanthal, Weiss, and Long have proposed that equity motivation is induced

---

1. I am grateful to Dr. M. Goldman, University of Missouri, for permission to use this material (cf. Goldman, Florez & Fuller, 1981). Sentences or phrases in italics denote material used in the experiment in Chapter 5.

whenever an individual experiences a discrepancy regardless of the reason for this discrepancy between inputs and outcomes: i.e., inequity can be the result of chance or deliberate intention. On the other hand, reciprocity motivation should be induced only when the discrepancy between input and outcome has been produced intentionally. *Thus, if A helps B unintentionally, equity theory would predict that B would be likely to help A, but the norm of reciprocity would predict that B would be likely to help A only if he thought A initially helped him intentionally.*

Social contact theory also appears to be related to helping behaviour. Zajonc has demonstrated that exposure to an object will induce a positive attitude towards that object. *Thus, an individual having increased contact with another person would be more likely to help that person than if little contact had occurred.*

If the norm of reciprocity is valid, then the intentional courtesy should produce more frequent courteous behaviour than no courtesy. Equity theory would also predict this. Further, equity theory would imply that courteous behaviour would occur to the same extent in both the unintentional and the intentional conditions, while the norm of reciprocity would predict that more courteous behaviour would occur in the intentional than in the unintentional condition.

The present study was designed to test these hypotheses. It also examined whether males more frequently than females engaged in courteous behaviour, since previous research has shown that females are less prosocial than males.

Eighty-four adult males and 84 adult females using a parking lot adjacent to a centrally located shopping center served unwittingly as Ss. A 2 (sex) X 4 (door holding behaviour: control, unintentional, contact, and intentional) experimental design was employed.

A male confederate (C) dressed similarly to the S population waited for an S to park his car and walk toward the exit. To exit from the parking lot required walking to an exit door, opening the door and proceeding through a 15-yard corridor, and finally opening a second door located at the opposite end of the corridor.

A trial was begun when the C saw an S walking alone toward the exit door. Data were collected during five two-hour weekday sessions. *Ss were randomly assigned to the four treatments* and one rotation of the four treatments was completed before the next rotation was run. In each rotation, the treatments were randomly ordered.

In the "control" treatment, the C preceded the S to the exit door with sufficient time to allow him to open the door and have it shut behind him, requiring that the S open the door himself. In the corridor the C carried out a delay tactic (bending down and adjusting his socks). This permitted the S to arrive at the second door with the C following behind.

In the "unintentional" treatment the C preceded the S to the exit door and holding the door open proceeded to tie his shoe lace. He took no notice of the S as the S walked through the open door. The C then followed the S to the second door.

In the "contact" treatment, the C preceded the S to the exit door and held the door open, allowing the S to pass through. As soon as the S had passed through the door, the C said, "Excuse me, can you tell me where he Embassy Theater is located?" The theater was located a block from the parking lot. *After the S responded, the C thanked him and followed behind the S to the second door.*

In the "intentional" treatment, the C again preceded the S to the exit door. He held the door open, allowed the S to walk through, and then followed the S to the second door.

In all four conditions the C was approximately 10 feet behind when the S opened the second door. *The dependent variable measure was whether, or not, the S held the door open for the C.*

The courteous response of the S consisted of holding the door open for the C who was following behind. An S holding the door open for the C was scored 1, and an S not holding the door open for the C was scored 0.

Lunney has shown that the analysis of variance test is appropriate for dichotomous data when the sample population (means) are between .20 and .80 and the sample sizes are reasonably large (for error term, $df > 20$). These conditions were satisfied in the present study.

The analysis of variance showed the mean courteous behaviour for males ($M = .48$) *to be significantly greater than that for females* ($M = .33$), $F(1, 160) = 3.69$, $p < .06$. *The different door holding treatments yielded a significant difference in the rate of courteous behaviour,* $F(3, 160) = 3.35$, $p < .05$. *The interaction was nonsignificant.*

The mean for the control ($M = .33$), *for the unintentional* ($M = .29$), *and for the contact* ($M = .40$) *treatments, when compared with each other, yielded nonsignificant results.* Differences between these three means were all well within chance variations. *However, the mean for the intentional treatment* ($M = .60$) *was significantly greater than the means for each of the other treatments* ($p < .05$ for all comparisons).

The present study found that it is possible for an individual to influence the likelihood that another will behave in a more courteous manner toward him. For courteous behaviour, the golden rule appears to be intentionally do unto others and they will do unto you.

Previous studies have reported that American females tended to be less prosocial than males. The results here also found that females engaged in less courteous behaviour than males.

Since the intentional treatment induced significantly more courteous behaviour than did the control or unintentional treatments, and the unintentional treatment did not produce significantly more courteous behaviour than the control treatment, the results support the norm of reciprocity and do not support equity theory. For courteous behaviour to be reciprocated by an individual, that individual needs to know that he was intentionally helped.

But intentional help must occur under specific conditions. In the contact treatment the S received intentional help, and this was followed by additional contact with the C asking for directions. Courteous behaviour did not increase above that in the control treatment. The norm of reciprocity could also be used in this instance to account for the absence of increased door-holding behaviour. Thus, the C acted courteously, holding open the door for the S, and the S then reciprocated by answering the C's questions. Any obligation that the S felt toward the C had been removed by his responding to the C's request.

Another reason why the contact treatment did not produce increased courteous behaviour might have been that the S was not certain that the C held the door open simply to be *polite. Since the C followed the door holding with a request for information, the door holding could have been* seen as incurring favour from someone whom you wish to ask for help. *Thus the S could have understood that C was holding the door open to reciprocate for the help for which the C was about to ask.*

Text2

(Title) *Age differences in peripheral letter perception*[2]

A number of developmental investigations lend support for the possibility that younger children may be less able than older children or adults to discriminate stimuli in the visual periphery. These include studies showing younger subjects to be less capable than older subjects of both visual search and peripheral recognition of brief exposures.

Whether these findings reflect actual age differences in peripheral visual capacity, however, is as yet unclear. The better performance of the older subjects on visual search tasks may reflect more efficient use of what is seen in peripheral vision, rather than better resolution within the visual system per se. Similarly, because most tachistoscopic studies have failed to compare visual capacity at a given distance into the periphery with that at the fixation point or at another peripheral distance, these results may stem from developmental trends in visual processing not specific to peripheral vision. An exception to this is provided by Fisher and Lefton, who examined age differences in speed of letter matching for pairs of letters presented at various peripheral locations and found a trend towards greater age differences at greater distances into the periphery. Unfortunately, this specific finding was part of a more complex interaction effect and was not reported as reliable in itself. An additional reason to question age differences in peripheral discrimination skill is provided by the failure of some recent studies to document them.

Nevertheless, no study has as yet explored this issue using a method that might optimise any potential differences. The present study was designed to do so by comparing age differences in threshold discriminability for letter arrays presented at the fixation point versus to the left and right of fixation. If indeed there are developmental changes in peripheral discrimination capacities above and beyond those due to changes in overall processing ability (i.e., those reflecting processes of general developmental relevance such as motivation, attention, and memory), *age differences should be greater for stimuli exposed in either visual half field than for stimuli appearing at the center of fixation.*

A further purpose was to explore other factors that might contribute to a greater ability of older as opposed to younger subjects to discriminate peripheral letter arrays. To determine whether age differences in peripheral vision might be introduced by a greater ability of older subjects to interpret cues received in peripheral vision, as suggested by Cohen and Haith, *letter arrays forming both words and nonwords were used. If older subjects make better use than younger subjects of the partial cues available only in words,* (e.g., sequential letter constraints), *these subjects should have a special advantage in discriminating them in peripheral vision. To investigate any further age advantage conferred to letters in the right half field by age trends in cerebral dominance and/or left-to-right postexposural scanning, letter arrays were presented to the left and right half fields in vertical as well as horizontal (normal) orientation. A greater age difference in peripheral letter discriminability for right half field relative to left half field arrays for both horizontal and vertical orientations would support the relevance of cerebral dominance. To the extent that any enhancement of age differences in the right half field relative to the left half field is greater for horizontal than for vertical arrays, a*

---

2. I am grateful to Dr. H.G. Taylor, Child Development Unit, Children's Hospital of Pittsburgh, for permission to use this material (cf. Taylor, 1982).

hypothetical left-to-right postexposural scanning mechanism would also bear consideration.

Subjects were 96 eight- and nine-year-old children (M age = 9.2 years) *and 96 university students* (M age = 21.5 years), *all right-handers.* Both age groups contained an equal number of males and females. Only subjects without significant visual problems were included (judged from a brief visual scanning using a Bausch & Lomb Orthorator). The university students received credits in partial fulfillment of elective laboratory requirements for general psychology courses. The children, who were given financial reimbursement for participating, were obtained from summer playgrounds in the Iowa City, Iowa, community and thus represented a broad sample from a largely middle-class setting. No specific criteria other than age, willingness to participate, and parental consent were used in obtaining the child sample. Nevertheless, reading ability data were collected on nearly half of the children, and these data may provide at least a rough indication of academic competency. Specifically, subsequent to the experimental session, children who received words as stimuli were asked to read a list of the words orally. Although one child was not examined in this regard, only 6 of the remaining 47 children were unable to read all of the words, and these 6 children misread only one word each. No children were excluded from the sample on the basis of whether or not they could read the list, but this observation suggests at least a minimal level of literacy among child participants.

Twenty-four children and 24 adults (half in each age group being of each sex) *were assigned to each of four stimulus conditions: horizontal words, vertical words, horizontal nonwords, and vertical nonwords. This resulted in an experimental design consisting of the within-subject factor, position of array* (left half field, center, right half field) *, and of the four between-subject factors, type of array* (word vs. nonword), *orientation of array* (horizontal vs. vertical), *age* (children vs. adults), *and sex.* The variable of sex was included primarily on a precautionary basis. Although clear sex differences were not predicted, neither could they be ruled out.

Stimuli consisted of 24 words and 24 nonwords, all containing four different letters. The words were of high readability and were familiar to the children as well as to the adults. Nonwords were constructed from the total set of letters in the words, with the restriction that letter sequences not include those commonly found in English words. *All arrays consisted of capital letters typed in black on a white field and presented in a Gerbrands two-field tachistoscope* (Model T-1C). Arrays were viewed binocularly, one at a time, in a field with an illuminance of approximately 17 cd/m. Their midpoints fell at the center of the field, 4 of visual angle to the left of center, or 4 to the right of center, with horizontal arrays subtending 2 48' from end to end and vertical ones subtending about the same angle from top to bottom. *At all times, except during stimulus presentation, a fixation dot was exposed in the center of the viewing field.*

On each trial the subject was told to look directly at the center dot when the experimenter said "look". After stimulus exposure the subject was asked to report four different letters. Six practise arrays were presented initially, all for 149 + 2 msec and at the fixation point. *The remaining 18 test arrays were divided into three groups of 6, with each group assigned to a different one of three visual field positions* — the assignment of arrays to positions being counterbalanced within each Age X Sex X Type X Orientation subgroup. *Test trials were administered in blocks of 18, with each of the 18 arrays presented at its assigned position once per block.* The order in which arrays appeared within each block was random, except for the restriction that arrays not appear more than three times successively at any one position. Exposure duration was constant for all arrays within a given block. *The threshold durations at which subjects recognized letters at a given position in the visual field were determined by exposing arrays for increased durations across trial blocks* (12, 22, 31, 52, 72, 101 and 149 msec, all durations + 2 msec). *After each block* the subject was given a rest and was reminded of the importance of looking directly at the fixation dot when asked to and of continuing to look at it until after stimulus presentation. At this time, *the experimenter also counted the number of letters correctly reported at each position* (out of 24 possible). *When on some trial block, the subject had correctly reported 12 or more letters* from arrays at a given position and when on this or earlier trial blocks, the subject had done the same with respect to arrays at other positions, *the session ended.*

Using this method of ascending limits, threshold exposure durations were obtained for letter arrays presented at each visual field position. Thresholds were defined as the durations yielding 50% letter recognition accuracy as calculated by linear interpolation. In the few instances in which accuracy either fell slightly above 50% on the intial trial block (ceiling effect) or did not reach the 50% level by the final block (floor effect), linear interpolation was not possible. In these cases a threshold corresponding to that for 50% accuracy at the first or last exposure duration (12 msec or 149 msec), respectively, was assigned. The number of instances in which either circumstance occurred is reported below. Subjects for whom the threshold for letters in the left or right half fields was lower than the threshold for center letters were replaced, since such an outcome was taken as evidence for improper fixation (n = 5 children and 2 adults).

Due to heterogeneity of threshold error variances, analysis of variance was conducted on reciprocal thresholds — a transformation among those suggested by Kirk and in this case conveniently interpretable as letter "perceptibility", with higher scores reflecting lower thresholds and thus greater sensory/perceptual skill. Means and standard deviations for this measure as a function of stimulus condition are presented in Table 1.

The results of the analysis of variance showed significant age differences in letter perceptibility for arrays in each of the three positions: For left half field letters $F(1, 176) = 26.91, p < .01$; *for center letter*, $F(1, 176) = 10.42, p < .01$; *and for right half field letters*, $F(1, 176) = 35.61, p < .01$. That a difference was apparent for central as well as peripheral stimuli argues for a contribution to these results by age trends in factors not specific to the locus of presentation in the visual field (e.g., memory, attention, motivation). *However, the analysis also revealed the predicted Age X Position interaction*, $F(2, 352) = 3.97, p < .05$; *for the left half field versus center comparison*, $F(1, 176) = 7.28, p < .01$. Mean perceptibility scores for left half field, center, and right half field letters were 24.94, 56.64, and 28.82, respectively, for children, and 36.54, 62.04, and 38.81 for adults. *As might be expected, letters presented in the left and right half fields were less discriminable than letters presented at the fixation point for both age groups* ($p < .01$ in all cases).

It is appropriate to point out that the greater age differences in letter perceptibility for peripheral as compared to center arrays (i.e. differences of 9.60 and 9.99 for left and right half field letters, respectively, in contrast to a difference of 5.40 for center letters) cannot reasonably be interpreted in terms of selection procedures. The fact that more children than adults (five vs. two) were excluded for a lower threshold n a peripheral position than at the fixation point suggests that children may have been more prone than adults to look to the side rather than fixate as instructed. The Position X Age effect, therefore, may well have been diminished rather than spuriously inflated by the possible failure of some subjects to comply with fixation instructions. The Age X Position interaction would also likely have been more prominent had it been possible to avoid floor and ceiling effects. The frequency of ceiling effects for center arrays was about the same for children and adults (five and eight, respectively), suggesting that the mean perceptibility for an age group for center letters was underestimated to a similar extent for both age groups. In contrast, the frequency of floor effects for peripheral arrays was much greater for children than adults (13 vs. 1), hence peripheral visual abilities would seem to have been overestimated to a greater extent for the children than for the adults.

Analyses of letter perceptibility failed to indicate a dependency of age differences on the position, type, or orientation of the letter array. As these results may have been obscured by the floor and ceiling effects noted above, an alternative measure that proved much less susceptible to these effects was analyzed. This measure, report accuracy, was derived by ascertaining the first trial block in which each subject correctly reported at least 50% of the letters from peripheral arrays (collapsing over left and right half field presentations) and then tallying the number of letters appearing in each half field that were correctly reported (out of 24 possible). Accuracy scores were then entered into an analysis of variance involving the same factors included in the first major analysis, the only exception being that position was reduced to a two-level factor (left vs. right half field). Results buttressed those of analysis of letter perceptibility in their failure to reveal any dependencies of age differences in peripheral letter perception on the type, orientation, or half field of the array.

With respect to the operation of these factors independent of age, both analyses revealed letters within words to be more discriminable than those within nonwords ($p < .05$). This effect, which did not vary according to the position of the array, agrees with the general finding that letters within arrays of higher order approximation to English are more discriminable than those within arrays of lower order approximation. *Results were consistent with the expectation that letters would be better perceived in the right half field than in the left half field. The only stimulus condition that failed to yield right half field superiority for either measure was vertical words. Analyses of neither letter perceptibility nor report accuracy revealed main effects for sex nor any interpretable interactions involving this factor,* this being the reason for collapsing across sex in Table 1.

Several previous developmental investigations suggest an increase with age in the sensitivity of the peripheral visual system during childhood. Though the existence of such a trend has not gone unchallenged, relevant studies include ones by Lakowski and Aspinall, who found that adolescents were unable to detect peripheral lights at lower luminescence levels than were young children; by Taylor, who found preschool children less accurate than adolescents in detecting peripheral lights of constant luminance; and by L.K. Miller, who observed that college students were able to react more quickly to peripheral lights than were eight-year-olds.

So far as the author is aware, however, *the present data are among the first supporting age differences in the near periphery using a task that explicitly required peripheral discrimination as opposed to mere detection. Since the abilities of each of the two age groups to discriminate peripheral letters were assessed relative to their abilities to* *discriminate letters presented at the fixation point,* these results cannot be readily attributed to variations with age in extraneous factors such as motivation, decision bias, attentiveness, or memory. *A developmental difference specific to peripheral visual processing seems indicated.*

These results challenge Cohen and Haith's conclusion that there may be no differences on tasks requiring static perception of familiar stimuli in the near periphery. The stimulus duration employed by these investigators (20 msec) may not have been one optimally sensitive to age differences. Moreover, the method of gradually decreasing stimulus duration from 150 msec to 20 msec during a pretraining phase and of presenting stimuli at increasingly greater distance from the fovea (2, 4, 6) may have attentuated potential age differences by reducing stimulus uncertainty or via a warm-up effect.

The fact that age differences in right field superiority were not fund argues against any contribution to peripheral vision of developmental trends in cerebral dominance or hypothesised left-to-right postexposural scanning, at least across the age span studied here and with the present stimuli. This is despite the fact that right half field superiority was obtained for three of the four stimulus conditions. Perhaps if arrays of more than four letters had been used or if children younger than eight- and nine-year-olds has been included as subjects, these mechanisms may have operated to enhance differences in peripheral vision in the right half field. Another possibility is that present age differences may have been more pronounced if peripheral stimuli had been presented at greater differences from the fixation point.

Age differences in the ability to make use of redundancies within peripheral words, as opposed to nonwords, also failed to appear. Words were more discriminable than nonwords, and the children were apparently as able as the adults to make use of additional cues available in identifying letters within words. Admittedly, had the orthographic complexity of the words been greater and an age advantage in identifying words over nonwords at the fixation point resulted, the hypothesis that children have a greater difficulty interpreting partial information in the periphery would have been put to a more rigorous test. This hypothesis cannot be dismissed on the basis of the present findings. Failure to show a dependency of present age differences on the type of letter array also fails to rule out developmental trends in the ability to make use of partial letter (as opposed to word) cues in identifying peripheral letters. With respect to both of these possibilities, however, it is important to recognise that merely proposing age differences in the ability to interpret partial information is not sufficient. Because there were greater age differences in the periphery than at the fixation point in the present study, explanations of this sort must specify that interpretation of partial cues seen in the periphery is somehow different from interpretation of such cues at the fixation point.

Several alternative explanations can be offered for the age differences observed here. *One possibility is that the visual systems of adults are better able than those of children to resolve peripheral stimuli. Improved resolution with age could be ascribed to (a) refinements in parts of the receptor apparatus, such as enlargement of the pupil, increasing accommodative power of the lens, or growth of the peripheral retina;* (b) increasing myelination of nerve fibers within extrafoveal visual pathways; (c) histological changes in portions of the cortex that subserve peripheral vision. Regarding a, Lakowski and Aspinall found no relation between pupil diameter and peripheral light sensitivity; and it seems doubtful that pupil diameter would have played much of a role in the present study, since stimuli appeared in the near periphery. Data from Atkinson and Braddick also argue against a. *These investigators suggested that neural processes rather than growth of the peripheral sensory system may account in large part for postinfancy development in acuity.* Evidence summarised by Eichorn also agrees more with alternatives b or c rather than with alternative a.

Another explanation for age differences in the ability to discriminate peripheral letters is that peripheral vision is suppressed in favour of foveal vision to a greater extent in children than in adults. According to this alternative, information arising from the entire visual field is too great for the young child to process all at once. As a result the child's effective vision is restricted to stimuli in the immediate vicinity of the fixation point. The notion that peripheral vision can be so suppressed, termed tunnel vision, is consistent with observations of Liebowitz and Appelle and Mackworth. These investigators found that the ability of adults to perceive stimuli in the periphery declined with increases in the amount of processing required at the fixation point. In the present study, of course, peripheral stimuli were not exposed simultaneously with stimuli at the fixation point. As an explanation for the present findings, tunnel vision would have to refer to a kind of attentional narrowing not directly related o ongoing processing at the center of fixation. Lakowski and Aspinall came closest to this notion when they speculated that "in young children there is cortical suppression of peripheral information, perhaps to allow perceptual development to take place in the foveal area."

Whatever its source, the present finding has several implications. *First, it suggests that explorations of age-related changes in the apparatus subserving peripheral vision may be worthwhile.* Further studies in this area might include developmental investigations of the peripheral retina, of the cortical projection areas, and of the mechanisms responsible for possible cortical suppression of peripheral vision. Second, it supports efforts to chart specific developmental changes in peripheral visual capacity so that performance limitations of subjects of various ages on tasks involving peripheral vision can be better understood. *Agewise limitations in peripheral discrimination capacity may, for example, imply a basic sensory constraint on the ability with which children can make use of contextual cues from the periphery during reading* (e.g. phrase endings, key words). Finally, the age differences in peripheral visual capacity suggested by this finding may be subtle, and their manifestation may therefore require sensitive, psychophysically based procedures.

APPENDIX B

REVISED MATERIAL

Text1

1 *The problems being investigated in this study* are the factors influencing *courteous* behaviour.[3]
61 *It also seems that* the factors *discussed here* are *indeed* those which influence *helping* behaviour.

2 *Researchers investigating prosocial behaviour have found* that it is possible *for a person* to influence the likelihood of receiving help.
60 *It seems, in conclusion,* that it is possible to influence the likelihood of receiving help.

3 Prosocial or altruistic behaviour involves activities where one person assists another but receives no obvious reward for his help.
58 Altruistic or prosocial behaviour involves activities where one person assists another but receives no obvious reward for his help.

4 These kind of behaviours are generally regarded as thoughtful or courteous.
59 These behaviours are generally regarded as thoughtful or courteous.

5 Social contact theory *is* related to helping behaviour.
50 Social contact theory *does appear to be* related to helping behaviour.

6 This suggests that if A had previously had brief verbal contact with B, then B would be more likely to help A than if this initial contact had not occurred.
51 This suggests that if A had previously had brief verbal contact with B, then B would be more likely to help A than if this initial contact had not occurred.

7 Norm-of-reciprocity theory states that people should help those who have helped them.
52 *The* norm-of-reciprocity theory states that people should help those who have helped them.

8 This *theory* suggests that if A has helped B, then B would be more likely to reciprocate by helping A.
53 This suggests that if A has helped B, then B would be more likely to reciprocate by helping A.

9 Equity theory states that interacting individuals try to balance their rewards and costs.
54 Equity theory states that interacting individuals try to balance their costs and

3. The second sentence of each pair is that which appears in the unstructured version. The numbers refer to the order of presentation. The words in italics indicate the words which were changed in order to create the unstructured version.

rewards.

10 This theory suggests that if A has helped B, then B has been rewarded at no cost to himself.
55 This theory suggests that if A has helped B, then B has been rewarded at no cost to himself.

11 Therefore, B would be motivated to re-establish *a balance* by helping A at the first opportunity.
56 Therefore, B would be motivated to re-establish *equity* by helping A at the first opportunity.

12 It has also been suggested that females are less prosocial than males.
57 It has also been suggested that, *on the basis of previous research*, females are less prosocial than males.

13 *At first it may seem* that the norm-of-reciprocity theory and the equity theory would make similar predictions.
38 *It seems at first sight* that the norm-of-reciprocity theory and the equity theory would make similar predictions.

14 Both theories would predict that if A helps B, then B would help A.
39 Both theories would predict that if A helps B, then B will help A.

15 However, it is possible to set up a situation in which the two theories would make alternative predictions.
40 It is possible, however, to set up a situation in which the two theories would make alternative predictions.

16 Equity theory implies that motivation to help will be induced regardless of whether or not the imbalance has been produced intentionally.
41 Equity theory implies that motivation to help will be induced regardless of whether or not the imbalance has been produced intentionally.

17 Therefore, if A helps B unintentionally, equity theory would predict that B would be likely to help A.
42 Therefore, if A helps B unintentionally, equity theory would predict that B would be likely to help A.

18 *However*, norm-of-reciprocity theory would predict that B would be likely to help A only if he thought that A had helped him intentionally.
43 *The* norm-of-reciprocity theory would predict that B would be likely to help A only if he thought that A had helped him intentionally.

19 If the norm-of-reciprocity theory is valid, then the intentional *courtesy* should produce more frequent courteous behaviour than no *courtesy*.
44 If the norm-of-reciprocity theory is valid, then the intentional *helping* should produce more frequent courteous behaviour than no *helping*.

20 Equity theory would also predict this.
45 Equity theory would also predict this.

21 Equity theory would imply that courteous behaviour should occur to the same

extent in both the unintentional and the intentional conditions.
46 Equity theory would imply that courteous behaviour should occur to the same extent in both the intentional and the unintentional conditions.

22 The norm-of-reciprocity theory would predict that more courteous behaviour should occur in the intentional than in the unintentional condition.
47 The norm-of-reciprocity theory would predict that more courteous behaviour should occur in the intentional than in the unintentional condition.

23 Social contact theory would predict that if B has had previous contact with A, then B would be more likely to help than if *no contact had occurred.*
48 Social contact theory would predict that if B has had previous contact with A, then B would be more likely to help A than if *this initial contact had not occurred.*

24 *It would* also be predicted that males would *show more frequent courteous behaviour* than females.
49 *It might* also be predicted that males would *engage in courteous behaviour more frequently* than females.

25 An investigation *of these hypotheses might* reveal which factors influence courteous behaviour.
37 *It was hoped that such* an investigation *would* reveal which factors influence courteous behaviour.

26 The *present* study was designed to test *the aforementioned* predictions.
36 The study was designed to test *various* predictions *concerning courteous behaviour.*

27 A two-by-four experimental design was used.
35 A two-by-four experimental design was used.

28 Eighty-four adult males and eighty-four adult females *who were* using a car-park next to a shopping centre served unwittingly as subjects.
34 Eighty-four adult males and eighty-four adult females using an *indoor* car-park next to a shopping centre served unwittingly as subjects.

29 Sex, *male and female,* was compared with door-holding behaviour under conditions of control, unintentional, contact and intentional.
32 Sex *of subject* was compared with door-holding behaviour under the conditions of control, unintentional, contact and intentional *helping.*

30 Subjects were randomly assigned to each of the four conditions.
33 Subjects were randomly assigned to each of the four conditions.

31 *A male confederate,* dressed similarly to the subject population, *waited for a subject to park his car and walk towards the car-park exit.*
28 *The confederate was male, and was* dressed similarly to the subject population.

32 A trial was begun when the confederate saw a subject *walking alone* towards the exit door.
29 A trial was begun when the confederate saw a subject *walk* towards the exit door *of an indoor car-park.*

33 *In order to leave* the car-park, *subjects had to walk towards* the exit door, open the

door, proceed through a corridor, and finally open a second door at the end of the corridor.
30 *Leaving* the car-park *required walking to* the exit door, opening the door and proceeding through a corridor, and finally opening a second door at the end of the corridor.

34 It was noted whether or not the subject held the second door open for the confederate.
31 It was noted whether or not the subject held the second door open for the confederate.

35 In the control condition, the confederate preceded the subject to the exit door, opened the door and shut it behind him.
18 In the control condition, the confederate preceded a subject to an exit door *which led to a corridor*, opened the door and shut it behind him,

36 *This* required the subject to open the door for himself.
18 *which* required the subject to open the door for himself.

37 In the corridor, the confederate bent down to adjust his socks, which allowed the subject to arrive at the second door before the confederate.
19 In the corridor, the confederate bent down to adjust his socks, which allowed the subject to arrive at the second door before the confederate.

38 In the unintentional condition, the confederate preceded the subject to the exit door and, whilst holding the door open, he bent down to tie his shoelaces.
20 In the unintentional *helping* condition, the confederate preceded the subject to the exit door and, whilst holding the door open, he bent down to tie his shoelace.

39 He took no notice of the subject as the subject walked through the open door.
21 He took no notice of the subject as the subject walked through the open door.

40 Then the confederate followed the subject to the second door.
22 Then the confederate followed the subject to the second door.

41 In the intentional condition, the confederate preceded the subject to the exit door.
23 In the intentional *helping* condition, the confederate preceded the subject to the exit door.

42 He held the door open, allowed the subject to walk through, and then followed the subject to the second door.
24 He held the door open, allowed the subject to walk through, and then followed the subject to the second door.

43 In the contact condition, the confederate preceded the subject to the exit door and held the door open, allowing the subject to pass through.
25 In the contact condition, the confederate preceded the subject to the exit door and held the door open, allowing the subject to pass through.

44 As soon as the subject had passed through the door, the confederate asked the subject for directions to the local theatre.
26 As soon as the subject had passed through the door, the confederate asked the

subject directions to the local theatre.

45 When the subject had responded the confederate thanked him and followed behind the subject to the second door.
27 When the subject had responded the confederate thanked him and followed behind the subject to the second door.

46 *The results of an analysis of variance* showed that the mean courteous behaviour for males was significantly greater than for females.
13 *Their results* showed that the mean courteous behaviour for males was significantly greater than that for females.

47 *The different door holding conditions* yielded a significant difference in the rate of courteous behaviour.
14 *Different conditions of helping* yielded a significant difference in the rate of courteous behaviour.

48 The interaction of the main effects was not significant.
15 The interaction of the main effects *for sex versus helping behavious* was not significant.

49 When compared with each other, the differences between means for control, unintentional and contact conditions were not significant.
16 When compared with each other, the differences between means for control, unintentional and contact conditions were not significant.

50 The mean for the intentional condition was significantly greater than the means for each of the other conditions.
17 *However*, the mean for intentional conditions *of helping* was significantly greater than the means for each of the other conditions.

51 *The present study* found that it is possible for an individual to influence the likelihood that another will behave in a courteous manner towards him.
12 *Goldman, Florez & Fuller (1981)* found that it is possible for an individual to influence the likelihood that another will behave in a courteous manner towards him.

52 *The* intentional condition induced significantly more courteous behaviour than did the control or unintentional conditions.
3 Intentional *helping* conditions induce significantly more courteous behaviour than control or unintentional conditions.

53 *The* unintentional condition did not produce significantly more courteous behaviour than the control condition.
4 Unintentional conditions do not produce significantly more courteous behaviour than control conditions.

54 However, *it appears that* intentional help *must* occur under specific conditions.
5 However, intentional help *apparently has to* occur under specific conditions.

55 *In the contact condition, the subject received intentional help.*
6 *In conditions with verbal contact, for example, where a subject receives intentional help,*

56 *This was* followed by additional contact with the confederate asking for directions.
6 followed by additional contact with the confederate asking for directions,

57 Courteous behaviour did not increase above that in the control condition.
6 courteous behaviour does not increase above that in control condition.

58 The norm-of-reciprocity theory can account for this absence of increased *door-holding* behaviour.
7 The norm-of-reciprocity theory can account for this absence of increased *courteous* behaviour.

59 The confederate acted courteously, holding the door open for the subject.
8 The confederate acts courteously, *for example by* holding a door open for the subject,

60 The subject then reciprocated by answering the confederate's questions.
8 *and* the subject then reciprocates by answering the confederate's questions.

61 Any obligation that the subject felt towards the confederate was removed by his responding to the confederate's request.
9 Any obligation that the subject feels towards the confederate is removed by his responding to the confederate's request.

62 Another reason why the contact *treatment* did not produce increased courteous behaviour might have been that the subject was not certain that the confederate held the door open simply to be polite.
10 Another reason why the contact *condition* does not produce increased courteous behaviour may be that the subject is not certain that the confederate held the door open simply to be polite.

63 The subject might *have understood* that the confederate was holding the door open to reciprocate for the help for which the confederate was asking.
11 The subject might *feel* that the confederate was holding the door open to reciprocate for the help for which the confederate was asking.

64 *It appears, therefore, that the results* support the norm-of-reciprocity theory and do not support the equity theory.
1 *Recent research into factors affecting courteous behaviour seems to* support the norm-of-reciprocity theory and not to support the equity theory.

65 If courteous behaviour is to be reciprocated by an individual, *then* that individual needs to know that he was helped intentionally.
2 *In other words,* if courteous behaviour is to be reciprocated by an individual, that individual needs to know that he was helped intentionally.

Text2

1 *It is possible* that younger children may be less able than older children or adults to discriminate stimuli in the visual periphery.
52 *All of these studies lend support for the possibility* that younger children may be less able than older children or adults to discriminate stimuli in the visual periphery.

2 *Several developmental studies have shown* younger subjects to be less capable than older subjects of both visual search and peripheral recognition of brief exposures.
51 *There have been numerous such studies showing* younger subjects to be less capable than older subjects of both visual search and peripheral recognition of brief exposures.

3 It is unclear, *however*, whether these findings reflect actual age differences in peripheral visual capacity.
50 It is *therefore* unclear whether their findings reflect actual age differences in peripheral visual capacity.

4 The better performance of the older subjects on visual search tasks may reflect more efficient use of what is seen in peripheral vision, rather than better resolution within the visual system.
48 The better performance of older subjects on visual search tasks may reflect more efficient use of what is seen in peripheral vision, rather than better resolution within the visual system, *as some have suggested*.

5 Because most tachistoscopic studies have failed to compare visual capacity at various distances into the periphery with that at the fixation point, these results may stem from other developmental trends in visual processing not specific to peripheral vision.
49 Because most tachistoscopic studies have failed to compare visual capacity at various distances into the periphery with that at the fixation point, their results may stem from other developmental trends in visual processing not specific to peripheral vision.

6 If there are developmental changes in peripheral discrimination capacities beyond those due to changes in overall processing ability, age differences should be greater for stimuli exposed in either visual half-field than for stimuli appearing at the centre of fixation.
44 *Taylor reasoned that* if there are developmental changes in peripheral discrimination capacities beyond those due to changes in overall processing ability, age differences should be greater for stimuli exposed in either visual half-field than for stimuli appearing at the centre of fixation.

7 If older subjects make better use than younger subjects of the partial cues available only in words, these subjects should have special advantage in discriminating them in peripheral vision.
45 If older subjects make better use than younger subjects of the partial cues available only in words, *then* these subjects should have a special advantage in discriminating them in peripheral vision.

8 A greater age difference in peripheral letter discriminability for right half-field

relative to left half-field arrays would support the relevance of cerebral dominance.
46 A greater age difference in peripheral letter discriminability for right half-field relative to left half-field arrays would support the relevance of cerebral dominance.

9 If any enhancement of age differences in the right half-field relative to the left half-field is greater for horizontal than for vertical arrays, a left-to-right postexposural scanning mechanism could be hypothesised.
47 If any enhancement of age differences in the right half-field relative to the left half-field was greater for horizontal than for vertical arrays, a left-to-right postexposural scanning mechanism could have been hypothesised.

10 *This study explores the issue of development in letter perception* using a method that will optimise any potential differences.
43 *None of the previous studies had explored this issue* using a method that would optimise any potential differences.

11 The study compared age differences in threshold discriminability for letter arrays presented at the fixation point versus to the right and left of fixation.
41 The study *was designed to* compare age differences in threshold discriminability for letter arrays presented at the fixation point versus to the right and left of fixation.

12 *A further* purpose was to explore other factors that might contribute to greater ability of older as compared with younger subjects to discriminate peripheral letter arrays.
42 *Another* purpose was to explore other factors that might contribute to greater ability of older as compared with younger subjects to discriminate peripheral letter arrays.

13 The experimental design consisted of the within-subject factor, position of array, and of the four between-subject factors, type of array, orientation of array, age and sex.
40 The experimental design *therefore,* consisted of the within-subject factor, position of array, and of the four between-subject factors, type of array, orientation of array, age and sex.

14 Subjects were ninety-six nine year old children and ninety-six university students, all of whom were right handed.
39 Subjects were ninety-six nine year old children and ninety-six university students, all of whom were right handed.

15 Twenty-four children and twenty-four adults were assigned to each of the four stimulus conditions.
34 Twenty-four children and twenty-four adults were assigned to each of the four stimulus conditions.

16 Stimuli consisted of eighteen words and eighteen nonwords, all containing four different letters.
35 Stimuli consisted of eighteen words and eighteen nonwords, all containing four different letters.

17 All arrays consisted of capital letters typed in black on a white field and presented via a two-field tachistoscope.

36 All arrays consisted of capital letters typed in black on a white field and presented via a two-field tachistoscope.

18 The threshold durations at which subjects recognised letters at a given position in the visual field were determined by exposing arrays for increased durations across trial blocks.
37 The threshold durations at which subjects recognised letters at a given position in the visual field were determined by exposing arrays for increased durations across trial blocks.

19 Thresholds were defined as the durations yielding fifty percent letter recognition accuracy.
38 Thresholds were defined as the durations yielding fifty percent letter recognition accuracy.

20 In order to determine whether age differences *in peripheral vision* might be *introduced by* a greater ability of older subjects to interpret cues received in peripheral vision, *letter arrays* forming both words and nonwords *were used.*
30 *Taylor used letters* forming both words and nonwords in order to determine whether age differences in peripheral vision might be *due to* a greater ability of older subjects to interpret cues received in peripheral vision.

21 In order to investigate any further age advantage given to letters in the right half-field by age trends in cerebral dominance and left-to-right postexposural scanning, letter arrays were presented to the left and right half-fields in vertical as well as horizontal orientation.
31 Letter arrays were presented to the left and right half-fields in vertical as well as horizontal orientation, in order to investigate any further age advantage given to letters in the right half-field by age trends in cerebral dominance and left-to-right postexposural scanning,

22 The conditions were therefore horizontal words, vertical words, horizontal nonwords and vertical nonwords.
32 The conditions *that Taylor used* were therefore horizontal words, vertical words, horizontal nonwords and vertical nonwords.

23 Threshold exposure durations were obtained for letter arrays presented at each visual field position.
33 *Using the method of ascending limits,* threshold exposure durations were obtained for letter arrays presented at each visual field position.

24 Between stimulus presentations a fixation dot was exposed in the centre of the viewing field.
23 Between stimulus presentations a fixation dot was exposed in the centre of the viewing field.

25 On each trial the subject was told to look directly into the centre dot when the experimenter said "Look".
24 On each trial the subject was told to look directly into the centre dot when the experimenter said "Look".

26 After stimulus exposure the subject was asked to report four different letters.
25 After stimulus exposure the subject was asked to report four different letters.

27 *The* eighteen test arrays were divided into three groups of six, with each group assigned to a different one of the three visual field positions.
26 Eighteen test arrays were divided into three groups of six, with each group assigned to a different one of three visual field positions.

28 Test trials were administered in blocks of eighteen, with each of the arrays presented only once per block.
27 Test trials were administered in blocks of eighteen, with each of the arrays presented only once per block.

29 After each block the experimenter counted the number of letters correctly reported at each position.
28 After each block the experimenter counted the number of letters correctly reported at each position.

30 When the subject had correctly reported *nine or more* letters, the session ended.
29 When the subject had correctly reported *fifty percent of the* letters, the session ended.

31 The results of the analysis of variance showed significant age differences in letter perceptibility for arrays in each of the three positions, left half-field letters, right half-field letters and centre letters.
14 These results showed significant age differences in letter perceptibility for arrays in each of the three positions, left half-field letters, right half-field letters and centre letters.

32 *The* analysis also revealed *the predicted* age versus position interaction for left half-field versus centre comparisons and for right half-field versus centre comparisons.
15 However, Taylor's analysis also revealed an age versus position interaction for left half-field versus centre comparisons and for right half-field versus centre comparisons.

33 Letters presented in the left and right half-fields were less discriminable than letters presented at the fixation point for both *age groups.*
16 Letters presented in the left and right half-fields were less discriminable than letters presented at the fixation point for both *adults and young children.*

34 Analyses of letter perceptibility failed to indicate a dependency of age differences on the position, type or orientation of the letter array.
17 Analyses of letter perceptibility failed to indicate a dependency of age differences on the position, type or orientation of the letter array.

35 When these factors were examined independently of age, analyses revealed letters within words to be more discriminable than those within nonwords.
18 When these factors were examined independently of age, analyses revealed letters within words to be more discriminable than those within nonwords.

36 Results *were consistent with the expectation that* letters would be better perceived in the right half-field than in the left half-field.
19 Results *showed that, as expected,* letters would be better perceived in the right half-field than in the left half-field.

37 The only stimulus condition that failed to yield right half-field superiority for either measure was *vertical words*.
20 The only stimulus condition that failed to yield right half-field superiority for either measure was *words presented vertically*.

38 *Analyses failed to reveal* main effects for sex nor any interpretable interactions involving this factor.
21 *There were no* main effects for sex nor any interpretable interactions involving this factor.

39 *The present data* appear to be among the first supporting age differences in the near periphery using a task that explicitly requires peripheral discrimination as opposed to mere detection.
13 *Results of a study by Taylor (1982)* appear to be among the first supporting age differences in the near periphery using a task that explicitly required peripheral discrimination as opposed to mere detection.

40 Since the abilities of *each of the two age groups* to discriminate peripheral letters were assessed relative to their abilities to discriminate letters presented at the fixation point, a developmental difference specific to peripheral visual processing seems indicated.
3 Since, *in these studies*, the abilities of *young children and adults* to discriminate peripheral letters were assessed relative to their abilities to discriminate letters presented at the fixation point, a developmental difference specific to peripheral visual processing seems indicated.

41 The fact that age differences in right half-field superiority were not found argues against any contribution to peripheral vision of developmental trends in cerebral dominance or left-to-right postexposural scanning.
4 The fact that age differences in right half-field superiority were not found argues against any contribution to peripheral vision of developmental trends in cerebral dominance or left-to-right postexposural scanning.

42 Age differences in the ability to make use of redundancies within peripheral words, as opposed to nonwords, *also failed to appear*.
5 *Neither have these studies found any* age differences in the ability to make use of redundancies within peripheral words, as opposed to nonwords.

43 Words *were* more discriminable than nonwords and children were apparently as able as adults to make use of additional cues in identifying letters within words.
6 Words *tend to be* more discriminable than nonwords and children are apparently as able as adults to make use of additional cues in identifying letters within words.

44 It is possible that the visual systems of adults are better able than those of children to resolve peripheral stimuli.
7 It is possible that the visual systems of adults are better able than those of children to resolve peripheral stimuli.

45 This could be due to refinements in receptor apparatus, such as enlargement of the pupil, increased accommodation of the lens, or growth of the peripheral retina.
8 This could be due to refinements in receptor apparatus, such as enlargement of the pupil, increased accommodation of the lens, or growth of the peripheral retina.

46 *However*, other researchers have argued that growth of neural processes play more of a role than growth of peripheral processes in postinfancy development of acuity.
9 Other researchers have argued that growth of neural processes play more of a role than growth of peripheral processes in postinfancy development of acuity.

47 Another explanation for age differences in the ability to discriminate peripheral letters is that peripheral vision is suppressed in favour of foveal vision to a greater extent in children than in adults.
10 Another explanation for age differences in the ability to discriminate peripheral letters is that peripheral vision is suppressed in favour of foveal vision to a greater extent in children than in adults.

48 According to this alternative, information arising from the entire visual field is too great for the young child to process all at once.
11 According to this alternative, information arising from the entire visual field is too great for the young child to process all at once.

49 As a result, the child's effective vision is restricted to stimuli in the immediate vicinity of the fixation point.
12 As a result, the child's effective vision is restricted to stimuli in the immediate vicinity of the fixation point.

50 *The present findings* suggest that exploration of age related changes in the apparatus subserving peripheral vision may be worthwhile.
1 *Recent research seems to* suggest that exploration of age related changes in the apparatus subserving peripheral vision may be a worthwhile *enterprise*.

51 *This study* also suggests that age limitations in peripheral discrimination capacity may *also* imply a basic sensory constraint on the ability with which young children can make use of contextual cues from the periphery during reading.
2 *Studies* also suggest that age limitations in peripheral discrimination capacity may imply a basic sensory constraint on the ability with which young children can make use of contextual cues from the periphery during reading.

APPENDIX C

ORGANISATION OF THE TEXTS

Text1: Structured

PROBLEM
1 **61** The problems being investigated in this study are the factors influencing courteous behaviour.[4]

SETTING
2 **60** Researchers investigating prosocial behaviour have found that it is possible for a person to influence the likelihood of receiving help.

ASSUMPTIONS
3 **58** Prosocial or altruistic behaviour involves activities where one person assists another but receives no obvious reward for his help.
4 **59** These kind of behaviours are generally regarded as thoughtful or courteous.

HYPOTHESES
5 **50** Social contact theory is related to helping behaviour.
6 **51** This suggests that if A had previously had brief verbal contact with B, then B would be more likely to help A than if this initial contact had not occurred.
7 **52** Norm-of-reciprocity theory states that people should help those who have helped them.
8 **53** This theory suggests that if A has helped B, then B would be more likely to reciprocate by helping A.
9 **54** Equity theory states that interacting individuals try to balance their rewards and costs.
10 **55** This theory suggests that if A has helped B, then B has been rewarded at no cost to himself.
11 **56** Therefore, B would be motivated to re-establish a balance by helping A at the first opportunity.
12 **57** It has also been suggested that females are less prosocial than males.

PREDICTIONS
13 **38** At first it may seem that the norm-of-reciprocity theory and the equity theory would make similar predictions.
14 **39** Both theories would predict that if A helps B, then B would help A.
15 **40** However, it is possible to set up a situation in which the two theories would make alternative predictions.
16 **41** Equity theory implies that motivation to help will be induced regardless of whether or not the imbalance has been produced intentionally.
17 **42** Therefore, if A helps B unintentionally, equity theory would predict that B would be likely to help A.
18 **43** However, norm-of-recprocity theory would predict that B would be likely to

4. The first number indicates order of presentation. The second number (bold) indicates the corresponding number of the sentence in the unstructured version.

help A only if he thought that A had helped him intentionally.
19 44 If the norm-of-reciprocity theory is valid, then the intentional courtesy should produce more frequent courteous behaviour than no courtesy.
20 45 Equity theory would also predict this.
21 46 Equity theory would imply that courteous behaviour should occur to the same extent in both the unintentional and the intentional conditions.
22 47 The norm-of-reciprocity theory would predict that more courteous behaviour should occur in the intentional than in the unintentional condition.
23 48 Social contact theory would predict that if B has had previous contact with A, then B would be more likely to help than if no contact had occurred.
24 49 It would also be predicted that males would show more frequent courteous behaviour than females.

SOLUTION
25 37 An investigation of these hypotheses might reveal which factors influence courteous behaviour.

EXPERIMENT
26 36 The present study was designed to test the aforementioned predictions.

DESIGN
27 35 A two-by-four experimental design was used.

SUBJECTS
28 34 Eighty-four adult males and eighty-four adult females who were using a car-park next to a shopping centre served unwittingly as subjects.

METHOD
29 32 Sex, male and female, was compared with door-holding behaviour under conditions of control, unintentional, contact and intentional.

EXECUTION
30 33 Subjects were randomly assigned to each of the four conditions.
31 28 A male confederate, dressed similarly to the subject population, waited for a subject to park his car and walk towards the car-park exit.
32 29 A trial was begun when the confederate saw a subject walking alone towards the exit door.
33 30 In order to leave the car-park, subjects had to walk towards the exit door, open the door, proceed through a corridor, and finally open a second door at the end of the corridor.
34 31 It was noted whether or not the subject held the second door open for the confederate.

ACTS
35 18 In the control condition, the confederate preceded the subject to the exit door, opened the door and shut it behind him.
36 18 This required the subject to open the door for himself.
37 19 In the corridor, the confederate bent down to adjust his socks, which allowed the subject to arrive at the second door before the confederate.
38 20 In the unintentional condition, the confederate preceded the subject to the exit door and, whilst holding the door open, he bent down to tie his shoelaces.
39 21 He took no notice of the subject as the subject walked through the open door.

40 22 Then the confederate followed the subject to the second door.
41 23 In the intentional condition, the confederate preceded the subject to the exit door.
42 24 He held the door open, allowed the subject to walk through, and then followed the subject to the second door.
43 25 In the contact condition, the confederate preceded the subject to the exit door and held the door open, allowing the subject to pass through.
44 26 As soon as the subject had passed through the door, the confederate asked the subject for directions to the local theatre.
45 27 When the subject had responded the confederate thanked him and followed behind the subject to the second door.

RESULTS
46 13 The results of an analysis of variance showed that the mean courteous behaviour for males was significantly greater than for females.
47 14 The different door holding conditions yielded a significant difference in the rate of courteous behaviour.
48 15 The interaction of the main effects was not significant.
49 16 When compared with each other, the differences between means for control, unintentional and contact conditions were not significant.
50 17 The mean for the intentional condition was significantly greater than the means for each of the other conditions.

EVALUATION
51 12 The present study found that it is possible for an individual to influence the likelihood that another will behave in a courteous manner towards him.

DISCUSSION
52 3 The intentional condition induced significantly more courteous behaviour than did the control or unintentional conditions.
53 4 The unintentional condition did not produce significantly more courteous behaviour than the control condition.
54 5 However, it appears that intentional help must occur under specifBc conditions.
55 6 In the contact condition, the subject received intentional help.
56 6 This was followed by additional contact with the confederate asking for directions.
57 6 Courteous behaviour did not increase above that in the control condition.
58 7 The norm-of-reciprocity theory can account for this absence of increased door-holding behaviour.
59 8 The confederate acted courteously, holding the door open for the subject.
60 8 The subject then reciprocated by answering the confederate's questions.
61 9 Any obligation that the subject felt towards the confederate was removed by his responding to the confederate's request.
62 10 Another reason why the contact treatment did not produce increased courteous behaviour might have been that the subject was not certain that the confederate held the door open simply to be polite.
63 11 The subject might have understood that the confederate was holding the door open to reciprocate for the help for which the confederate was asking.

CONCLUSIONS
64 1 It appears, therefore, that the results support the norm-of-reciprocity theory and do not support the equity theory.

65 **2** If courteous behaviour is to be reciprocated by an individual, then that individual needs to know that he was helped intentionally.

Text1: Unstructured

1 **64** Recent research into factors affecting courteous behaviour seems to support the norm-of-reciprocity theory and not to support the equity theory.[5]

2 **65** In other words, if courteous behaviour is to be reciprocated by an individual, that individual needs to know that he was helped intentionally.

3 **52** Intentional helping conditions induce significantly more courteous behaviour than control or unintentional conditions.

4 **53** Unintentional conditions do not produce significantly more courteous behaviour than control conditions.

5 **54** However, intentional help apparently has to occur under specific conditions.

6 **55-57** In conditions with verbal contact, for example, where a subject receives intentional help, followed by additional contact with the confederate asking for directions, courteous behaviour does not increase above that in control conditions.

7 **58** The norm-of-reciprocity theory can account for this absence of increased courteous behaviour.

8 **59-60** The confederate acts courteously, for example by holding a door open for the subject, and the subject then reciprocates by answering the confederate's questions.

9 **61** Any obligation that the subject feels towards the confederate is removed by his responding to the confederate's request.

10 **62** Another reason why the contact condition does not produce increased courteous behaviour may be that the subject is not certain that the confederate held the door open simply to be polite.

11 **63** The subject might feel that the confederate was holding the door open to reciprocate for the help for which the confederate was asking.

12 **51** Goldman, Florez & Fuller (1981) found that it is possible for an individual to influence the likelihood that another will behave in a courteous manner towards him.

13 **46** Their results showed that the mean courteous behaviour for males was significantly greater than that for females.

14 **47** Different conditions of helping yielded a significant difference in the rate of

---

5. The numbers in bold refer to the sentences from the original (structured) version.

courteous behaviour.

15 48 The interaction of the main effects for sex versus helping behaviour was not significant.

16 49 When compared with each other, the differences between means for control, unintentional and contact conditions were not significant.

17 50 However, the mean for intentional conditions of helping was significantly greater than the means for each of the other conditions.

18 35-36 In the control condition, a confederate preceded a subject to an exit door which led to a corridor, opened the door, and shut it behind him, which required the subject to open the door himself.

19 37 In the corridor, the confederate bent down to adjust his socks, which allowed the subject to arrive at the second door before the confederate.

20 38 In the unintentional helping condition, the confederate preceded the subject to the exit door and, whilst holding the door open, he bent down to tie his shoelace.

21 39 He took no notice of the subject as the subject walked through the open door.

22 40 Then the confederate followed the subject to the second door.

23 41 In the intentional helping condition, the confederate preceded the subject to the exit door.

24 42 He held the door open, allowed the subject to walk through, and then followed the subject to the second door.

25 43 In the contact condition, the confederate preceded the subject to the exit door and held the door open, allowing the subject to pass through.

26 44 As soon as the subject had passed through the door, the confederate asked the subject directions to the local theatre.

27 45 When the subject had responded the confederate thanked him and followed behind the subject to the second door.

28 31 The confederate was male, and was dressed similarly to the subject population.

29 32 A trial was begun when the confederate saw a subject walk towards the exit door of an indoor car-park.

30 33 Leaving the car-park required walking to the exit door, opening the door and proceeding through a corridor, and finally opening a second door at the end of the corridor.

31 34 It was noted whether or not the subject held the second door open for the confederate.

32 29 Sex of subject was compared with door-holding behaviour under the conditions of control, unintentional, contact and intentional helping.

33 30 Subjects were randomly assigned to each of the four conditions.

34 28 Eighty-four adult males and eighty-four adult females using the indoor car-park next to a shopping centre served unwittingly as subjects.

35 27 A two-by-four experimental design was used.

36 26 The study was designed to test various predictions concerning courteous behaviour.

37 25 It was hoped that such an investigation would reveal which factors influence courteous behaviour.

38 13 It seems at first sight that the norm-of-reciprocity theory and the equity theory would make similar predictions.

39 14 Both theories would predict that if A helps B, then B will help A.

40 15 It is possible, however, to set up a situation in which the two theories would make alternative predictions.

41 16 Equity theory implies that motivation to help will be induced regardless of whether or not an imbalance has been produced intentionally.

42 17 Therefore, if A helps B unintentionally, equity theory would predict that B would be likely to help A.

43 18 The norm-of-recprocity theory would predict that B would be likely to help A only if he thought that A had helped him intentionally.

44 19 If the norm-of-reciprocity theory is valid, then the intentional helping should produce more frequent courteous behaviour than no helping.

45 20 Equity theory would also predict this.

46 21 Equity theory would imply that courteous behaviour should occur to the same extent in both the intentional and the unintentional conditions.

47 22 The norm-of-reciprocity theory would predict that more courteous behaviour should occur in the intentional condition than in the unintentional condition.

48 23 Social contact theory would predict that if B has had previous contact with A, then B would be more likely to help A than if this intial contact had not occurred.

49 24 It might also be predicted that males would engage in courteous behaviour more frequently than females.

50 5 Social contact theory does appear to be related to helping behaviour.

51 6 This suggests that if A had previously had brief verbal contact with B, then B would be more likely to help A than if this initial contact had not occurred.

52 7 The norm-of-reciprocity theory states that people should help those who have helped them.

53 8 This suggests that if A has helped B, then B would be more likely to reciprocate by helping A.

54 9 Equity theory states that interacting individuals try to balance their costs and rewards.

55 10 This theory suggests that, if A has helped B, then B has been rewarded at no cost to himself.

56 11 Therefore, B would be motivated to re-establish equity by helping A at the first opportunity.

57 12 It has also been suggested that, on the basis of previous research, females are less prosocial than males.

58 3 Altruistic or prosocial behaviour involves activities where one person assists another but receives no obvious reward for his help.

59 4 These behaviours are generally regarded as thoughtful or courteous.

60 2 It seems, in conclusion, that it is possible to influence the likelihood of receiving help.

61 1 It also seems that the factors discussed here are indeed those which influence helping behaviour.

Text2: Structured

PROBLEM

1 52 It is possible that younger children may be less able than older children or adults to discriminate stimuli in the visual periphery.

SETTING

2 51 Several developmental studies have shown younger subjects to be less capable than older subjects of both visual search and peripheral recognition of brief exposures.

ASSUMPTIONS

3 50 It is unclear, however, whether these findings reflect actual age differences in peripheral visual capacity.

HYPOTHESES

4 48 The better performance of the older subjects on visual search tasks may reflect more efficient use of what is seen in peripheral vision, rather than better resolution within the visual system.

5 49 Because most tachistoscopic studies have failed to compare visual capacity at various distances into the periphery with that at the fixation point, these results may stem from other developmental trends in visual processing not specific to peripheral vision.

PREDICTIONS

6 44 If there are developmental changes in peripheral discrimination capacities beyond those due to changes in overall processing ability, age differences should be greater for stimuli exposed in either visual half-field than for stimuli appearing at the centre of fixation.

7 45 If older subjects make better use than younger subjects of the partial cues available only in words, these subjects should have special advantage in discriminating them in peripheral vision.

8 46 A greater age difference in peripheral letter discriminability for right half-field relative to left half-field arrays would support the relevance of cerebral dominance.

9 47 If any enhancement of age differences in the right half-field relative to the left half-field is greater for horizontal than for vertical arrays, a left-to-right postexposural scanning mechanism could be hypothesised.

SOLUTION

10 43 This study explores the issue of development in letter perception using a method that will optimise any potential differences.

EXPERIMENT

11 41 The study compared age differences in threshold discriminability for letter arrays presented at the fixation point versus to the right and left of fixation.

12 42 A further purpose was to explore other factors that might contribute to greater ability of older as compared with younger subjects to discriminate peripheral letter arrays.

DESIGN

13 40 The experimental design consisted of the within-subject factor, position of array, and of the four between-subject factors, type of array, orientation of array,

age and sex.

SUBJECTS
14 39 Subjects were ninety-six nine year old children and ninety-six university students, all of whom were right handed.

METHOD
15 34 Twenty-four children and twenty-four adults were assigned to each of the four stimulus conditions.
16 35 Stimuli consisted of eighteen words and eighteen nonwords, all containing four different letters.
17 36 All arrays consisted of capital letters typed in black on a white field and presented via a two-field tachistoscope.
18 37 The threshold durations at which subjects recognised letters at a given position in the visual field were determined by exposing arrays for increased durations across trial blocks.
19 38 Thresholds were defined as the durations yielding fifty percent letter recognition accuracy.

EXECUTION
20 30 In order to determine whether age differences in peripheral vision might be introduced by a greater ability of older subjects to interpret cues received in peripheral vision, letter arrays forming both words and nonwords were used.
21 31 In order to investigate any further age advantage given to letters in the right half-field by age trends in cerebral dominance and left-to-right postexposural scanning, letter arrays were presented to the left and right half-fields in vertical as well as horizontal orientation.
22 32 The conditions were therefore horizontal words, vertical words, horizontal nonwords and vertical nonwords.

ACTS
23 22 Threshold exposure durations were obtained for letter arrays presented at each visual field position.
24 23 Between stimulus presentations a fixation dot was exposed in the centre of the viewing field.
25 24 On each trial the subject was told to look directly into the centre dot when the experimenter said "Look".
26 25 After stimulus exposure the subject was asked to report four different letters.
27 26 The eighteen test arrays were divided into three groups of six, with each group assigned to a different one of the three visual field positions.
28 27 Test trials were administered in blocks of eighteen, with each of the arrays presented only once per block.
29 28 After each block the experimenter counted the number of letters correctly reported at each position.
30 29 When the subject had correctly reported nine or more letters, the session ended.

RESULTS
31 14 The results of the analysis of variance showed significant age differences in letter perceptibility for arrays in each of the three positions, left half-field letters, right half-field letters and centre letters.
32 15 The analysis also revealed the predicted age versus position interaction for left half-field versus centre comparisons and for right half-field versus centre

comparisons.
33 **16** Letters presented in the left and right half-fields were less discriminable than letters presented at the fixation point for both age groups.
34 **17** Analyses of letter perceptibility failed to indicate a dependency of age differences on the position, type or orientation of the letter array.
35 **18** When these factors were examined independently of age, analyses revealed letters within words to be more discriminable than those within nonwords.
36 **19** Results were consistent with the expectation that letters would be better perceived in the right half-field than in the left half-field.
37 **20** The only stimulus condition that failed to yield right half-field superiority for either measure was vertical words.
38 **21** Analyses failed to reveal main effects for sex nor any interpretable interactions involving this factor.

EVALUATION
39 **13** The present data appear to be among the first supporting age differences in the near periphery using a task that explicitly requires peripheral discrimination as opposed to mere detection.

DISCUSSION
40 **3** Since the abilities of each of the two age groups to discriminate peripheral letters were assessed relative to their abilities to discriminate letters presented at the fixation point, a developmental difference specific to peripheral visual processing seems indicated.
41 **4** The fact that age differences in right half-field superiority were not found argues against any contribution to peripheral vision of developmental trends in cerebral dominance or left-to-right postexposural scanning.
42 **5** Age differences in the ability to make use of redundancies within peripheral words, as opposed to nonwords, also failed to appear.
43 **6** Words were more discriminable than nonwords and children were apparently as able as adults to make use of additional cues in identifying letters within words.
44 **7** It is possible that the visual systems of adults are better able than those of children to resolve peripheral stimuli.
45 **8** This could be due to refinements in receptor apparatus, such as enlargement of the pupil, increased accommodation of the lens, or growth of the peripheral retina.
46 **9** However, other researchers have argued that growth of neural processes play more of a role than growth of peripheral processes in postinfancy development of acuity.
47 **10** Another explanation for age differences in the ability to discriminate peripheral letters is that peripheral vision is suppressed in favour of foveal vision to a greater extent in children than in adults.
48 **11** According to this alternative, information arising from the entire visual field is too great for the young child to process all at once.
49 **12** As a result, the child's effective vision is restricted to stimuli in the immediate vicinity of the fixation point.

CONCLUSIONS
50 **1** The present findings suggest that exploration of age related changes in the apparatus subserving peripheral vision may be worthwhile.
51 **2** This study also suggests that age limitations in peripheral discrimination capacity may also imply a basic sensory constraint on the ability with which young children can make use of contextual cues from the periphery during reading.

Text2: Unstructured

1 50 Recent research seems to suggest that exploration of age related changes in the apparatus subserving peripheral vision may be a worthwhile enterprise.

2 51 Studies also suggest that age limitations in peripheral discrimination capacity may imply a basic sensory constraint on the ability with which young children can make use of contextual cues from the periphery during reading.

3 40 Since, in these studies, the abilities of young children and adults to discriminate peripheral letters were assessed relative to their abilities to discriminate letters presented at the fixation point, a developmental difference specific to peripheral visual processing seems indicated.

4 41 The fact that age differences in right half-field superiority were not found argues against any contribution to peripheral vision of developmental trends in cerebral dominance or left-to-right postexposural scanning.

5 42 Neither have these studies found any age differences in the ability to make use of redundancies within peripheral words, as opposed to nonwords.

6 43 Words tend to be more discriminable than nonwords and children are apparently as able as adults to make use of additional cues available in identifying letters within words.

7 44 It is possible that the visual systems of adults are better able than those of children to resolve peripheral stimuli.

8 45 This could be due to refinements in receptor apparatus, such as enlargement of the pupil, increased accommodation of the lens, or growth of the peripheral retina.

9 46 Other researchers have argued that growth of neural processes play more of a role than growth of peripheral processes in postinfancy development of acuity.

10 47 Another explanation for age differences in the ability to discriminate peripheral letters is that peripheral vision is suppressed in favour of foveal vision to a greater extent in children than in adults.

11 48 According to this alternative, information arising from the entire visual field is too great for the young child to process all at once.

12 49 As a result, the child's effective vision is restricted to stimuli in the immediate vicinity of the fixation point.

13 39 Results of a study by Taylor (1982) appear to be among the first reporting age differences in the near periphery using a task that explicitly required peripheral

discrimination as opposed to mere detection.

14 31 These results showed significant age differences in letter perceptibility for arrays in each of three positions, left half-field letters, right half-field letters and centre letters.

15 32 However, Taylor's analysis also revealed an age versus position interaction for left half-field versus centre comparisons and for right half-field versus centre comparisons.

16 33 Letters presented in the left and right half-fields were less discriminable than letters presented at the fixation point for both adults and young children.

17 34 Analyses of letter perceptibility failed to indicate a dependency of age differences on the position, type or orientation of the letter array.

18 35 When these factors were examined independently of age, analyses revealed letters within words to be more discriminable than those within nonwords.

19 36 Results showed that, as expected, letters were better perceived in the right half-field than in the left half-field.

20 37 The only stimulus condition that failed to yield right half-field superiority was words presented vertically.

21 38 There were no main effects for sex nor any interpretable interactions involving this factor.

22 23 Threshold exposure durations were obtained for letter arrays presented at each visual field position.

23 24 Between stimulus presentations a fixation dot was exposed in the centre of the viewing field.

24 25 On each trial the subject was told to look directly into the centre dot when the experimenter said "Look".

25 26 After stimulus exposure the subject was asked to report four different letters.

26 27 Eighteen test arrays were divided into three groups of six, with each group assigned to a different one of three visual field positions.

27 28 Test trials were administered in blocks of eighteen, with each of the arrays presented only once per block.

28 29 After each block the experimenter counted the number of letters correctly reported at each position.

29 30 When the subject had correctly reported fifty percent of the letters, the session ended.

30 20 Taylor used letters forming both words and nonwords in order to determine

whether age differences might be due to a greater ability of older subjects to interpret cues received in peripheral vision.

31 21 Letter arrays were presented to the left and right half-fields in vertical as well as horizontal orientation, in order to investigate any further age advantage given to letters in the right half-field by age trends in cerebral dominance and left-to-right postexposural scanning.

32 22 The conditions that Taylor used were therefore horizontal words, vertical words, horizontal nonwords and vertical nonwords.

33 23 Using the method of ascending limits, threshold exposure durations were obtained for letter arrays presented at each visual field position.

34 15 Twenty-four children and twenty-four adults were assigned to each of the four stimulus conditions.

35 16 Stimuli consisted of eighteen words and eighteen nonwords, all containing four different letters.

36 17 All arrays consisted of capital letters typed in black on a white field and presented via a two-field tachistoscope.

37 18 The threshold durations at which subjects recognised letters at a given position in the visual field were determined by exposing arrays for increased durations across trial blocks.

38 19 Thresholds were defined as the durations yielding fifty percent letter recognition accuracy.

39 14 Subjects were ninety-six nine year old children and ninety-six university students, all of whom were right handed.

40 13 The experimental design therefore, consisted of the within-subject factor, position of array, and of the four between-subject factors, type of array, orientation of array, age and sex.

41 11 The study was designed to compare age differences in threshold discriminability for letter arrays presented at the fixation point versus to the right and left of fixation.

42 12 Another purpose was to explore other factors that might contribute to greater ability of older as compared with younger subjects to discriminate peripheral letter arrays.

43 10 None of the previous studies had explored this issue using a method that would optimise any potential differences.

44 6 Taylor reasoned that if there are developmental changes in peripheral

discrimination capacities beyond those due to changes in overall processing ability, age differences should be greater for stimuli exposed in either visual half-field than for stimuli appearing at the centre of fixation.

45 7 If older subjects make better use than younger subjects of the partial cues available only in words, then these subjects should have a special advantage in discriminating them in peripheral vision.

46 8 A greater age difference in peripheral letter discriminability for right half-field relative to left half-field arrays would support the relevance of cerebral dominance.

47 9 If any enhancement of age differences in the right half-field relative to the left half-field was greater for horizontal than for vertical arrays, a left-to-right postexposural scanning mechanism could have been hypothesised.

48 4 The better performance of older subjects on visual search tasks may reflect more efficient use of what is seen in peripheral vision, rather than better resolution within the visual system, as some have suggested.

49 5 Because most tachistoscopic studies have failed to compare visual capacity at various distances into the periphery with that at the fixation point, their results may stem from other developmental trends in visual processing not specific to peripheral vision.

50 3 It is therefore unclear whether their findings reflect actual differences in peripheral visual capacity.

51 2 There have been numerous such studies showing younger subjects to be less able than older subjects of both visual search and peripheral recognition of brief exposures.

52 1 All of these studies lend support for the possibility that younger children may be less able than older children or adults to discriminate stimuli in the visual periphery.

APPENDIX D

INSTRUCTIONS FOR SUBJECTS

Condition A

STRUCTURED TEXT

Instructions For Reading

A number of sentences will appear one at a time, on the screen in front of you. You are asked to read each sentence carefully, in order to understand it, and then to press the RETURN key once, when you are ready to move on to the next sentence.

The text concerns a piece of research in psychology. When you are ready to begin reading, press the RETURN key.

Instructions For Recall

Please write down as many of the sentences as you can recall, word for word if possible, and in the order in which you have read them. You may have as much time as you like.

In order to help you in this task, you are asked to bear in mind the following points concerning the reporting of psychological research:

> A research report usually begins with a statement of the *problem* to be investigated. There then follows an outline of the research conducted previously, whose findings are relevant to the current problem; this is the *setting* or *background*. A number of *assumptions* are then outlined, upon which various *theories* or *hypotheses* are based. From these hypotheses the researcher can draw various *predictions*. These predictions lead to a proposed *solution* to the problem outlined previously. This solution consists of an outline of the *experiment* to be conducted. The *design* of the experiment is outlined, with details of *subjects* and *method* of testing used. A description of the *execution* or procedure of the experiment is then given, with details of *acts* performed, and *results* obtained. These results are then *evaluated*, with a detailed *discussion* of the findings, from which various *conclusions* are drawn.

In summary, then, a psychological research report usually takes the following form:

Statement of problem
Setting or background of previous research
Assumptions
Theory or hypotheses
Predictions

Proposed solution to problem
Outline of experiment
Design used
Subjects
Method for testing
Procedure or execution of experiment
Details of acts performed
Results obtained
Outline evaluation of experiment
Detailed discussion
Conclusions

Remember that you are asked to recall the sentences word for word, an in the correct order, as much as possible. You have as much time as you like.

UNSTRUCTURED TEXT

<u>Instructions For Reading</u>

A number of sentences will appear one at a time, on the screen in front of you. You are asked to read each sentence carefully, in order to understand it, and then to press the RETURN key once, when you are ready to move on to the next sentence.

The text concerns a piece of research in psychology. The text is taken from an actual research report, but has been slightly reworded and reordered. When you are ready to begin reading, press the RETURN key.

<u>Instructions For Recall</u>

You are asked to rewrite the text you have read, without altering the actual content as far as possible, so that it takes the form of a research report. You have as much time as you like.

In order to help you to organise the text in your mind, you are asked to bear in mind the following points concerning the reporting of psychological research:

> A research report usually begins with a statement of the *problem* to be investigated. There then follows an outline of the research conducted previously, whose findings are relevant to the current problem; this is the *setting* or *background*. A number of *assumptions* are then outlined, upon which various *theories* or *hypotheses* are based. From these hypotheses the researcher can draw various *predictions*. These predictions lead to a proposed *solution* to the problem outlined previously. This solution consists of an outline of the *experiment* to be conducted. The *design* of the experiment is outlined, with details of *subjects* and *method* of testing used. A description of the *execution* or procedure of the experiment is then given, with details of *acts* performed, and *results* obtained. These results are then *evaluated*, with a detailed *discussion* of the findings, from which various *conclusions* are drawn.

In summary, then, a psychological research report usually takes the following form:

Statement of problem
Setting or background of previous research
Assumptions
Theory or hypotheses
Predictions
Proposed solution to problem
Outline of experiment
Design used
Subjects
Method for testing
Procedure or execution of experiment
Details of acts performed
Results obtained
Outline evaluation of experiment

Detailed discussion
Conclusions

Condition B

STRUCTURED TEXT

Instructions For Reading

A number of sentences will appear one at a time, on the screen in front of you. You are asked to read each sentence carefully, in order to understand it, and then to press the RETURN key once, when you are ready to move on to the next sentence.

The text concerns a piece of research in psychology. In order to help you to organise the text in your mind, you are asked to bear in mind the following points concerning the reporting of psychological research:

> A research report usually begins with a statement of the *problem* to be investigated. There then follows an outline of the research conducted previously, whose findings are relevant to the current problem; this is the *setting* or *background*. A number of *assumptions* are then outlined, upon which various *theories* or *hypotheses* are based. From these hypotheses the researcher can draw various *predictions*. These predictions lead to a proposed *solution* to the problem outlined previously. This solution consists of an outline of the *experiment* to be conducted. The *design* of the experiment is outlined, with details of *subjects* and *method* of testing used. A description of the *execution* or procedure of the experiment is then given, with details of *acts* performed, and *results* obtained. These results are then *evaluated*, with a detailed *discussion* of the findings, from which various *conclusions* are drawn.

In summary, then, a psychological research report usually takes the following form:

Statement of problem
Setting or background of previous research
Assumptions
Theory or hypotheses
Predictions
Proposed solution to problem
Outline of experiment
Design used
Subjects
Method for testing
Procedure or execution of experiment
Details of acts performed
Results obtained
Outline evaluation of experiment
Detailed discussion
Conclusions

When you are ready to begin reading, press the RETURN key.

<u>Instructions For Recall</u>

Please write down as many of the sentences as you can recall, word for word if possible, and in the order in which you have read them. You may have as much time as you like.

UNSTRUCTURED TEXT

Instructions For Reading

A number of sentences will appear one at a time, on the screen in front of you. You are asked to read each sentence carefully, in order to understand it, and then to press the RETURN key once, when you are ready to move on to the next sentence.

The text concerns a piece of research in psychology. The text is taken from an actual research report, but has been slightly reworded and reordered. In order to help you to organise the text in your mind, you are asked to bear in mind the following points concerning the reporting of psychological research:

> A research report usually begins with a statement of the *problem* to be investigated. There then follows an outline of the research conducted previously, whose findings are relevant to the current problem; this is the *setting* or *background*. A number of *assumptions* are then outlined, upon which various *theories* or *hypotheses* are based. From these hypotheses the researcher can draw various *predictions*. These predictions lead to a proposed *solution* to the problem outlined previously. This solution consists of an outline of the *experiment* to be conducted. The *design* of the experiment is outlined, with details of *subjects* and *method* of testing used. A description of the *execution* or procedure of the experiment is then given, with details of *acts* performed, and *results* obtained. These results are then *evaluated*, with a detailed *discussion* of the findings, from which various *conclusions* are drawn.

In summary, then, a psychological research report usually takes the following form:

Statement of problem
Setting or background of previous research
Assumptions
Theory or hypotheses
Predictions
Proposed solution to problem
Outline of experiment
Design used
Subjects
Method for testing
Procedure or execution of experiment
Details of acts performed
Results obtained
Outline evaluation of experiment
Detailed discussion
Conclusions

When you are ready to begin reading, press the RETURN key.

<u>**Instructions For Recall**</u>

You are asked to rewrite the text you have read, without altering the actual content as far as possible, so that it takes the form of a research report, as outlined to you previously. You have as much time as you like.

Condition N

STRUCTURED TEXT

Instructions For Reading

A number of sentences will appear one at a time, on the screen in front of you. You are asked to read each sentence carefully, in order to understand it, and then to press the RETURN key once, when you are ready to move on to the next sentence.

The text concerns a piece of research in psychology. When you are ready to begin reading, press the RETURN key.

Instructions For Recall

Please write down as many of the sentences as you can recall, word for word if possible, and in the order in which you have read them. You may have as much time as you like.

UNSTRUCTURED TEXT

Instructions For Reading

A number of sentences will appear one at a time, on the screen in front of you. You are asked to read each sentence carefully, in order to understand it, and then to press the RETURN key once, when you are ready to move on to the next sentence.

The text concerns a piece of research in psychology. When you are ready to begin reading, press the RETURN key.

Instructions For Recall

You are asked to rewrite the text you have read, without altering the actual content as far as possible, so that it takes the form of a research report. You have as much time as you like.

APPENDIX E

INSTRUCTIONS FOR JUDGES

Instructions for Coding Texts

You are asked to go through four texts, of about 1000 words each, and for each text, to divide it into "idea units". An idea unit encompasses a complete thought; and while idea units often coincide with sentences, clauses, or are set off by punctuation, this is not always the case. More than one idea unit can occur in a single sentence. Sometimes you will not be certain where one idea ends and another begins. I would like you to make judgements despite this uncertainty. Please place a slash mark ("/") after each idea unit. You may go back over the text as many times as you like until you are happy about the divisions you have made.

When you have performed this procedure for each of the texts, please go through each one of the 70 protocols (the subject sheets) and do the same as you did for the original texts.

In order to help you to decide what you are being asked to do, on the next page is a sample paragraph, together with the same paragraph divided into possible "idea units".

Wolf packs are territorial, each pack occupying a home range of 50 square miles or more. A wolf or wolf pack intruding on another's home range is attacked aggressively. Two adjacent ranges may overlap, but since the ranges of overlap are small and the territories very large, the chance of two packs encountering one another is not great. The boundaries of territories are scent marked with urine: scent marking may convey information about both the extent of the territory and the amount of time that has passed since the pack passed by, thus making it easy for packs to avoid one another. It also might provide information about the number of wolves in the pack, since it is likely that the urine of each wolf has a different odor.

Wolf packs are territorial,/
each pack occupying a home range of 50 square miles or more./
A wolf or wolf pack intruding on another's home range is attacked aggressively./
Two adjacent ranges may overlap,/
but since the ranges of overlap are small and the territories very large,/
the chance of two packs encountering one another is not great./
The boundaries of territories are scent marked with urine:/
scent marking may convey information about both the extent of the territory/
and the amount of time that has passed since the pack passed by,/
thus making it easy for packs to avoid one another./
It also might provide information about the number of wolves in the pack,/
since it is likely that the urine of each wolf has a different odor./

Instructions For Scoring Recall Protocols

You are being given a number of recall protocols to score, together with a copy of the original text, with which to compare the protocols. The original text is divided into "idea units", based on the coding which you did previously. Each unit is numbered. Each of the protocols is also divided into "idea units". These units are delimited by a vertical bar or slash mark, and a blank line. They are not numbered.

You are asked to go through each protocol and compare it with the corresponding original text. For each unit in the protocol decide whether or not it is correctly recalled from the original text, and note the corresponding unit number beside it. A unit is counted as being correctly recalled if it captures the same meaning as the original (it doesn't have to be in exactly the same words). If the unit only captures part of the meaning of the original, number it as above, but place a "?" beside the number. If the unit is a generalisation or summary of several of the original units, note the numbers it corresponds to, as above, but place a "G" beside it.

APPENDIX F

DIVISIONS USED IN CODING

Text1: Structured

The problems being investigated in this study #1
are the factors influencing courteous behaviour. #3[6]
Researchers investigating prosocial behaviour #1
have found that it is possible for a person to influence the likelihood of receiving help. #3

Prosocial or altruistic behaviour involves activities #1
where one person assists another but receives no obvious reward for his help. #3

These kind of behaviours are generally regarded as thoughtful #1
or courteous. #3

Social contact theory is related to helping behaviour. #3

This suggests that if A had previously had brief verbal contact with B, #1
then B would be more likely to help A #1
than if this initial contact had not occurred. #3

Norm-of-reciprocity theory states #1
that people should help those who have helped them. #3

This theory suggests that if A has helped B, #1
then B would be more likely to reciprocate by helping A. #3

Equity theory states that interacting individuals try to balance their rewards and costs. #3

This theory suggests that if A has helped B, #1
then B has been rewarded at no cost to himself. #3

Therefore, B would be motivated to re-establish a balance by helping A #1
at the first opportunity. #3

It has also been suggested that females are less prosocial than males. #3

At first it may seem that the norm-of-reciprocity theory and the equity theory would make similar predictions. #3

Both theories would predict that if A helps B, #1
then B would help A. #3

---

6. The numbers refer to the number of judges in agreement concerning the division.

However, it is possible to set up a situation #1
in which the two theories would make alternative predictions. #3

Equity theory implies that motivation to help will be induced #1
regardless of whether or not the imbalance has been produced
intentionally. #3

Therefore, if A helps B unintentionally, #1
equity theory would predict that B would be likely to help A. #3

However, norm-of-recprocity theory would predict that B would be
likely to help A #1
only if he thought that A had helped him intentionally. #3

If the norm-of-reciprocity theory is valid, #1
then the intentional courtesy should produce more frequent courteous
behaviour
than no courtesy. #3

Equity theory would also predict this. #3

Equity theory would imply that courteous behaviour should occur #1
to the same extent in both the unintentional and the intentional
conditions. #3

The norm-of-reciprocity theory would predict that more courteous
behaviour should occur #1
in the intentional than in the unintentional condition. #3

Social contact theory would predict #1
that if B has had previous contact with A, #1
then B would be more likely to help #1
than if no contact had occurred. #3

It would also be predicted #1
that males would show more frequent courteous behaviour than females. #3

An investigation #1
of these hypotheses #1
might reveal which factors influence courteous behaviour. #3

The present study was designed #1
to test the aforementioned predictions. #3

A two-by-four experimental design was used. #3

Eighty-four adult males #1
and eighty-four adult #1
females who were using a car-park #1
next to a shopping centre #1
served unwittingly as subjects. #3

Sex, male and female, was compared with door-holding behaviour #1
under conditions of control, #2
unintentional, contact and intentional. #3

Subjects were randomly assigned to each of the four conditions. #3

A male confederate, #1
dressed similarly to the subject population, #1
waited for a subject to park his car #1
and walk towards the car-park exit. #3

A trial was begun #2
when the confederate saw a subject walking alone #1
towards the exit door. #3

In order to leave the car-park, #2
subjects had to walk towards the exit door, #2
open the door, #2
proceed through a corridor, #2
and finally open a second door #1
at the end of the corridor. #3

It was noted whether or not the subject held the second door open for the confederate. #3

In the control condition, #1
the confederate preceded the subject to the exit door, #3
opened the door #2
and shut it behind him. #3

This required the subject to open the door for himself. #2

In the corridor, #1
the confederate bent down to adjust his socks, #2
which allowed the subject to arrive at the second door before the confederate. #3

In the unintentional condition, #1
the confederate preceded the subject to the exit door #3
and, whilst holding the door open, #2
he bent down to tie his shoelaces. #3

He took no notice of the subject #2
as the subject walked through the open door. #3

Then the confederate followed the subject #1
to the second door. #3

In the intentional condition, #1
the confederate preceded the subject #1
to the exit door. #3

He held the door open, #3

allowed the subject to walk through, #3
and then followed the subject #1
to the second door. #3

In the contact condition, #1
the confederate preceded the subject #1
to the exit door #2
and held the door open, #2
allowing the subject to pass through. #3

As soon as the subject had passed through the door, #2
the confederate asked the subject for directions to the
local theatre. #3

When the subject had responded #2
the confederate thanked him #3
and followed behind the subject #1
to the second door. #2

The results of an analysis of variance showed that #1
the mean courteous behaviour for males was significantly greater than
for females. #3

The different door holding conditions yielded a significant difference #1
in the rate of courteous behaviour. #3

The interaction of the main effects was not significant. #3

When compared with each other, #1
the differences between means for control, #1
unintentional and contact conditions were not significant. #3

The mean for the intentional condition was significantly greater #1
than the means for each of the other conditions. #3

The present study found that it is possible for an individual to
influence the likelihood #1
that another will behave in a courteous manner towards him. #3

The intentional condition induced significantly more courteous behaviour #1
than did the control #1
or unintentional conditions. #3

The unintentional condition did not produce significantly more
courteous behaviour #1
than the control condition. #3

However, it appears that intentional help must occur #1
under specific conditions. #3

In the contact condition, #1
the subject received intentional help. #3

This was followed by additional contact #1
with the confederate asking for directions. #3

Courteous behaviour did not increase above #1
that in the control condition. #3

The norm-of-reciprocity theory can account for this absence of increased door-holding behaviour. #3

The confederate acted courteously, #3
holding the door open for the subject. #3

The subject then reciprocated #1
by answering the confederate's questions. #2

Any obligation that the subject felt #1
towards the confederate was removed #1
by his responding to the confederate's request. #3

Another reason why the contact treatment did not produce increased courteous behaviour #1
might have been that the subject was not certain #1
that the confederate held the door open simply to be polite. #3

The subject might have understood #1
that the confederate was holding the door open to reciprocate for the help #1
for which the confederate was asking. #3

It appears, therefore, that the results support the norm-of-reciprocity theory #2
and do not support the equity theory. #3

If courteous behaviour is to be reciprocated by an individual, #1
then that individual needs to know that he was helped intentionally. #3

Text1: Unstructured

Recent research into factors affecting courteous behaviour seems to support
the norm-of-reciprocity theory #2
and not to support the equity theory. #3

In other words, if courteous behaviour is to be reciprocated by an
individual, #1
that individual needs to know that he was helped intentionally. #3

Intentional helping conditions induce significantly more courteous
behaviour than
control or unintentional conditions. #3

Unintentional conditions do not produce significantly more courteous behaviour
than control conditions. #3

However, intentional help apparently has to occur #1
under specific conditions. #3

In conditions with verbal contact, for example, #1
where a subject receives intentional help, #3
followed by additional contact with the confederate asking for directions, #3
courteous behaviour does not increase above that in control conditions. #3

The norm-of-reciprocity theory can account for this absence of increased
courteous behaviour. #3

The confederate acts courteously, #2
for example by holding a door open for the subject, #3
and the subject then reciprocates by answering the
confederate's questions. #3

Any obligation that the subject feels towards the confederate is removed #1
by his responding to the confederate's request. #3

Another reason why the contact condition does not produce
increased courteous behaviour #1
may be that the subject is not certain that the confederate held the
door open simply to be polite. #3

The subject might feel that the confederate was holding the door open
to reciprocate #1
for the help for which the confederate was asking. #3

Goldman, Florez & Fuller (1981) found that it is possible for an individual
to influence the likelihood #1
that another will behave in a courteous manner towards him. #3

Their results showed that the mean courteous behaviour for males was
significantly greater #1
than that for females. #3

Different conditions of helping yielded a significant difference in the rate of courteous behaviour. #3

The interaction of the main effects for sex versus helping behaviour was not significant. #3

When compared with each other, #1
the differences between means for control, unintentional and contact conditions were not significant. #3

However, the mean for intentional conditions of helping was significantly greater #1
than the means for each of the other conditions. #3

In the control condition, a confederate preceded a subject to an exit door #1
which led to a corridor, #3
opened the door, #3
and shut it behind him, #3
which required the subject to open the door himself. #3

In the corridor, the confederate bent down to adjust his socks, #2
which allowed the subject to arrive at the second door before the confederate. #3

In the unintentional helping condition, the confederate preceded the subject to the exit door #3
and, whilst holding the door open, #2
he bent down to tie his shoelace. #3

He took no notice of the subject #2
as the subject walked through the open door. #3

Then the confederate followed the subject to the second door. #3

In the intentional helping condition, the confederate preceded the subject to the exit door. #3

He held the door open, #3
allowed the subject to walk through, #3
and then followed the subject to the second door. #3

In the contact condition, the confederate preceded the subject to the exit door #2
and held the door open, #3
allowing the subject to pass through. #3

As soon as the subject had passed through the door, #2
the confederate asked the subject directions to the local theatre. #3

When the subject had responded #2
the confederate thanked him #2

and followed behind the subject to the second door. #3

The confederate was male, #3
and was dressed similarly to the subject population. #3

A trial was begun #2
when the confederate saw a subject walk towards the exit door of an indoor car-park. #3

Leaving the car-park required walking to the exit door, #2
opening the door #2
and proceeding through a corridor, #3
and finally opening a second door at the end of the corridor. #3

It was noted whether or not the subject held the second door open for the confederate. #3

Sex of subject was compared with door-holding behaviour #1
under the conditions of control, unintentional, contact and intentional helping. #3

Subjects were randomly assigned to each of the four conditions. #3

Eighty-four adult males and eighty-four adult females using the indoor car-park #1
next to a shopping centre #1
served unwittingly as subjects. #3

A two-by-four experimental design was used. #3

The study was designed to test various predictions concerning courteous behaviour. #3

It was hoped that such an investigation would reveal which factors influence courteous behaviour. #3

It seems at first sight that the norm-of-reciprocity theory
and the equity theory
would make similar predictions. #3

Both theories would predict that if A helps B, #1
then B will help A. #3

It is possible, however, to set up a situation #1
in which the two theories would make alternative predictions. #3

Equity theory implies that motivation to help will be induced #1
regardless of whether or not an imbalance has been produced intentionally. #3

Therefore, if A helps B unintentionally, #1
equity theory would predict that B would be likely to help A. #3

The norm-of-recprocity theory would predict that B would be likely to

help A #1
only if he thought that A had helped him intentionally. #3

If the norm-of-reciprocity theory is valid, #1
then the intentional helping should produce more frequent courteous behaviour #1
than no helping. #3

Equity theory would also predict this. #3

Equity theory would imply that courteous behaviour should occur #1
to the same extent #1
in both the intentional and the unintentional conditions. #3

The norm-of-reciprocity theory would predict that more courteous behaviour should occur in the intentional condition #1
than in the unintentional condition. #3

Social contact theory would predict that if B has had previous contact with A, #1
then B would be more likely to help A #1
than if this intial contact had not occurred. #3

It might also be predicted that males would engage in courteous behaviour #1
more frequently than females. #3

Social contact theory does appear to be related to helping behaviour. #3

This suggests that if A had previously had brief verbal contact with B, #1
then B would be more likely to help A #1
than if this initial contact had not occurred. #3

The norm-of-reciprocity theory states that people should help those who have helped them. #3

This suggests that if A has helped B, #1
then B would be more likely to reciprocate by helping A. #3

Equity theory states that interacting individuals try to balance their costs #1
and rewards. #3

This theory suggests that, if A has helped B, #1
then B has been rewarded at no cost to himself. #3

Therefore, B would be motivated to re-establish equity by helping A #1
at the first opportunity. #3

It has also been suggested that, #1
on the basis of previous research, #1
females are less prosocial than males. #3

Altruistic or prosocial behaviour involves activities #1

where one person assists another #1
but receives no obvious reward for his help. #3

These behaviours are generally regarded as thoughtful #1
or courteous. #3

It seems, in conclusion, #1
that it is possible to influence the likelihood of receiving help. #3

It also seems that the factors discussed here are indeed those which
influence helping behaviour. #3

Text2: Structured

It is possible that younger children may be less able than
older children or adults #1
to discriminate stimuli in the visual periphery. #3

Several developmental studies have shown younger subjects to be
less capable than older subjects #2
of both visual search and peripheral recognition of brief exposures. #3

It is unclear, however, whether these findings reflect actual
age differences #1
in peripheral visual capacity. #3

The better performance of the older subjects #1
on visual search tasks #1
may reflect more efficient use of what is seen in peripheral vision, #2
rather than better resolution within the visual system. #3

Because most tachistoscopic studies have failed to compare visual capacity at
various distances into the periphery with that at the fixation point, #2
these results may stem from other developmental trends in visual
processing #2
not specific to peripheral vision. #3

If there are developmental changes in peripheral discrimination capacities #1
beyond those due to changes in overall processing ability, #2
age differences should be greater for stimuli exposed in either
visual half-field #2
than for stimuli appearing at the centre of fixation. #3

If older subjects make better use than younger subjects of the
partial cues #1
available only in words, #2
these subjects should have special advantage in discriminating them #1
in peripheral vision. #3

A greater age difference in peripheral letter discriminability #1
for right half-field relative to left half-field #1
arrays would support the relevance of cerebral dominance. #3

If any enhancement of age differences #1
in the right half-field relative to the left half-field #1
is greater for horizontal than for vertical arrays, #1
a left-to-right postexposural scanning mechanism could be hypothesised. #3

This study explores the issue of development in letter perception #2
using a method that will optimise any potential differences. #3

The study compared age differences #1
in threshold discriminability for letter arrays #1
presented at the fixation point #2

versus to the right and left of fixation. #3

A further purpose was to explore other factors that might
contribute to greater ability of older #1
as compared with younger subjects #1
to discriminate peripheral letter arrays. #3

The experimental design consisted of the within-subject factor,
position of array, #1
and of the four between-subject factors, #1
type of array, orientation of array, age and sex. #3

Subjects were ninety-six nine year old children #1
and ninety-six university students, #3
all of whom were right handed. #3

Twenty-four children and twenty-four adults were assigned to each
of the four stimulus conditions. #3

Stimuli consisted of eighteen words and eighteen nonwords, #3
all containing four different letters. #3

All arrays consisted of capital letters typed in black on a white field #3
and presented via a two-field tachistoscope. #3

The threshold durations #1
at which subjects recognised letters at a given position in the
visual field #1
were determined by exposing arrays for increased durations across
trial blocks. #3

Thresholds were defined as the durations #1
yielding fifty percent letter recognition accuracy. #3

In order to determine whether age differences in peripheral vision
might be introduced
by a greater ability of older subjects to interpret cues received in
peripheral vision, #2
letter arrays forming both words and nonwords were used. #3

In order to investigate any further age advantage given to letters
in the right half-field #1
by age trends in cerebral dominance #1
and left-to-right postexposural scanning, #3
letter arrays were presented to the left and right half-fields #1
in vertical as well as horizontal orientation. #3

The conditions were therefore horizontal words, vertical words, horizontal
nonwords and vertical nonwords. #3

Threshold exposure durations were obtained for letter arrays
presented at each visual field position. #3

Between stimulus presentations a fixation dot was exposed #1
in the centre of the viewing field. #3

On each trial the subject was told to look directly into the centre dot #1
when the experimenter said "Look". #3

After stimulus exposure #1:
the subject was asked to report four different letters. #3

The eighteen test arrays were divided into three groups of six, #3
with each group assigned to a different one of the three visual field positions. #3

Test trials were administered in blocks of eighteen, #3
with each of the arrays presented only once per block. #3

After each block #1
the experimenter counted the number of letters correctly reported at each position. #3

When the subject had correctly reported nine or more letters, #2
the session ended. #3

The results of the analysis of variance showed significant age differences in letter perceptibility #1
for arrays in each of the three positions, #2
left half-field letters, right half-field letters and centre letters. #3

The analysis also revealed the predicted age versus position interaction #1
for left half-field versus centre comparisons #2
and for right half-field versus centre comparisons. #3

Letters presented in the left and right half-fields #1
were less discriminable #1
than letters presented at the fixation point #1
for both age groups. #3

Analyses of letter perceptibility failed to indicate a dependency of age differences #1
on the position, type or orientation of the letter array. #3

When these factors were examined independently of age, #2
analyses revealed letters within words to be more discriminable #1
than those within nonwords. #3

Results were consistent #1
with the expectation that letters would be better perceived in the right half-field #1
than in the left half-field. #3

The only stimulus condition that failed to yield right half-field superiority for either measure was vertical words. #3

Analyses failed to reveal main effects for sex #2
nor any interpretable interactions involving this factor. #3

The present data appear to be among the first supporting age
differences in the near periphery #2
using a task that explicitly requires peripheral discrimination #1
as opposed to mere detection. #3

Since the abilities of each of the two age groups to discriminate
peripheral letters
were assessed relative to their abilities to discriminate letters
presented at the fixation point, #3
a developmental difference specific to peripheral visual processing
seems indicated. #3

The fact that age differences in right half-field superiority
were not found #1
argues against any contribution to peripheral vision of developmental
trends in cerebral dominance #1
or left-to-right postexposural scanning. #3

Age differences in the ability to make use of redundancies within
peripheral words, #1
as opposed to nonwords, #1
also failed to appear. #3
Words were more discriminable than nonwords #3
and children were apparently as able as adults to make use of additional cues
in identifying letters within words. #3

It is possible that the visual systems of adults are better able
than those of children #1
to resolve peripheral stimuli. #3

This could be due to refinements in receptor apparatus, #2
such as enlargement of the pupil, #2
increased accommodation of the lens, #2
or growth of the peripheral retina. #3

However, other researchers have argued that growth of neural processes
play more of a role #1
than growth of peripheral processes in postinfancy development of acuity. #3

Another explanation for age differences in the ability to discriminate
peripheral letters #1
is that peripheral vision is suppressed in favour of foveal vision #1
to a greater extent in children than in adults. #3

According to this alternative, #1
information arising from the entire visual field is too great for
the young child to process all at once. #3

As a result, the child's effective vision is restricted to stimuli #1
in the immediate vicinity of the fixation point. #3

The present findings suggest that exploration of age related changes in the apparatus subserving peripheral vision #1
may be worthwhile. #3

This study also suggests #1
that age limitations in peripheral discrimination capacity may also imply a basic
sensory constraint on the ability #1
with which young children can make use of contextual cues from the periphery during reading. #2

Text2: Unstructured

Recent research seems to suggest #1
that exploration of age related changes in the apparatus
subserving peripheral
vision may be a worthwhile enterprise. #3

Studies also suggest #1
that age limitations in peripheral discrimination capacity may imply a basic
sensory constraint on the ability #1
with which young children can make use of contextual cues from the
periphery during reading. #3

Since, in these studies, #1
the abilities of young children and adults to discriminate peripheral
letters were assessed #1
relative to their abilities to discriminate letters presented at
the fixation point, #3
a developmental difference specific to peripheral visual
processing seems indicated. #3

The fact that age differences in right half-field superiority
were not found #1
argues against any contribution to peripheral vision of developmental
trends in cerebral dominance #1
or left-to-right postexposural scanning. #3

Neither have these studies found any age differences in the ability to make
use of redundancies within peripheral words, #2
as opposed to nonwords. #3

Words tend to be more discriminable than nonwords #3
and children are apparently as able as adults to make use of additional cues
available in identifying letters within words. #3

It is possible that the visual systems of adults are better able
than those of children #1
to resolve peripheral stimuli. #3

This could be due to refinements in receptor apparatus, #2
such as enlargement of the pupil, #2
increased accommodation of the lens, #2
or growth of the peripheral retina. #3

Other researchers have argued #1
that growth of neural processes play more of a role than growth of peripheral
processes in postinfancy development of acuity. #3

Another explanation for age differences in the ability to discriminate
peripheral letters #1
is that peripheral vision is suppressed in favour of foveal vision to a greater
extent in children than in adults. #3

According to this alternative, #1
information arising from the entire visual field is too great for the young child to process all at once. #3

As a result, #1
the child's effective vision is restricted to stimuli in the immediate vicinity of the fixation point. #3

Results of a study by Taylor (1982) appear to be among the first #1
reporting age differences in the near periphery #1
using a task that explicitly required peripheral discrimination #1
as opposed to mere detection. #3

These results showed significant age differences in letter perceptibility #1
for arrays in each of three positions, #1
left half-field letters, right half-field letters and centre letters. #3

However, Taylor's analysis also revealed an age versus position interaction #1
for left half-field versus centre comparisons #2
and for right half-field versus centre comparisons. #3

Letters presented in the left and right half-fields were less discriminable #1
than letters presented at the fixation point #1
for both adults and young children. #3

Analyses of letter perceptibility failed to indicate a dependency of age differences on the position, type or orientation of the letter array. #3

When these factors were examined independently of age, #2
analyses revealed letters within words to be more discriminable than those within nonwords. #3

Results showed that, #1
as expected, letters were better perceived in the right half-field than in the left half-field. #3

The only stimulus condition that failed to yield right half-field superiority #1
was words presented vertically. #3

There were no main effects for sex #3
nor any interpretable interactions involving this factor. #3

Threshold exposure durations were obtained for letter arrays presented at each visual field position. #3

Between stimulus presentations #1
a fixation dot was exposed in the centre of the viewing field. #3

On each trial the subject was told to look directly into the centre dot #1
when the experimenter said "Look". #3

After stimulus exposure #1
the subject was asked to report four different letters. #3

Eighteen test arrays were divided into three groups of six, #3
with each group assigned to a different one of three visual
field positions. #3

Test trials were administered in blocks of eighteen, #3
with each of the arrays presented only once per block. #3

After each block the experimenter counted the number of letters
correctly reported at each position. #3

When the subject had correctly reported fifty percent of the letters, #2
the session ended. #3

Taylor used letters forming both words and nonwords #1
in order to determine whether age differences might be due to a
greater ability of older subjects #1
to interpret cues received in peripheral vision. #3

Letter arrays were presented to the left and right half-fields #1
in vertical as well as horizontal orientation, #2
in order to investigate any further age advantage given to letters in the
right half-field by age trends in cerebral dominance #1
and left-to-right postexposural scanning. #3

The conditions that Taylor used were therefore horizontal words,
vertical words,
horizontal nonwords and vertical nonwords. #3

Using the method of ascending limits, #2
threshold exposure durations were obtained for letter arrays
presented at each visual field position. #3

Twenty-four children and twenty-four adults were assigned to
each of the four stimulus conditions. #3

Stimuli consisted of eighteen words and eighteen nonwords, #3
all containing four different letters. #3

All arrays consisted of capital letters typed in black on a white field #3
and presented via a two-field tachistoscope. #3

The threshold durations at which subjects recognised letters at
a given position
in the visual field were determined #1
by exposing arrays for increased durations across trial blocks. #3

Thresholds were defined as the durations yielding fifty percent
letter recognition accuracy. #3

Subjects were ninety-six nine year old children #1
and ninety-six university students, #3
all of whom were right handed. #3

The experimental design therefore, consisted of the within-subject
factor, position of array, #1
and of the four between-subject factors, #1
type of array, orientation of array, age and sex. #3

The study was designed to compare age differences in threshold discriminability
for letter arrays presented at the fixation point #1
versus to the right and left of fixation. #3

Another purpose was to explore other factors that might contribute to greater
ability of older as compared with younger subjects #1
to discriminate peripheral letter arrays. #3

None of the previous studies had explored this issue #1
using a method that would optimise any potential differences. #3

Taylor reasoned #1
that if there are developmental changes in peripheral discrimination capacities
beyond those due to changes in overall processing ability, #3
age differences should be greater for stimuli exposed in either visual half-field #1
than for stimuli appearing at the centre of fixation. #3

If older subjects make better use than younger subjects of the
partial cues available only in words, #2
then these subjects should have a special advantage in
discriminating them in peripheral vision. #3

A greater age difference in peripheral letter discriminability for right
half-field relative to left half-field arrays #1
would support the relevance of cerebral dominance. #3

If any enhancement of age differences in the right half-field relative to the
left half-field was greater for horizontal than for vertical arrays, #2
a left-to-right postexposural scanning mechanism could have been
hypothesised. #3

The better performance of older subjects on visual search tasks may reflect
more efficient use of what is seen in peripheral vision, #3
rather than better resolution within the visual system, #2
as some have suggested. #3

Because most tachistoscopic studies have failed to compare visual
capacity at
various distances into the periphery with that at the fixation point, #2
their results may stem from other developmental trends in visual processing #1
not specific to peripheral vision. #3

It is therefore unclear #1
whether their findings reflect actual differences in peripheral

visual capacity. #3

There have been numerous such studies showing younger subjects to be less able than older subjects #1
of both visual search and peripheral recognition of brief exposures. #3

All of these studies lend support for the possibility #1
that younger children may be less able than older children or adults to discriminate stimuli in the visual periphery. #3

APPENDIX G

DIVISIONS USED IN SCORING

Text1: Structured

1 (1) The problems being investigated in this study are the factors influencing courteous behaviour. [7]

2 (2) Researchers investigating prosocial behaviour have found that it is possible for a person to influence the likelihood of receiving help.

3 (3) Prosocial or altruistic behaviour involves activities where one person assists another but receives no obvious reward for his help.

4 (4) These kind of behaviours are generally regarded as thoughtful or courteous.

5 (5) Social contact theory is related to helping behaviour.

6 (6) This suggests that if A had previously had brief verbal contact with B, then B would be more likely to help A than if this initial contact had not occurred.

7 (7) Norm-of-reciprocity theory states that people should help those who have helped them.

8 (8) This theory suggests that if A has helped B, then B would be more likely to reciprocate by helping A.

9 (9) Equity theory states that interacting individuals try to balance their rewards and costs.

10 (10) This theory suggests that if A has helped B, then B has been rewarded at no cost to himself.

11 (11) Therefore, B would be motivated to re-establish a balance by helping A at the first opportunity.

12 (12) It has also been suggested that females are less prosocial than

---

7. The numbers in parentheses refer to the sentence numbers; the other numbers refer to the proposition numbers used for scoring.

males.

13 (13) At first it may seem that the norm-of-reciprocity theory and the equity theory would make similar predictions.

14 (14) Both theories would predict that if A helps B, then B would help A.

15 (15) However, it is possible to set up a situation in which the two theories would make alternative predictions.

16 (16) Equity theory implies that motivation to help will be induced regardless of whether or not the imbalance has been produced intentionally.

17 (17) Therefore, if A helps B unintentionally, equity theory would predict that B would be likely to help A.

18 (18) However, norm-of-reciprocity theory would predict that B would be likely to help A only if he thought that A had helped him intentionally.

19 (19) If the norm-of-reciprocity theory is valid, then the intentional courtesy should produce more frequent courteous behaviour than no courtesy.

20 (20) Equity theory would also predict this.

21 (21) Equity theory would imply that courteous behaviour should occur to the same extent in both the unintentional and the intentional conditions.

22 (22) The norm-of-reciprocity theory would predict that more courteous behaviour should occur in the intentional than in the unintentional condition.

23 (23) Social contact theory would predict that if B has had previous contact with A, then B would be more likely to help than if no contact had occurred.

24 (24) It would also be predicted that males would show more frequent courteous behaviour than females.

25 (25) An investigation of these hypotheses might reveal which factors influence courteous behaviour.

26 (26) The present study was designed to test the aforementioned predictions.

27 (27) A two-by-four experimental design was used.

28 (28) Eighty-four adult males and eighty-four adult females who were using a car-park next to a shopping centre served unwittingly as subjects.

29 (29) Sex, male and female, was compared with door-holding behaviour under conditions of control, unintentional, contact and intentional.

30 (30) Subjects were randomly assigned to each of the four conditions.

31 (31) A male confederate,

31 (32) dressed similarly to the subject population,

31 (33) waited for a subject to park his car and walk towards the car-park exit.

32 (34) A trial was begun

32 (35) when the confederate saw a subject walking alone towards the exit door.

33 (36) In order to leave the car-park,

33 (37) subjects had to walk towards the exit door,

33 (38) open the door,

33 (39) proceed through a corridor,

33 (40) and finally open a second door at the end of the corridor.

34 (41) It was noted whether or not the subject held the second door open for the confederate.

35 (42) In the control condition, the confederate preceded the subject to the exit door,

35 (43) opened the door

35 (44) and shut it behind him.

36 (45) This required the subject to open the door for himself.

37 (46) In the corridor, he confederate bent down to adjust his socks,

37 (47) which allowed the subject to arrive at the second door before the confederate.

38 (48) In the unintentional condition, the confederate preceded the subject to the exit door

38 (49) and, whilst holding the door open,

38 (50) he bent down to tie his shoelaces.

39 (51) He took no notice of the subject

39 (52) as the subject walked through the open door.

40 (53) Then the confederate followed the subject to the second door.

41 (54) In the intentional condition, the confederate preceded the subject to the exit door.

42 (55) He held the door open,

42 (56) allowed the subject to walk through,

42 (57) and then followed the subject to the second door.

43 (58) In the contact condition, the confederate preceded the subject to the exit door

43 (59) and held the door open,

43 (60) allowing the subject to pass through.

44 (61) As soon as the subject had passed through the door,

44 (62) the confederate asked the subject for directions to the local theatre.

45 (63) When the subject had responded

45 (64) the confederate thanked him

45 (65) and followed behind the subject to the second door.

46 (66) The results of an analysis of variance showed that the mean courteous behaviour for males was significantly greater than for females.

47 (67) The different door holding conditions yielded a significant difference in the rate of courteous behaviour.

48 (68) The interaction of the main effects was not significant.

49 (69) When compared with each other, the differences between means for control, unintentional and contact conditions were not significant.

50 (70) The mean for the intentional condition was significantly greater than the means for each of the other conditions.

51 (71) The present study found that it is possible for an individual to influence the likelihood that another will behave in a courteous manner towards him.

52 (72) The intentional condition induced significantly more courteous behaviour than did the control or unintentional conditions.

53 (73) The unintentional condition did not produce significantly more courteous behaviour than the control condition.

54 (74) However, it appears that intentional help must occur under specific conditions.

55 (75) In the contact condition, the subject received intentional help.

56 (76) This was followed by additional contact with the confederate asking for directions.

57 (77) Courteous behaviour did not increase above that in the control condition.

58 (78) The norm-of-reciprocity theory can account for this absence of increased door-holding behaviour.

59 (79) The confederate acted courteously,

59 (80) holding the door open for the subject.

60 (81) The subject then reciprocated by answering the confederate's questions.

61 (82) Any obligation that the subject felt towards the confederate was removed by his responding to the confederate's request.

62 (83) Another reason why the contact treatment did not produce increased courteous behaviour might have been that the subject was not certain that the confederate held the door open simply to be polite.

63 (84) The subject might have understood that the confederate was holding the door open to reciprocate for the help for which the confederate was asking.

64 (85) It appears, therefore, that the results support the norm-of-reciprocity theory

64 (86) and do not support the equity theory.

65 (87) If courteous behaviour is to be reciprocated by an individual, then that individual needs to know that he was helped intentionally.

Text1: Unstructured

1 (1) Recent research into factors affecting courteous behaviour seems to support the norm-of-reciprocity theory

1 (2) and not to support the equity theory.

2 (3) In other words, if courteous behaviour is to be reciprocated by an individual, that individual needs to know that he was helped intentionally.

3 (4) Intentional helping conditions induce significantly more courteous behaviour than control or unintentional conditions.

4 (5) Unintentional conditions do not produce significantly more courteous behaviour than control conditions.

5 (6) However, intentional help apparently has to occur under specific conditions.

6 (7) In conditions with verbal contact, for example, where a subject receives intentional help,

6 (8) followed by additional contact with the confederate asking for directions,

6 (9) courteous behaviour does not increase above that in control conditions.

7 (10) The norm-of-reciprocity theory can account for this absence of increased courteous behaviour.

8 (11) The confederate acts courteously,

8 (12) for example by holding a door open for the subject,

8 (13) and the subject then reciprocates by answering the confederate's questions.

9 (14) Any obligation that the subject feels towards the confederate is removed by his responding to the confederate's request.

10 (15) Another reason why the contact condition does not produce increased courteous behaviour may be that the subject is not certain that the confederate held the door open simply to be polite.

11 (16) The subject might feel that the confederate was holding the door open to reciprocate for the help for which the confederate was asking.

12 (17) Goldman, Florez & Fuller (1981) found that it is possible for an individual to influence the likelihood that another will behave in a courteous

manner towards him.

13 (18) Their results showed that the mean courteous behaviour for males was significantly greater than that for females.

14 (19) Different conditions of helping yielded a significant difference in the rate of courteous behaviour.

15 (20) The interaction of the main effects for sex versus helping behaviour was not significant.

16 (21) When compared with each other, the differences between means for control, unintentional and contact conditions were not significant.

17 (22) However, the mean for intentional conditions of helping was significantly greater than the means for each of the other conditions.

18 (23) In the control condition, a confederate preceded a subject to an exit door which led to a corridor,

18 (24) opened the door,

18 (25) and shut it behind him,

18 (26) which required the subject to open the door himself.

19 (27) In the corridor, the confederate bent down to adjust his socks,

19 (28) which allowed the subject to arrive at the second door before the confederate.

20 (29) In the unintentional helping condition, the confederate preceded the subject to the exit door

20 (30) and, whilst holding the door open,

20 (31) he bent down to tie his shoelace.

21 (32) He took no notice of the subject

21 (33) as the subject walked through the open door.

22 (34) Then the confederate followed the subject to the second door.

23 (35) In the intentional helping condition, the confederate preceded the subject to the exit door.

24 (36) He held the door open,

24 (37) allowed the subject to walk through,

24 (38) and then followed the subject to the second door.

25 (39) In the contact condition, the confederate preceded the subject to the exit door

25 (40) and held the door open,

25 (41) allowing the subject to pass through.

26 (42) As soon as the subject had passed through the door,

26 (43) the confederate asked the subject directions to the local theatre.

27 (44) When the subject had responded

27 (45) the confederate thanked him

27 (46) and followed behind the subject to the second door.

28 (47) The confederate was male,

28 (48) and was dressed similarly to the subject population.

29 (49) A trial was begun

29 (50) when the confederate saw a subject walk towards the exit door of an indoor car-park.

30 (51) Leaving the car-park

30 (52) required walking to the exit door,

30 (53) opening the door

30 (54) and proceeding through a corridor,

30 (55) and finally opening a second door at the end of the corridor.

31 (56) It was noted whether or not the subject held the second door open for the confederate.

32 (57) Sex of subject was compared with door-holding behaviour under the conditions of control, unintentional, contact and intentional helping.

33 (58) Subjects were randomly assigned to each of the four conditions.

34 (59) Eighty-four adult males and eighty-four adult females using the indoor car-park next to a shopping centre served unwittingly as subjects.

35 (60) A two-by-four experimental design was used.

36 (61) The study was designed to test various predictions concerning courteous behaviour.

37 (62) It was hoped that such an investigation would reveal which factors influence courteous behaviour.

38 (63) It seems at first sight that the norm-of-reciprocity theory and the equity theory would make similar predictions.

39 (64) Both theories would predict that if A helps B, then B will help A.

40 (65) It is possible, however, to set up a situation in which the two theories would make alternative predictions.

41 (66) Equity theory implies that motivation to help will be induced regardless of whether or not an imbalance has been produced intentionally.

42 (67) Therefore, if A helps B unintentionally, equity theory would predict that B would be likely to help A.

43 (68) The norm-of-recprocity theory would predict that B would be likely to help A only if he thought that A had helped him intentionally.

44 (69) If the norm-of-reciprocity theory is valid, then the intentional helping should produce more frequent courteous behaviour than no helping.

45 (70) Equity theory would also predict this.

46 (71) Equity theory would imply that courteous behaviour should occur to the same extent in both the intentional and the unintentional conditions.

47 (72) The norm-of-reciprocity theory would predict that more courteous behaviour should occur in the intentional condition than in the unintentional condition.

48 (73) Social contact theory would predict that if B has had previous contact with A, then B would be more likely to help A than if this intial contact had not occurred.

49 (74) It might also be predicted that males would engage in courteous behaviour more frequently than females.

50 (75) Social contact theory does appear to be related to helping behaviour.

51 (76) This suggests that if A had previously had brief verbal contact with B, then B would be more likely to help A than if this initial contact had not occurred.

52 (77) The norm-of-reciprocity theory states that people should help those who have helped them.

53 (78) This suggests that if A has helped B, then B would be more likely to reciprocate by helping A.

54 (79) Equity theory states that interacting individuals try to balance their costs and rewards.

55 (80) This theory suggests that, if A has helped B, then B has been rewarded at no cost to himself.

56 (81) Therefore, B would be motivated to re-establish equity by helping A at the first opportunity.

57 (82) It has also been suggested that, on the basis of previous research, females are less prosocial than males.

58 (83) Altruistic or prosocial behaviour involves activities where one person

assists another but receives no obvious reward for his help.

59 (84) These behaviours are generally regarded as thoughtful or courteous.

60 (85) It seems, in conclusion, that it is possible to influence the likelihood of receiving help.

61 (86) It also seems that the factors discussed here are indeed those which influence helping behaviour.

Text2: Structured

1 (1) It is possible that younger children may be less able than older children or adults to discriminate stimuli in the visual periphery.

2 (2) Several developmental studies have shown younger subjects to be less capable than older subjects

2 (3) of both visual search and peripheral recognition of brief exposures.

3 (4) It is unclear, however, whether these findings reflect actual age differences in peripheral visual capacity.

4 (5) The better performance of the older subjects on visual search tasks may reflect more efficient use of what is seen in peripheral vision,

4 (6) rather than better resolution within the visual system.

5 (7) Because most tachistoscopic studies have failed to compare visual capacity at various distances into the periphery with that at the fixation point,

5 (8) these results may stem from other developmental trends in visual processing

5 (9) not specific to peripheral vision.

6 (10) If there are developmental changes in peripheral discrimination capacities beyond those due to changes in overall processing ability,

6 (11) age differences should be greater for stimuli exposed in either visual half-field

6 (12) than for stimuli appearing at the centre of fixation.

7 (13) If older subjects make better use than younger subjects of the partial cues available only in words,

7 (14) these subjects should have special advantage in discriminating them in peripheral vision.

8 (15) A greater age difference in peripheral letter discriminability for right half-field relative to left half-field arrays would support the relevance of cerebral dominance.

9 (16) If any enhancement of age differences in the right half-field relative to the left half-field is greater for horizontal than for vertical arrays,

9 (17) a left-to-right postexposural scanning mechanism could be hypothesised.

10 (18) This study explores the issue of development in letter perception

10 (19) using a method that will optimise any potential differences.

11 (20) The study compared age differences in threshold discriminability for letter arrays presented at the fixation point

11 (21) versus to the right and left of fixation.

12 (22) A further purpose was to explore other factors that might contribute to greater ability of older as compared with younger subjects to discriminate peripheral letter arrays.

13 (23) The experimental design consisted of the within-subject factor, position of array, and of the four between-subject factors, type of array, orientation of array, age and sex.

14 (24) Subjects were ninety-six nine year old children and ninety-six university students,

14 (25) all of whom were right handed.

15 (26) Twenty-four children and twenty-four adults were assigned to each of the four stimulus conditions.

16 (27) Stimuli consisted of eighteen words and eighteen nonwords,

16 (28) all containing four different letters.

17 (29) All arrays consisted of capital letters typed in black on a white field

17 (30) and presented via a two-field tachistoscope.

18 (31) The threshold durations at which subjects recognised letters at a given position in the visual field were determined by exposing arrays for increased durations across trial blocks.

19 (32) Thresholds were defined as the durations yielding fifty percent letter recognition accuracy.

20 (33) In order to determine whether age differences in peripheral vision might be introduced by a greater ability of older subjects to interpret cues received in peripheral vision,

20 (34) letter arrays forming both words and nonwords were used.

21 (35) In order to investigate any further age advantage given to letters in the right half-field by age trends in cerebral dominance and left-to-right postexposural scanning,

21 (36) letter arrays were presented to the left and right half-fields in vertical as well as horizontal orientation.

22 (37) The conditions were therefore horizontal words, vertical words,

horizontal nonwords and vertical nonwords.

23 (38) Threshold exposure durations were obtained for letter arrays presented at each visual field position.

24 (39) Between stimulus presentations a fixation dot was exposed in the centre of the viewing field.

25 (40) On each trial the subject was told to look directly into the centre dot when the experimenter said "Look".

26 (41) After stimulus exposure the subject was asked to report four different letters.

27 (42) The eighteen test arrays were divided into three groups of six,

27 (43) with each group assigned to a different one of the three visual field positions.

28 (44) Test trials were administered in blocks of eighteen,

28 (45) with each of the arrays presented only once per block.

29 (46) After each block the experimenter counted the number of letters correctly reported at each position.

30 (47) When the subject had correctly reported nine or more letters,

30 (48) the session ended.

31 (49) The results of the analysis of variance showed significant age differences in letter perceptibility for arrays in each of the three positions,

31 (50) left half-field letters, right half-field letters and centre letters.

32 (51) The analysis also revealed the predicted age versus position interaction for left half-field versus centre comparisons

32 (52) and for right half-field versus centre comparisons.

33 (53) Letters presented in the left and right half-fields were less discriminable than letters presented at the fixation point for both age groups.

34 (54) Analyses of letter perceptibility failed to indicate a dependency of age differences on the position, type or orientation of the letter array.

35 (55) When these factors were examined independently of age,

35 (56) analyses revealed letters within words to be more discriminable than those within nonwords.

36 (57) Results were consistent with the expectation that letters would be better perceived in the right half-field than in the left half-field.

37 (58) The only stimulus condition that failed to yield right half-field superiority for either measure was vertical words.

38 (59) Analyses failed to reveal main effects for sex

38 (60) nor any interpretable interactions involving this factor.

39 (61) The present data appear to be among the first supporting age differences in the near periphery

39 (62) using a task that explicitly requires peripheral discrimination as opposed to mere detection.

40 (63) Since the abilities of each of the two age groups to discriminate peripheral letters were assessed relative to their abilities to discriminate letters presented at the fixation point,

40 (64) a developmental difference specific to peripheral visual processing seems indicated.

41 (65) The fact that age differences in right half-field superiority were not found argues against any contribution to peripheral vision of developmental trends in cerebral dominance or left-to-right postexposural scanning.

42 (66) Age differences in the ability to make use of redundancies within peripheral words, as opposed to nonwords, also failed to appear.

43 (67) Words were more discriminable than nonwords

43 (68) and children were apparently as able as adults to make use of additional cues in identifying letters within words.

44 (69) It is possible that the visual systems of adults are better able than those of children to resolve peripheral stimuli.

45 (70) This could be due to refinements in receptor apparatus,

45 (71) such as enlargement of the pupil,

45 (72) increased accommodation of the lens,

45 (73) or growth of the peripheral retina.

46 (74) However, other researchers have argued that growth of neural processes play more of a role than growth of peripheral processes in postinfancy development of acuity.

47 (75) Another explanation for age differences in the ability to discriminate peripheral letters is that peripheral vision is suppressed in favour of foveal vision to a greater extent in children than in adults.

48 (76) According to this alternative, information arising from the entire visual field is too great for the young child to process all at once.

49 (77) As a result, the child's effective vision is restricted to stimuli in the immediate vicinity of the fixation point.

50 (78) The present findings suggest that exploration of age related changes in the apparatus subserving peripheral vision may be worthwhile.

51 (79) This study also suggests that age limitations in peripheral discrimination capacity may also imply a basic sensory constraint on the ability with which young children can make use of contextual cues from the periphery during reading.

Text2: Unstructured

1 (1) Recent research seems to suggest that exploration of age related changes in the apparatus subserving peripheral vision may be a worthwhile enterprise.

2 (2) Studies also suggest that age limitations in peripheral discrimination capacity may imply a basic sensory constraint on the ability with which young children can make use of contextual cues from the periphery during reading.

3 (3) Since, in these studies, the abilities of young children and adults to discriminate peripheral letters were assessed relative to their abilities to discriminate letters presented at the fixation point,

3 (4) a developmental difference specific to peripheral visual processing seems indicated.

4 (5) The fact that age differences in right half-field superiority were not found argues against any contribution to peripheral vision of developmental trends in cerebral dominance or left-to-right postexposural scanning.

5 (6) Neither have these studies found any age differences in the ability to make use of redundancies within peripheral words, as opposed to nonwords.

6 (7) Words tend to be more discriminable than nonwords

6 (8) and children are apparently as able as adults to make use of additional cues available in identifying letters within words.

7 (9) It is possible that the visual systems of adults are better able than those of children to resolve peripheral stimuli.

8 (10) This could be due to refinements in receptor apparatus,

8 (11) such as enlargement of the pupil,

8 (12) increased accommodation of the lens,

8 (13) or growth of the peripheral retina.

9 (14) Other researchers have argued that growth of neural processes play more of a role than growth of peripheral processes in postinfancy development of acuity.

10 (15) Another explanation for age differences in the ability to discriminate peripheral letters is that peripheral vision is suppressed in favour of foveal vision to a greater extent in children than in adults.

11 (16) According to this alternative, information arising from the entire visual field is too great for the young child to process all at once.

12 (17) As a result, the child's effective vision is restricted to stimuli in the immediate vicinity of the fixation point.

13 (18) Results of a study by Taylor (1982) appear to be among the first reporting age differences in the near periphery

13 (19) using a task that explicitly required peripheral discrimination as opposed to mere detection.

14 (20) These results showed significant age differences in letter perceptibility for arrays in each of three positions,

14 (21) left half-field letters, right half-field letters and centre letters.

15 (22) However, Taylor's analysis also revealed an age versus position interaction for left half-field versus centre comparisons

15 (23) and for right half-field versus centre comparisons.

16 (24) Letters presented in the left and right half-fields were less discriminable than letters presented at the fixation point for both adults and young children.

17 (25) Analyses of letter perceptibility failed to indicate a dependency of age differences on the position, type or orientation of the letter array.

18 (26) When these factors were examined independently of age,

18 (27) analyses revealed letters within words to be more discriminable than those within nonwords.

19 (28) Results showed that, as expected, letters were better perceived in the right half-field than in the left half-field.

20 (29) The only stimulus condition that failed to yield right half-field superiority was words presented vertically.

21 (30) There were no main effects for sex

21 (31) nor any interpretable interactions involving this factor.

22 (32) Threshold exposure durations were obtained for letter arrays presented at each visual field position.

23 (33) Between stimulus presentations a fixation dot was exposed in the centre of the viewing field.

24 (34) On each trial the subject was told to look directly into the centre dot when the experimenter said "Look".

25 (35) After stimulus exposure the subject was asked to report four different letters.

26 (36) Eighteen test arrays were divided into three groups of six,

26 (37) with each group assigned to a different one of three visual field

positions.

27 (38) Test trials were administered in blocks of eighteen,

27 (39) with each of the arrays presented only once per block.

28 (40) After each block the experimenter counted the number of letters correctly reported at each position.

29 (41) When the subject had correctly reported fifty percent of the letters,

29 (42) the session ended.

30 (43) Taylor used letters forming both words and nonwords

30 (44) in order to determine whether age differences might be due to a greater ability of older subjects to interpret cues received in peripheral vision.

31 (45) Letter arrays were presented to the left and right half-fields in vertical as well as horizontal orientation,

31 (46) in order to investigate any further age advantage given to letters in the right half-field by age trends in cerebral dominance and left-to-right postexposural scanning.

32 (47) The conditions that Taylor used were therefore horizontal words, vertical words, horizontal nonwords and vertical nonwords.

33 (48) Using the method of ascending limits,

33 (49) threshold exposure durations were obtained for letter arrays presented at each visual field position.

34 (50) Twenty-four children and twenty-four adults were assigned to each of the four stimulus conditions.

35 (51) Stimuli consisted of eighteen words and eighteen nonwords,

35 (52) all containing four different letters.

36 (53) All arrays consisted of capital letters typed in black on a white field

36 (54) and presented via a two-field tachistoscope.

37 (55) The threshold durations at which subjects recognised letters at a given position in the visual field were determined by exposing arrays for increased durations across trial blocks.

38 (56) Thresholds were defined as the durations yielding fifty percent letter recognition accuracy.

39 (57) Subjects were ninety-six nine year old children and ninety-six university students,

39 (58) all of whom were right handed.

40 (59) The experimental design therefore, consisted of the within-subject factor, position of array, and of the four between-subject factors, type of array, orientation of array, age and sex.

41 (60) The study was designed to compare age differences in threshold discriminability for letter arrays presented at the fixation point

41 (61) versus to the right and left of fixation.

42 (62) Another purpose was to explore other factors that might contribute to greater ability of older as compared with younger subjects to discriminate peripheral letter arrays.

43 (63) None of the previous studies had explored this issue

43 (64) using a method that would optimise any potential differences.

44 (65) Taylor reasoned that if there are developmental changes in peripheral discrimination capacities beyond those due to changes in overall processing ability,

44 (66) age differences should be greater for stimuli exposed in either visual half-field

44 (67) than for stimuli appearing at the centre of fixation.

45 (68) If older subjects make better use than younger subjects of the partial cues available only in words,

45 (69) then these subjects should have a special advantage in discriminating them in peripheral vision.

46 (70) A greater age difference in peripheral letter discriminability for right half-field relative to left half-field arrays would support the relevance of cerebral dominance.

47 (71) If any enhancement of age differences in the right half-field relative to the left half-field was greater for horizontal than for vertical arrays,

47 (72) a left-to-right postexposural scanning mechanism could have been hypothesised.

48 (73) The better performance of older subjects on visual search tasks may reflect more efficient use of what is seen in peripheral vision,

48 (74) rather than better resolution within the visual system,

48 (75) as some have suggested.

49 (76) Because most tachistoscopic studies have failed to compare visual capacity at various distances into the periphery with that at the fixation point,

49 (77) their results may stem from other developmental trends in visual

processing

49 (78) not specific to peripheral vision.

50 (79) It is therefore unclear whether their findings reflect actual differences in peripheral visual capacity.

51 (80) There have been numerous such studies showing younger subjects to be less able than older subjects

51 (81) of both visual search and peripheral recognition of brief exposures.

52 (82) All of these studies lend support for the possibility that younger children may be less able than older children or adults to discriminate stimuli in the visual periphery.