

THE UNIVERSITY OF HULL

ANALYTIC AND CONSTRUCTIVE PROCESSES  
IN THE COMPREHENSION OF TEXT

being a Thesis submitted for the Degree of

DOCTOR OF PHILOSOPHY

in the University of Hull

by

Patricia Anne Lindley (B.A., Open)

February, 1983.

## PREFACE.

### Analytic and Constructive processes in the Comprehension of Text.

This thesis explores the process of comprehension as a purposeful interaction between a reader and the information in a text. The review begins by discussing the difference between educational and psychological perspectives on comprehension. Approaches to the analysis of text structure are then described and models and theories of the representation of knowledge are evaluated. It is argued that these are limited in that they tend to focus either on the text or the reader: they either examine those procedures that are necessary for text analysis or the knowledge structures required for comprehension, storage and retrieval. Those that come nearest to examining the interaction between text and knowledge structures tend to be limited in terms of the texts they can deal with and they do not deal adequately with the predictive aspects of comprehension.

Experiments are reported which look at the ongoing predictions made by readers, and how these are affected by factors such as text structure and "Interestingness". The experiments provided the opportunity for examining the potential of alternative methodologies (such as the content analysis of open-ended questions). It is felt that it is necessary to examine comprehension using methods which are direct but not intrusive. The studies reported demonstrate that it is possible to obtain reliable measures of a reader's predictions and that these are systematically affected by the structure and content of the text.

## ACKNOWLEDGEMENTS

I would like to thank Dr. David Bartram for his consistent supervision, advice and encouragement during the formation of this thesis. I would also like to thank him and both Mrs. Lorraine Orvis and Mrs. Linda Randerson for their advice and help in the final preparation of the manuscript.

The experiments reported in the thesis were carried out between October 1979 and September 1980, when I was seconded by the Wakefield Metropolitan Education Authority and the Home Office Prison Education Department, and between September 1980 and October 1982 when I was in receipt of a University of Hull Postgraduate Award. I am grateful to all the above for the financial support they afforded me.

Finally I would like to thank my husband, my son Carl and especially my daughter Susan, for their personal support and encouragement throughout.

## CONTENTS

## INTRODUCTION.

What a competent reader would need to have to comprehend a text.	1
Comprehension as a learning process.	7
A summary of the contents.	10
CHAPTER ONE. Educational perspectives and research into comprehension.	15
Comprehension: a part of a hierarchy of skills.	15
Comprehension: central to the reading process.	17
Statistical Analyses of the Skills of Comprehension.	19
Summary of the educational perspectives and research.	26
CHAPTER TWO. Linguistic models which have practical and theoretical implications for models and theories of comprehension.	28
The theories of Chomsky (1957, 1965) and Fillmore (1968)	29
Chomsky's (1957, 1965) models of generative transformational grammar.	29
Fillmore's (1968) Case Grammar.	39
The limitations of the linguistic models as models of comprehension.	45
CHAPTER THREE. Semantic Memory.	47
Early network models of semantic memory.	48
Alternatives to semantic network models.	57
More recent models of semantic networks.	60
CHAPTER FOUR. Concept-driven Processes.	66
Knowledge based structures.	67
Story grammars and narrative structures in simple stories.	78
General discussion of the models.	89

Contents	(ii)
CHAPTER FIVE. A Model of Text Comprehension Developed by Kintsch and his Associates.	93
The Microstructural Process: The establishment of a coherent text base (Kintsch, 1974).	94
The Macrostructural process.	96
Kintsch and van Dijk's (1978) model.	101
A Modification of the 1978 model (Kintsch and Vipond (1979)	106
Problems connected with the 1979 model.	108
CHAPTER SIX. Psychological evidence for Kintsch's models.	112
Text characteristics and the empirical evidence which supports the existence of microstructural processes related to them.	112
The processing capacity of the reader.	117
Evidence that the reader processes the text as proposed by the Kintsch and Vipond (1979) model.	119
Summary of the evidence for the existence of the processes described by the models.	125
Conclusions.	127
CHAPTER SEVEN. Scenarios.	130
The model of text comprehension produced by Sanford and Garrod (1981).	
The development of the Scenario model.	131
The Scenario model.	138
A comparison of the Kintsch (1979) model with the Scenario model.	143
CHAPTER EIGHT. Elaborations: Associations and Constructions.	145
Introduction	145
The construction of meaning.	148
Elaborations.	159
Elaborations as semantic associations.	160
Elaborations generated in context.	164
Conclusions.	168

Contents	(iii)
CHAPTER NINE. Adequacy of the models and measurements reviewed for the analysis and comprehension of Complex Tests.	169
Relationship between the content and structure in the Comprehension of a text.	169
Representation of the content of text in Memory.	178
The processes which may be involved in the Comprehension of text.	179
Implications of the Issues/Models for the comprehension of Complex tests.	185
CHAPTER TEN. Experiment One: A methodological exploration.	191
Introduction.	191
Methodological Problems.	197
The Experiment.	202
CHAPTER ELEVEN. Experiment Two: The effect of Background and Outcome on predictions about Events and hypotheses about Characters.	222
CHAPTER TWELVE. The effects of sequential order on the processing of "Interesting" content.	265
Experiment Three: An experiment to investigate the concept of absolute interestingness.	267
Experiment Four: Ratings of interestingness of sentences by subjects who have previously read the sentence in context.	271
Experiment Five: The effects of "Interestingness" on sentence choice in text.	275
CHAPTER THIRTEEN. Experiment Six: The effects of text and variables on predictions.	287
CONCLUSIONS.	326
Comprehension: the interaction between text and reader.	326
General and educational implications.	330
Methodology.	332

REFERENCES.	335
APPENDIX 1. Questions given to subjects in Experiment One.	352
APPENDIX 2. Questions given to subjects in Experiment Two.	355
APPENDIX 3. Experiment 2a: the reliability of response categorisation.	359
3a. Description of categories.	362
3b. The text.	364
3c. The responses.	365
3d. Table of responses assigned to each category.	366
APPENDIX 4. Texts used in Experiment 5.	367

## TABLES

## EXPERIMENT ONE:

10:1 (i)	Open-ended questions about the events.	209
10:1 (ii)	Open-ended questions about the characters.	211
10:1 (iii)	Open-ended questions about the wider issues.	211
10:2 (i)	Multi-choice questions.	213
10:2 (ii)	The fixed choice question.	214
10:3	Rating Scales.	215

## EXPERIMENT TWO:

11:1	Question One.	232
11:2	Question Two.	234
11:3	Question Three (the soldier).	235
11:4	Question Three (Jenny).	236
11:5	Question 4a.	238
11:6	Question 4b.	239
11:7	Question 6a.	242
11:8	Question 6b.	243
11:9	Question 12.	244
11:10	Question 8 (Self-esteem: Jenny).	246
11:11	Question 8 (Self-esteem: the soldier).	247
11:12	Question 8 (Attractiveness: Jenny).	249
11:13	Question 8 (Attractiveness: the soldier).	250
11:14	Question 10a (Jenny).	251
11:15	Question 10a (the soldier).	252
11:16	Question 10b (Jenny).	254
11:17	Question 10b (the soldier).	255
11:18	Question 11.	257
11:19	Question 5.	259



## EXPERIMENT THREE:

12:1 Mean ratings of "Interestingness". 270

## EXPERIMENT FOUR:

12:2 Mean ratings of "Interestingness". 273

12:3 Consistency of sentence ratings for each group. 274

12:4 Intergroup correlations of ratings. 274

## EXPERIMENT FIVE:

12:5 Balance of rated "Interestingness" and choice of 281 sentences.

12:6 Texts presented in canonical order. 282

12:7 Texts presented in reverse order. 282

12:8 Texts with and without time order in Condition B. 283

12:9 Texts with a time sequence in Conditions A and B. 284

## EXPERIMENT SIX:

13:1 Mean ratings for male and female subjects under each 300 condition for each of the 18 scales.

13:2 Matrix of correlations between group means. 302

13:3 Comparison between pairs of conditions. 302

13:4 Results of Analyses of variance between sex and role 307 of characters for pairs of scales.

13:5 Summary of results. 310

13:6 Assigned sex of character A. 315

13:7 Assigned sex of character B. 316

13:8 Attribution of responsibility. 318

## FIGURES

10:1 Text used in Experiments One and Two. 205

13:1 Texts used in Experiment Six. 293

INTRODUCTION

This thesis is concerned with both theoretical analyses and empirical investigations of the comprehension of text. The definition of "text" in this context is a loose one. It is pragmatically defined as the stimulus material which subjects have been asked to "comprehend" during experimental investigations or the written material which is used to illustrate a theoretical model. As a result a "text" can vary in length from a sentence to a simple short story. In all cases it is longer than a single word and more connected than a list of words.

The definition of "comprehension" is far less simple. It could be defined simply as whatever the reader does when reading a text. However in an experimental situation, comprehension is being assessed by the experimenter according to a preselected criterion and in relation to a text which has been preselected for a particular purpose. What the reader does in this situation may not be what he/she would do in reading a self selected text for a self determined purpose. Therefore any complete definition of comprehension will have to include the purpose for reading the text. A very tentative and general working definition of comprehension at this point is that it is a purposeful interaction between a reader and a text leading to the goal of "meaningful" (to the reader) interpretation. Bartlett (1932) called this process "effort after meaning". Such a definition leaves open the question of what the reader might be doing in the course of the interaction, and the purpose for reading the text and what is meant by "meaningful".

What a competent reader would need to have to comprehend a text.

If a reader is to comprehend a text he or she must have sufficient knowledge about the text structure and content to be able to decide whether the text is comprehensible. Comprehensibility is not all or nothing but lies somewhere on a continuum from easy to impossible. The position on the

continuum where any particular text is placed is a function of the interaction between the reader's knowledge and features in the text. A text cannot be easy to comprehend unless the reader has the necessary knowledge to understand it. The greater the "effort" the reader must expend to find or create a "meaning", the further along the continuum towards impossible the text will lie.

If the rather tentative definition of comprehension is to be expanded, features of the text and the knowledge the reader must have in order to make comprehension easy will have to be explored. This means that both the text and the reader's knowledge base will have to be analysed in a complex manner. While aspects of the analysis will be described separately, this is not to imply that they are independent aspects or that they are stages in a process. The separation is for convenience of description. A series of questions may be posed to aid these analyses:

- (i) Are the words known to the reader?
- (ii) Are the words combined in an acceptable order?
- (iii) Are the concepts related to each other in an appropriate manner?
- (iv) Is the content acceptable, possible or plausible in the wider context of the reader's world knowledge?

These questions require that the reader has sufficient and specific linguistic knowledge in addition to some pre-existing knowledge about the topic to be able to analyse the text. Comprehension will be easy if the reader has all the necessary knowledge to answer these questions. It will be impossible if he/she has none.

- (i) Are the words known to the reader?

The familiarity of the words used in the text depends upon the reader's lexical knowledge (that is, knowledge about the meaning of individual words in the text). This is most important when single and unconnected words are being read. However, if connected text is written in such a way (and this will be

discussed in relation to other aspects of text analysis) that there is redundancy in the text then knowing the lexical meaning of each word is not essential. Only having lexical knowledge of each word is not sufficient to understand the overall meaning but it may be better than nothing. This becomes apparent if a word by word translation is made from a foreign language and it is easy to end up with lists of words which make no sense as a connected text. However in connected text some lexical knowledge is important because knowledge of none of the words would mean that the reader could not begin to analyse the text. Lexical knowledge then is a necessary but not a sufficient component of text analysis.

(ii) Are the words combined in an acceptable order?

The acceptability of the order of words or events depends upon the reader's syntactic knowledge. Syntactic knowledge at sentence level is knowledge about the rules which determine the order and combination of parts of speech. A competent reader is able to decide that sentences such as:

"The dog ate the bone" or "The bone ate the dog"

are grammatically correct, whereas:

"Dog the the ate bone"

is not.

The labels which have been attached to parts of speech by linguists (for example, noun and verb phrases, nouns, verbs and adjectives) are convenient as short-hand descriptions but are arbitrary and ill-defined when a deeper analysis of their functions is attempted. Yet sentence syntax has been described with the aid of these labels and they do serve to define a set of rules which attempts to distinguish an acceptable word order. What is being asserted at this point is that if the reader is to comprehend a sentence he must be able to decide whether the order is acceptable: no claims are being made about the rules that he uses.

Syntactic knowledge can be applied in isolation from lexical or other

knowledge. Esoteric scientific or sociological texts can be recognised as consisting of acceptable sentences even if the lexical meaning is obscure. But even if the lexical meaning is clear and the syntax (or grammatical structures) are acceptable, this is not a sufficient condition for the text to be comprehended. This is illustrated by the example used above ("The bone ate the dog"). Some additional analysis is required. This will be based on the appropriateness of the relationships between the concepts (see iii).

In addition to being able to decide upon the acceptability of word order, the reader must also be able to decide whether the structure of the complete text is acceptable. To make this decision he must have some knowledge of the conventions of text, so that he can recognise the form, for example whether it is a short story, a scientific report or a newspaper article and whether the content of the text can be understood according to these conventions.

(iii) Are the concepts related to each other in an appropriate manner?

The appropriateness of relationships between concepts depends upon the reader's semantic knowledge, that is abstract knowledge about words at their conceptual level. This includes abstract knowledge about the properties of concepts. For example "mother" is animate, human, female, adult, has children. "Mother" can only be linked to actions that require an animate relationship. For example, "breathes" requires an animate subject, the subject can be human or non human and what is breathed must be a gas/vapour. Therefore "mother breathes" is a semantically acceptable relationship, but "mother breathes water" is not (but see (iv) below). The example used earlier ("The bone eats the dog") is not acceptable in the light of conceptual and semantic knowledge because "bone " is inanimate, and "eats" requires an animate subject.

If the reader is reading a longer-than-sentence text, then this knowledge must extend to between sentence relationships if the text is to be comprehensible. There must be acceptable links between the sentences and these

must overlap if the text is to be comprehensible. The competent reader has to be able to discern these links and confirm their appropriateness if he/she is to decide that the text is comprehensible.

A series of sentences such as:

"Mother came into the room. She sat down by the table. It was covered with a clean cloth."

is well linked. But the reader needs to know about pronominalisation to be able to discern the links and to decide that the text is coherent. Even at this point the processing may not be sufficient for the reader to decide if the text is comprehensible. Further analysis is required in order to decide if the content is possible in terms of the reader's world knowledge.

(iv) Is the content acceptable, possible or plausible in the wider context of the reader's world knowledge?

The words in the text may be familiar in a lexical sense, the sentences and the overall text acceptable grammatically and according to typical text conventions, the relationships between the concepts permissible and the sentences within the text overlap conceptually, but the text may not be fully comprehended because the information which is given does not accord with the reader's knowledge about what is possible in the world. For example, "The lion slept on the table" is comprehensible in terms of the words, the grammatical structure, and the relationships between the concepts. However, lions live in jungles or safari parks or in cages at the zoo. Tables are kept in houses, or in restaurants and are generally used by people, so the sentence does not make sense in terms of general world knowledge. However, the reader may persist and try to make sense of the sentence believing that an author generally will wish the text to be sensible to the reader (this belief of course may not exist when the reader is in an experimental situation). At such times the reader may try to create a situation in which such information would be sensible. The source of the information can then be in the text, where he may discover

further references to the lion performing weird acts such as drinking milk from a saucer, eating fish and hunting mice and perhaps from this context decide either that this is a lion which has been trained to act like a domestic cat and is therefore unusual, or that the author is using the term lion metaphorically to refer to a domestic cat. Alternatively he may create a context from his own stored knowledge and either create or recall a situation in which this act might be plausible. It may well be an unusual and idiosyncratic situation such as that described above or a very general explanation such as the lion was receiving veterinary treatment and was lying on the table in the surgery.

The reader's world knowledge may enable him/her to create a context for sentences which are syntactically and semantically anomalous or ambiguous. In this sense, world knowledge is the most important factor in comprehension. For example, the sentence "Mother breathes water" was shown to be semantically unacceptable. However the reader is able to create either realistic or unreal situations in which the sentence can be acceptable. The realistic situation would accept the semantic anomaly (water is a liquid, not a gas, and therefore cannot be breathed) and create a situation in which this was maintained (for example, "Mother breathes water and dies"). The unreal situation would create a bizarre or fantastic situation in which breathing a liquid becomes acceptable:

"My mother, the mermaid, breathes water."

or: "Mrs Trout, the biggest fish in the river, is my mother and she breathes water."

A single sentence is easily ~~accommodated~~ accommodated in this way but if it is a part of a longer text this becomes increasingly difficult unless the situation created by the reader accords with that which has been chosen by the author. In reading longer texts, the reader would need to be able to change and reconstruct his/her interpretation.

Comprehension as a learning process.

A text which is lexically, syntactically, semantically and pragmatically (in terms of world knowledge) acceptable will offer no challenge to the reader because it can be passively absorbed into the existing knowledge structures. "Easy" comprehension would seem to imply that the reader need expend no effort in establishing meaning and that he/she will learn nothing from the text.

If something new is to be learned, the process must demand something more of the reader: the introduction of new terminology; a novel construction of the sentence or the text which perhaps will place the outcome of the sequence at the beginning as a flashback; a figurative use of language which may juxtapose semantically incompatible concepts and perhaps produce a metaphor such as "the leaves danced on the trees" or information which adds to or challenges the reader's knowledge of the world.

Whether or not such changes move the comprehension of text to the "impossible" end of the continuum depends on the knowledge of the reader. If it can be assumed that a skilled reader is aware of the linguistic conventions of text (that is, he has a good working vocabulary, can recognise non-grammatical constructions, is aware of semantic constraints) then comprehension will be a function of his experience and knowledge of the world.

The world knowledge that is required for the reader to check the reality of the information is idiosyncratic and personal and may well depend upon the his/her experience in life. Experiences differ at many levels. In very general terms a person who went to a church school may well know more about religious matters a person who did not. On the other hand those who have attended a Jewish School will have quite different religious knowledge than one who attended a Catholic School. In either case the extent of their knowledge will depend on their commitment to their faith. If they were given the same text about a religious festival in an educational test or in an experimental situation their comprehension would be very different because the world



knowledge which would be the basis of their criteria for judging what was possible would be different.

What is comprehended: the effect of purpose, instruction and experience on the interaction between reader and text.

(i) Experience.

The role of experience in text analysis has already been discussed and it is clear from the discussion that experience can strongly influence what is comprehended in terms of a passive acceptance of what is already known and known to be reality. Although common experience at a national level (for example, going through state education and being exposed to the the media) has provided for a great amount of overlap in certain areas of knowledge, experience at other levels (for example, professional and family influence ) can lead to wide differences in knowledge. Experience can have an additional influence in the part it plays in the choice and interpretation of self selected texts and can certainly influence the purpose for the reading the text (that is, in influencing what it is that the reader wishes to know when he begins to read a text). In experimental and educational situations differences in experience are controlled to a large extent by the selection of text and task.

(ii) Purpose and instruction.

Readers usually have a reason for reading a text. In a normal reading situation the purpose of the reader may be very specific (for example, reading the newspaper to find out what is on television or reading a journal or magazine article which has information about a professional or leisure time activity). On the other hand the purpose can be very general (to pass the time in the dentists waiting room, or to read for pleasure). Purposes may be related to roles in life such as spouse, parent, worker, or to leisure activities such

as gardening, knitting or computer programming and so on. What is comprehended (the outcome of the reading) is related to the purpose. The purpose which initiated the reading may not be constant or consistent throughout; information that is incidental to the initial purpose may change the purpose during reading and comprehension.

In an educational test or task, the choice of material is controlled by the teacher or by the text book or reading laboratory material. In an experimental situation the choice of material is controlled by the person who designs the experiment. Thus the reader's experience is incidental to the choice of material, unless it is a part of the overall plan of the chooser to include this experience. Therefore the overt purpose for reading a particular text is usually decided by someone other than the reader and the reader may or may not be aware of the reason for the choice of a particular text. Any personal and private outcome which results from reading the text is ignored unless by chance it corresponds with the outcome selected by the experimenter.

The reading outcome which is taken as a measure of comprehension is shaped and constrained by the instructions given at the outset, and the nature of the task the reader is asked to perform. The assessment of comprehension will be determined in part by the match between the reader's response and the expectations of the experimenter/ teacher.

The reader may be required to answer questions based on the text; to make truth judgements about the text; to recall the text or specific aspects of the text in cued or free recall conditions; or to recognise portions of the text. Time taken to read a text or a variety of texts may be measured. What is inferred from these measures is related to the theoretical stance of the experimenter.

So, in a controlled reading situation what the reader might be doing in addition to the task is largely ignored. It is easy to deduce from experimental results and test scores whether or not the reader is performing

the activity predicted by the task and /or the text. What it is not permissible to deduce is that what is being measured is necessarily "comprehension".

### A summary of the contents.

In their measures of comprehension, educationalists have largely been concerned with the content of what is comprehended, analysing the skills required for successful comprehension and the production of measures to assess the attainment of those skills. A very <sup>brief</sup> summary of their work in this area appears as Chapter One

Psychologists and Psycholinguists have been more concerned with the processes involved in comprehension and in providing empirical evidence for the existence of these processes. Their work is reviewed according to their theoretical perspective.

In Chapter Two the psycholinguistic models of Chomsky and Fillmore are reviewed. These models focus on the production and analysis of sentences. They are not primarily concerned with explaining what readers do but focus on the rules which will produce ideal sentences (those which are semantically and syntactically acceptable) and more importantly will not produce unacceptable sentences. In terms of the definition of comprehension which has been explored in this chapter, these models would produce sentences which (given a competent reader) are "easy" if the reader has the required knowledge of the world to comprehend their content.

Chapter Three reviews models of semantic memory. The models propose a set of universal abstract associations by which concepts in all sentences or propositions can be encoded. The models begin with an analysis of text according to experimenter defined relationships. They propose that the features and associations used in the analysis are paralleled in the reader. According to these models, the reader is a text analyser who processes the text in the same way as the model: that is, he /she is passive and encodes and

retrieves text in accordance with a pre-existing structure or an automatic process. The models are not concerned with ways in which the structures and processes might be established in the reader; they rest on the assumption that the structures and processes exist and that they are automatic and universal. Empirical investigations which support such models assume that the semantic processing described by that model is the main variable of reading time.

Models which have proposed global representations of knowledge are reviewed in Chapter Four. These models are similar to models of semantic memory except that they are concerned with the reader as an analyser of texts which are larger-than-proposition length. Generally ~~the~~ models are concerned with the consistent features of texts and situations and how they are represented in the reader (whereas the models in the Chapter Three were concerned with consistent features and associations of sentences or propositions) . Like ~~those~~ of semantic memory, the models of global knowledge assume that the proposed structures which can aid the analysis of texts and situations are present in the reader. Some models propose that the structures might be created by experience, others assume their existence. All assume an efficient analyser who encodes and retrieves large sections of text according to pre-existing structures. Some of the models are concerned with general knowledge structure, others with structures which will analyse a specific type of text. Neither these nor the semantic memory models are concerned with the purpose of the reader or with information which will not fit the pre-existing structures.

Chapter Five reviews the model of comprehension which has been developed by Kintsch and his associates. The early version of the model is concerned with modelling the process by which a text representation can be produced. The basic unit of the process is a proposition derived from the text by use of Fillmore's Case Grammar. The reader again is seen as a text analyser who passively and uncritically analyses all input from the text, his main purpose

being to establish textual coherence. Further development of the model describes the process by which the gist of a text is abstracted under the constraints of the reader's knowledge of text structure. A final development is the process by which the text is made "meaningful" to the reader in terms of stored structures of world knowledge. The model sees the reader as a processor who abstracts semantic, global and world knowledge from a text by embedding the content into a flexible structure which is constrained by stored information about the structure and content of texts.

Chapter Six examines the empirical evidence which supports Kintsch's model of comprehension. The methodology is again concerned with matching reading times and recall of text to a pre-existing model. The model is based on an analysis of the text and therefore the evidence compares the results obtained from the reader to properties in the text. It assumes that reading time is utilised in carrying out the proposed processes and proceeds on the basis that only recall material that matches the input is interesting.

Chapter Seven reviews the model of comprehension produced by Sanford and Garrod. Their model is concerned with the representation of the text as it might be created by the use of linguistic referents between concepts in the text. It models the process by which the reader creates and stores the text representation and how this relates (in terms of what is accessed and what is stored) to general long term memory. In this model the reader extracts the semantic relationships from the text by relating input to stored primitive meanings. Again the reader is seen as a processor of text who relates the text to a pre-existing structure to extract "meaning". According to this model all information accesses stored primitive structures. New structures are not created when a mismatch is made, instead a different structure is accessed.

In Chapter Eight the previously reviewed "association" models (which see the reader as a processor of text who passively relates all input to previously stored structures) are contrasted with models which present the

reader as an active constructor of "meaning". The chapter traces the development of the "constructivist" approach. This approach sees the reader as an active seeker of "meaning" who uses information in the text to create a "meaning" for the particular situation that the text or the context described. In this model the reader does not absorb input into pre-existing structures but creates structures from stored knowledge of the world which will make the content "meaningful" even when in strictly analytical terms it is not "easy" to comprehend. In the final part of the chapter the differences in approach of the "constructivist" and the "associationist" models are discussed by reference to models of elaborative processing and the experimental work which supports the two approaches.

Chapter Nine is a summary of the models presented in the review. It contrasts their assumptions for: the relationship between the structure and the content of text; the representation of text in memory and the processes used by the reader in the comprehension of the text and discusses the implications of those models for the comprehension of a specific extract of text.

Chapter Ten describes a pilot experiment which explores methods by which the processes involved in the comprehension of complex text might be examined. It particularly examines the type of question which is most useful for eliciting quantifiable data about the predictions and hypotheses made about narrative texts on the basis of varying amounts of the same text.

The methods which were established as being useful in the pilot experiment are utilised in an experiment described in Chapter Eleven. The experiment analyses the predictions made about the events and characters on the basis of various amounts of the same text and discusses the relative contributions of information about structure and content in the predictions that are made.

Chapter Twelve examines the differences in processing of the same

information when it is presented in canonical and "flashback" order. The chapter describes a series of experiments whereby the concept of "Interestingness" as described by Schank (1978) was investigated and used as a measure of the processing involved in the comprehension of two structures of text. It is concluded that variations in the structure influence the processing of the content to the extent that what is processed is influenced by the temporal order in which it is presented.

Chapter Thirteen examines the effects of text and reader variables on predictions by manipulating the text used in Experiments One and Two. The amount and type of information given to male and female subjects is systematically varied and the effects measured by the ratings of likeliness given to a set of predictions. It is concluded that readers use general stored world knowledge to comprehend narrative texts and that this is elicited by minimal information in the text.

The Conclusions present some general conclusions from the experimental work and their methodological and educational implications.

CHAPTER ONE.Educational perspectives and research into comprehension.

Educational research is concerned with comprehension as a part of the reading process. Two alternative perspectives have developed. According to one perspective comprehension forms the top level or levels of a hierarchy of skills through which the reader progresses either developmentally (in the process of learning to read) or sequentially (as an experienced reader). The other perspective sees comprehension as being central to reading with specific subskills being learned or practised as an aid to comprehension.

Comprehension: a part of a hierarchy of skills.

This perspective sees reading as involving a progression through a hierarchy of skills each level of which is a necessary prerequisite to the next level in the hierarchy. The hierarchy is devised by a logical or intuitive analysis of the reading process. According to this perspective, reading is a process which begins with the decoding of marks on paper (letters, phonemes or words) into speech sounds. Reading ability is developed by the gradual introduction of skills for decoding more vocabulary as single words and as sentences and finally by the introduction of skills for the extraction of meaning. Skills to work at these levels are developed by the use of structured texts. Methods of teaching which adopt this perspective (Gagne, 1965; Downing, 1972; Gattegno, 1962; Fries, 1962) assume that there is need for proficiency in certain skills at each level or stage of the process before the reader can progress to another stage. Comprehension is a part of the hierarchy and may have three or more levels or stages according to the various models. Gray (1960) for example proposes that comprehension has three levels: "the lines" (a clear grasp of what was read); "between the lines" (determining what is implied) and "beyond the lines" (implications beyond what is stated). Lunzer (1979) uses the analogy of a target with four concentric



rings, each ring representing a different skill. As the reader works outwards from the central skill of decoding words or marks on the paper, the levels of comprehension become less general and more idiosyncratic. The DES report (1975), Moyle (1972), Robinson (1972) and Barrett (1972) each propose that the comprehension process has five stages. The stages in the DES model and those proposed by Barrett are identical. They propose that the comprehension process has literal, reorganisational, inferential, evaluative, and appreciative stages (these stages are explained below). The DES report implies that these stages must be practised sequentially in each encounter with text: "When he has achieved a grasp of the literal content the reader is then in a position to analyse, paraphrase, synthesise and summarise it" (DES report, 1975, para 6:4). It is also implied that these are stages which can be separately taught and tested.

The stages in the Moyle and Robinson models involve going beyond the literal, interpretational and evaluational stages of the Barrett and the DES models to the integration of the information into what is already known and then next to 'acting upon' (Moyle) or 'Utilisation' (Robinson) of the text. The description of the five stages of the Barrett and the DES models is based on the descriptions in the DES report (Paras 6:40, 6:41.)

- (i) Literal: The reader identifies material which is explicitly set down in the text.
- (ii) Reorganisational: The reader can analyse, paraphrase, synthesise and summarise the material in the text.
- (iii) Inferential: The reader goes beyond what is explicitly stated in the text and interprets ideas and thoughts which might have been included or made explicit but were not.
- (iv) Evaluational: At this level the reader applies "truth" tests to the material to assess the logic or the authenticity, adequacy, and appropriateness in comparison with other sources.

(v) Appreciative: At this level the reader responds to the author's use of language, to imagery, to style and structure using aesthetic and affective judgements.

Barrett (1972) devised a taxonomy of thirty three tasks which relate to and are suitable for teaching and testing the five stages in the model. According to this perspective a successful reader is one who can apply lexical, syntactic and semantic skills to the analysis of text and as a result obtain the "meaning" from the text.

Comprehension: central to the reading process.

The alternative perspective (which has become known as the Psycholinguistic or Language Experience approach to reading) begins with the assumption that potential readers have had life experiences and that they possess lexical, syntactic and semantic knowledge so that these do not need to be taught as separate skills, instead they are applied to the understanding of the text. According to this perspective, reading is a unitary process which utilises prior experience and knowledge in the processing of text. When he/she is reading, the reader selects cues from the text which enable him/her to utilise world knowledge and recode the text into experience (either real or imagined). This contrasts with the hierarchy of skills model which sees reading as a process of decoding print into speech.

The psycholinguistic approach does not see learning to read as a steady progression through a standardised hierarchy of skills which lead from perception of print to sound and then to meaning. Instead it is seen as a process which differs for each individual as a function of prior experience and skill in applying knowledge of vocabulary, syntax and semantics to the text.

Teaching of reading (according to this perspective) is based on the exploitation of the skills and experience which the reader possesses and

begins with the encoding of the child's experience into a written form. The child then recodes the written word back into the experience. This is usually carried out through the medium of pictures or familiar objects (see Carrillo, 1967). From these early stages the child acquires a lexicon of written words which are associated with experience and meaning. New vocabulary is learned by applying "guess and check" strategies based on all available cues (Neisser, 1967; Goodman 1967, 1969, 1970).

Goodman (1969) has devised a method of analysing errors in children's reading which can be utilised when the child is reading aloud in a one-to-one situation. This method called "miscue analysis" assumes that errors are the result of applying knowledge inappropriately. Successful diagnosis based on the analysis of miscues allows the teacher to aid the child to develop better strategies for utilising the most appropriate cues and to recode the text successfully. In this perspective a successful reader is one who uses least cues to construct the "meaning" from the text in terms of his own knowledge and experience.

#### A comparison between educational and psychological perspectives on comprehension.

The two perspectives which have briefly been outlined above represent two very different approaches to the teaching of reading and two differing theoretical views of comprehension or understanding the meaning of text. The two theoretical positions (one that the "meaning" of a text can be obtained through an analysis of the printed word, and the other that "meaning" is constructed from the reader's knowledge and experience on the basis of cues in the text) are repeated in the psychological approach to comprehension. (The differences between the theories will be discussed in depth in Chapter Eight in relation to the role of elaborations in comprehension).

Statistical Analyses of the Skills of Comprehension.

Educational research into comprehension since the 1940's has been preoccupied with the question of whether comprehension is a single skill or a collection of subskills. Lennon (1962) reviewed some of the research which was carried out between 1940 and 1962. Each of the twelve pieces of research reviewed used factor analysis to provide a solution to the question. The studies in Lennon's review can be divided into three categories:

- (i) Those which carried out factor analysis on pre-existing data from established tests (Traxler, 1941; Langsam, 1941; Artley, 1942; Stoker and Cropp, 1960);
- (ii) Those in which results from new tests were combined with pre-existing data from established tests (Gans, 1940; Conant, 1942; Maney and Sochor, 1952) ;
- (iii) Those which analysed the data from newly created tests (Davis, 1941, 1944, 1946; Thurstone, 1946; Hall and Robinson, 1945; Harris, 1948 ;and Hunt, 1952).

Although the last category differs from the others in that the studies use new tests, the basis of these tests is a set of skills which were derived from a factor analysis of previous tests of comprehension.

Standardised tests of comprehension (for example, the Edinburgh reading test 1972, or the Neale Analysis of Reading, 1959) and Reading Laboratories (SRA, 1961; Ward Lock 1969, 1971) assess comprehension by means of questions that can be marked as right or wrong according to predetermined answers. The answers are derived from lexical, syntactic and semantic information which is explicit in the text or which requires a very general level of knowledge of the world. Such tests utilise at the most the first four stages of the DES model of comprehension and quite often (in the interest of standardised marking ) only the first two. In tests such as these idiosyncratic knowledge is not explored, and reasoning which is based on information which contradicts

that stated in the text is scored as if it were incorrect. The outcome is that in these tests the definition of comprehension is narrowed and restricted in the interests of establishing the reliability of the tests and the standardisation of marking.

The statistical research (reviewed by Lennon, 1962) which explored the question of whether comprehension is a process involving a unitary skill or a complexity of multiple skills was based upon definitions of comprehension derived from similar tests. Davis (1941) surveyed and analysed the questions which were used in comprehension tests and the skills that they claimed to test. He identified several hundred separate skills. These were sorted into nine groups which had high inter-group correlations and low between group correlations. He developed test questions to measure these nine basic skills and (Davis 1944) administered the tests and computed the intercorrelations among the results. He used a multi-dimensional method of factor analysis and weighted the questions according to the importance of the skill (judged by reading experts) prior to the analysis. From his interpretation of the factor analysis, he was able to identify nine factors, six of them statistically significant. He was able to conclude that two of the factors (word knowledge and reasoning ability) were measured with sufficient reliability to be used as they were for teaching and testing comprehension. Three other factors (literal comprehension, inference and the ability to follow the organisation of a passage) needed further development.

Thurstone (1946) reanalysed Davis's (1944) data using a unidimensional method of analysis. He found that one common factor accounted for most of the variance in six of the skills and that there was high single factor loading with a very small residual. Hence he concluded that there was only one skill underlying the comprehension process (General Reading Ability) and that there was no evidence for the existence of components of this ability.

Davis (1946) reiterated the multi-skill perspective and claimed that

comprehension involves five separate and identifiable mental abilities:

- (i) Word Knowledge;
- (ii) Ability to reason (identifying the relationship between ideas);
- (iii) Ability to follow the organisation of the text;
- (iv) Recognition of literary devices;
- (v) Ability to focus on explicit information in the text.

In addition, Davis pointed out the three basic differences in approach that were underlined by the separate analyses carried out by himself and Thurstone.

1. Thurstone's analysis was of the common variance between the factors, whereas Davis analysed the total variance. Thurstone was interested in finding significant differences in the specific variance between factors while Davis wished to analyse any differences that remained over and above the general reading ability.

2. As a consequence, Thurstone used a centroid method and Davis used a principal axes method of factor analysis.

3. Thurstone weighted the tests in terms of their communalities while Davis weighted his in terms of their importance according to the judgements of reading experts.

These differences in approach were guided by different goals and philosophies. Davis was concerned with results which would enable him to find differences between the skills so that he could devise tests and instructional methods for the teaching and testing of comprehension. Thurstone was concerned with the method of analysis employed to analyse the given data with no practical application in view.

Lennon (1962) reviewed these and the other studies cited and concluded on an intuitive basis that a multiplicity of subskills underlies comprehension in reading. Lennon's review is useful to the study of analysis of comprehension skills mainly because it underlines the intuitive basis of definitions of

comprehension skills, the variety of methods that have been used to test the assertion that comprehension is a multiplicity of subskills and the inconclusive nature of the statistical evidence supplied.

Davis (1968) undertook a further study of the single/multi-skill question. He selected eight skills which he derived from reviews of earlier experimental studies of reading comprehension. He constructed test items to measure each of these skills and a separate passage of texts to measure each of these items. The test items were divided into two equivalent forms and the same subjects were tested with one or two days interval between each form of the test. A 'uniqueness-analysis' multiple regression technique was used to analyse the combined data and to determine the proportion of variance of a particular skill which can be accounted for by the best weighted combination of the other seven skills. From the results he concluded that comprehension is not a single mental ability but that the mental abilities used in the eight skills are independent of each other. He claimed that five of the skills could be experimentally differentiated. These were:

- (i) Recalling word meanings;
- (ii) Finding answers to explicit questions;
- (iii) Drawing inferences from the content;
- (iv) Recognising the writer's mood;
- (v) Following the structure of the passage.

Davis (1971) subsequently analysed the same data using a principal components analysis after rotation to the normalised varimax criteria. The analysis was applied separately to the two subtest forms. In each of the matrices the same four factors were identified. These were equivalent to skills 1, 2, and 3 of the initial analysis and in addition the factor identified as the ability to infer the meaning of a word from the context. Other factors were found separately in each matrix.

Thorndike (1971) reanalysed Davis's (1968) data using a principal

components analysis of the reliability co-efficients in the diagonal cells of the matrix. In this way he demonstrated that the non-chance variance in the eight tests could be accounted for completely by three factors and that one single factor accounted for 93% of the non-chance variance. Only one skill (word knowledge) could be accounted for in terms of its factor pattern. No one other of the eight skills could be distinguished separately.

The differences between the methods of analysis used by Davis (1968, 1971) and Thorndike (1971) are similar to those between Davis (1944, 1946) and Thurstone (1946). In both cases Davis has based his analysis on the total variance and his critic has based his analysis on the common variance.

Spearritt (1972) analysed the Davis (1971) data yet again, in an attempt to establish the differentiability of reading skills. His method of analysis differed from those used both by Davis (1968, 1971) and Thorndike (1971) in that he combined the two eight-variable matrices that they used in their analyses and he produced a sixteen variable matrix based on the sixteen scores of each subject. The data points were then subdivided into Sample One and Sample Two and two matrices produced. Spearritt then used a maximum likelihood factor analysis procedure (Joreskog and van Thillo, 1971) to analyse the data. Davis's (1968) hypothesis of eight separate reading skills was used to specify the theoretical factor model. The procedure placed the unique variance from each skill into the common factor space in the model, and the variance was allowed to emerge as a factor if the factorial solution was acceptable at the required level of statistical significance. The results for Sample One and Sample Two varied: Sample One needed three factors to account for the correlations between skills and factors, Sample Two needed four factors.

Spearritt (1972) used the results from this exploratory analysis to create and test a further hypothetical structure. To do this he took from the sample those skills which had the highest loadings on the factors identified in the exploratory analysis. He analysed these results using a restricted



maximum likelihood test. From this analysis, four skills were identified which could be differentiated in terms of maximum likelihood factor analysis. These four skills were similar to those identified by Davis (1971). An interesting conclusion from Spearritt was that although four skills can be identified as a part of the comprehension process, most current reading tests tend to measure only one ability apart from word knowledge and that may well be "reasoning in reading".

It would seem from the analyses cited that the conclusions arrived at from research which relies on the analysis of the same data depends upon the method which is used in the analysis. The method employed in turn relies upon the theoretical stance of the researcher and the aims of the research. However recent work which was carried out by Lunzer, Waite and Dolan (1979) has shown that this is not necessarily the case. Lunzer et al began their research to establish that a multiplicity of skills could be identified as a part of the reading process. They constructed their own test items with the dual aim of demonstrating that subskills did exist, and investigating whether methods of instruction could be devised to promote the teaching of these skills.

In their test eight subskills were assessed on each of four different texts. The subskills were derived from the New York List (Rauch, 1972) although this was modified in such a way that some of the items on the list were combined to form a single subskill. The test was administered and the results analysed using the Kaiser (1970) method of factor analysis. This test assumes that the factors are oblique and therefore correlated.

The results of the test were inconsistent because although six factors were defined, they were not present in the data from all of the texts. What was found was that two of the factors related to two of the texts and two other factors related to the other two texts. Only the first general factor, which accounted for 31% of the variance, related to all four texts. Further analysis of the data (in which the results from all of the texts were combined

and subskill scores were averaged out across the texts) confirmed the previous conclusion that only one general factor existed. From this Lunzer et al have concluded that comprehension is a unitary ability, that is the ability to reflect upon what one is reading and that any one of the test items will allow the reader to do this.

The conclusions of Lunzer and his associates are interesting for two reasons:

1. They have demonstrated that even given a particular stance in the question of skills and subskills, and a method of analysis which seeks to support this position, it is not always possible to substantiate the stated hypothesis. (The previous research which has been discussed in this chapter had seemed to force the conclusion that this was possible.) From this and the other results it is now possible to conclude that more is required of the research than that a 'suitable' method of factor analysis is applied to the results. Some more rigorous theoretical base for the description and assessment of skills will be required if the question is to be pursued further.

2. The main point about the change in their theoretical position is not merely that comprehension is a unitary skill, but that it is a mental ability which can be encouraged by any method which requires that the reader reflects upon what is read. This conclusion would seem to change the nature of the question being asked about the comprehension process: instead of asking 'how can comprehension be taught?' as if it were a set of skills that once learned could be transferred from one text to another, the question now being asked is 'how can the reader best be encouraged to reflect upon this particular text?'. From this perspective comprehension would seem to be a reader centred process in which his/her involvement, purpose and own knowledge and perspective are of more importance than the application of a standardised method of analysis to the text.

Summary of the educational perspectives and research.

On the surface it appeared that the two educational perspectives presented at the beginning of the chapter were diametrically opposed. However a more careful examination of the differences has revealed that although there are differences between them, there is also some common ground.

Theories which maintain that comprehension is a unitary process would seem to be saying that comprehension is a general process of understanding things and is not necessarily tied to reading. They would argue that comprehension is related to the understanding of visual information as well as of spoken and written language. In those who are beginning to read this general comprehension is applied to the written text. Even so, without reading skills written text can not be comprehended because it is impossible to translate the unknown in terms of one's own experience. Just as a blind person must learn to "read" Braille, so the beginning reader must learn the conventions of the medium of print. Such things as word shapes, spelling patterns, punctuation and permissible combinations of words must be learned. Those who would maintain that comprehension is the upper layer or layers of a hierarchy of skills would also maintain that the skills were necessary, furthermore they would argue that mastery of the skills is essential for comprehension of the text; that comprehension is a matter of decoding the meaning from the text. In this view comprehension of text is separate from general comprehension and meaning can only be derived from print when the reader can use the decoding skills lower in the hierarchy.

Both would agree that skills are necessary for the reader to be able to interact with the text. In this respect the two perspectives overlap. Where they have differed is in their approach to how the skills should be taught (whether the reader is taught to decode the print into sound and then into meaning, or whether he/she is taught to apply his experience and knowledge to recode the print in a deductive manner); and more pertinently (to this thesis)

how the relationship between text, comprehension and the reader can be defined. Their differences in this area are: the unitary-skill view is that the reader brings comprehension to the text; the hierarchy-of-skills view is that the reader takes the meaning from the text.

In the face of these similarities and differences the factor analysis of skills is almost totally irrelevant, for the difference centres on philosophy and methods of teaching and not on the question of the existence of separate skills. How far these skills can foster comprehension ultimately depends on the definition of comprehension adopted, that is whether it is seen as an integral part of a general concept of comprehension or a separate text based process. The definition adopted is influential in decisions which are made at many levels in the choice of tests, texts and methods of teaching.

In a recent report to the Schools Council, Gardener (1982) changed the focus of the comprehension debate. Instead of continuing the long standing debate of whether it is a single or a multi-skill process, he has now adopted the Lunzer et al (1979) view that comprehension is a general ability to reflect upon the written word. The new question is whether comprehension of a text is sufficient to ensure that learning takes place.

If comprehension is separated from learning in this way, then the implication must be that it is a passive process in which those aspects of text which are already a part of the knowledge base are absorbed. From this it follows that learning will be the active process in which analytic, integrative and inferential operations are carried out so that the reader can learn something new from the text.

This shift of focus seems to have put the "skills" problem onto a more manageable and operational footing because the questions asked of the text can now be related to a particular problem or purpose. But in so doing the definition of comprehension has been diminished and restricted to a process by which text is passively related to what the reader already knows.

CHAPTER TWOLinguistic models which have practical and theoretical implications for models and theories of comprehension.

Educational theories of comprehension have focussed on the relationship between the reader and the text. The issue of whether comprehension of texts is a specific process or an integral part of a more general concept of comprehension and the related issue of whether it is brought to or taken from the text were central to educational theories which were discussed in the previous chapter. The linguistic theories which are to be discussed here are united on these issues. Their perspective is that meaning is taken from and determined by the text. If this view was to be extended it would seem that comprehension is solely a text bound issue and meaning is accessed from any specific text by means of analysis of the semantic and syntactic structure. This would mean that the reader's contribution would be knowledge of the word meanings and of the semantic and syntactic rules which constrain the way in which they can be combined.

However the linguistic theories which are to be reviewed do not focus on the reader/listener and his part in the process, nor particularly on the text itself but on the underlying rules which would determine the generation and comprehension of each and every sentence or proposition in a language. The theories are concerned with spoken language but they assume implicitly that written language is no more than a transcription of spoken language and that the two forms are interchangeable. This aspect parallels the hierarchy-of-skills perspective which was described in the previous chapter and which sees reading primarily as a decoding of the written to the spoken.

The concern of the linguist is not to provide an explanation of any specific example of spoken language, but to provide rules which are generally applicable to the production and comprehension of all grammatical sentences or

propositions. Texts which are used in these theories and models are merely illustrations for the display of the underlying rules.

### The theories of Chomsky (1957, 1965) and Fillmore (1968)

Two particular linguistic theories are reviewed. These are Chomsky's two models (1957, 1965) which are mainly concerned with syntactic structures and Fillmore's (1968) model of case grammar which is concerned with semantic rules. Both these theories have been important in the development of models of text comprehension. Chomsky's models were important at three levels:

1. They provided a model from which psychological studies of the production and comprehension of continuous text could be developed. (Prior to Chomsky experimental psychology had been almost exclusively concerned with the perception of and memory for lists of single words).
2. They were the source of specific hypotheses about the nature of the processes which underlie the comprehension of text.
3. The theoretical ideas about deep and surface structure have been influential in consequent models of text processing. (See Schank and Ableson's (1975) theory of primitives [Chapter Four] and Sanford and Garrod's 1981 Scenarios [Chapter Seven]).

Fillmore's (1968) model of case grammar has become central to subsequent models of text analysis. (See the 1972 Rumelhart, Lindsay and Norman model of semantic networks [Chapter Three] and Kintsch's 1974 model of text coherence [Chapter Five])

### Chomsky's (1957, 1965) models of generative transformational grammar.

Chomsky's models attempted to find rules which could create all grammatical and no ungrammatical sentences in a language. The models were theoretical and provided a description of the way in which an infinite number of sentences could be created from a finite number of words. In this respect

they were generative. They were also concerned to provide rules which would be able to combine the same word string in all its possible forms; active/passive or interrogative/declarative, and would be able to combine simple sentences into complex sentences. In this respect the rules were transformational.

### 1. The 1957 model.

This model was concerned with describing how sentences were produced. The model had four components:

- (i) An initial element;
- (ii) A string;
- (iii) The kernel or transformed sentence;
- (iv) A final form of a grammatically acceptable sentence.

The initial element was the starting point for the generation of the sentence. This would seem to be merely the notion that a sentence was to be created, or a signal for the process to begin. This was represented symbolically (in the model) as "S".

Phrase structure rules operated on the sentence idea and systematically (by a process of binary division) broke it down into its constituent parts (for example, the initial division was into noun phrase and verb phrase) until the terminal stage was reached. At this point the nouns and verbs could be replaced by words. Because this was a theoretical and abstract model it was possible to substitute any nouns and verbs at this point. Therefore any combination could illustrate the model. What has been generated at this stage is the syntactic or grammatical structure. This describes permissible noun verb combinations in a simple (single verb) sentence. The construction at this stage is called the string.

Transformational rules operated on the string and were of three kinds:

a. Obligatory rules.

These were applied to all strings and supplied compulsory syntactic features (such as subject-verb agreement) to the string. These obligatory rules produced the kernel sentence or the simple sentence in its active declarative form.

b. Optional rules.

These rules operated on the kernel sentence to produce a negative/passive/interrogative form (or a combination of these forms) of sentence.

c. Generalised rules.

These allowed the strings which underlie two or more kernel sentences to be joined either by embedding (the cat, whose coat was silky, drank the cream) or by conjoining (the cat ate the fish and drank the cream).

At this stage the model had produced the kernel or transformed sentence (that is the words plus the symbols which indicate mood, form, tense and noun/verb agreements). The morphophonemic rules operated on the sentence at this stage and produce the words in their appropriate form for speech or written presentation.

Investigations arising from the 1957 model.

Although Chomsky produced a set of linguistic rules, the models were adopted by psychologists (who rapidly became known as Psycholinguists) as models of human processes in the production and comprehension of language. The 1957 model gave rise to a number of hypotheses which tested a general notion that the greater the complexity of the rules which are necessary to produce the sentence, the longer it will take to produce or comprehend a sentence and the more storage space the process will occupy.

If these hypotheses were supported then it could be predicted that:

- a. Kernel sentences would take least processing time and occupy least processing space.



- b. Sentences which require one optional transform (either negative, passive or interrogative ) would take longer and require more processing space than a kernel sentence (but there should be no difference in processing time between the options).
- c. Sentences which require more than one option will take even longer to process and require more processing space (the increase being a function of the number of optional rules which must be applied).
- d. Recursive sentences should take the longest time and require more processing space as a function of the number of options and the number of strings involved.

Experiments by Miller and McKean (1964) and Savin and Perchonock (1965) illustrate the kinds of experiment that were carried out to test these aspects of the model.

Miller and McKean (1964), using reading time as evidence, found support for the hypothesis that reading time increased as a function of the complexity of the sentences which were produced and comprehended. Their results suggested that transformations were individual operations that were carried out independently and serially during encoding and decoding. They did find however that passive transformations took more time than negative transformations.

Savin and Perchonock (1965) used a recall paradigm and measured the number of extra unrelated words that could be recalled when the transformations were being carried out. They found that there was a decrease in the number of non-related words that could be recalled as the number of transformations increased. This suggested that the transformations were independent processes that required individual amounts of processing space and that increasing the number of transformations encroached upon the storage space available. These results however found that negative transformations required more processing space than passive transformations. There was therefore a basic contradiction in the results and the experiments failed to

establish a one to one relationship between the complexity of the transformations and the performance of the subjects.

The 1957 model was a formal and abstract description of rules which would produce grammatical sentences and conversely could be used to analyse any grammatical sentence. Although experiments tested the model for its psychological reality, it did not claim to model the actual process used by a speaker of the language. In addition the rules were very general and were not intended to generate or analyse any particular utterance. The rules however could only be illustrated by using examples of actual language and the use of actual language produced the problem of meaning that the non syntactic (or semantic ) aspects of words which made them sensible or nonsensical in combination. The idea of producing a user element was not an afterthought but a part of the initial plan;

"Having determined the syntactic structures of the language, we can study the way in which this syntactic structure is put to use in the actual functioning of language". (Chomsky, 1957.p.101.)

#### Chomsky's 1965 model.

The 1965 model introduced the concept of meaning into the abstract structure. A direct comparison between the 1957 and 1965 models is confusing because Chomsky uses the same terminology (for example transformational rules) for different procedures in the two models. Conversely he uses different terminology for the same processes (for example, phrase structure rules and transformational rules in the 1957 model become incorporated as base rules in the 1965 model). However some comparison is necessary to explain the terminology and to point out the developments in the 1965 model. To minimise confusion the 1965 model will be discussed initially as though it were a new model and comparisons will only be made as is necessary.

The 1965 model had three components: a syntactic component (which created

the sentence structures); a semantic component (which created the meaning ) and a phonological component (which provided the final spoken form of the sentence) .

The processes operated in the following way. The syntactic component was the first and the fundamental element of the process. At this level the following procedures were carried out. An initial element (or sentence idea ) was the starting point. Base rules operated on this and generated the string or strings of the simple or complex (embedded or conjoined) sentence. Additional obligatory rules determined the mood (active or passive) and the form (declarative or interrogative). These rules created the deep structure of the sentence which was symbolically represented as a tree structure which integrated the embedded or conjoined sentences and contained information about the mood and form of the sentence. Lexical selection rules (a part of the base rules ) provided the appropriate words to the terminal string. These operations provided the deep structure of the sentence. Transformational rules changed the sequencing of the words in the sentence to the order which would be used in the final spoken form. In addition to the correct sequence of words the structure included the markers for tense and agreement. This was the surface structure of the sentence. The two structures (deep and surface) produced by the syntactic component became the input for the semantic and phonological components:

- a. The deep structure became the input for the semantic component. At this stage semantic interpretation rules extracted the meaning from the structure.
- b. The surface structure became the input for the phonological component. At this stage phonological rules removed the markers and presented the words of the sentence in a final form for a spoken presentation.

Chomsky related the syntactic component to human structures:

"There are two aspects to this syntactic structure. It consists of a surface directly related to the phonetic form, and a deep structure that underlies the semantic interpretation. The deep structure is represented in the mind but is rarely indicated in the physical signal."(Chomsky,1965).

The syntactic component also included a lexicon and the base rules included rules by which individual words were selected from the lexicon (to replace the symbols in the formation of the deep structure). The rules had to provide restrictions upon selections so that meaningful and non-deviant sentences would be produced. Chomsky proposed two ways in which selections might be restricted: one related to restrictions at the syntactic level, the other related to restrictions placed on words in the lexicon:

1. Restrictions which are imposed by the syntactic structure. Selection restrictions at a syntactic level could be imposed by the addition of "strict subcategorisation features". This would entail marking the syntactic elements in the lexicon with their restriction. For example verbs could be labelled transitive with the information that transitive verbs only appear in sentences with a direct object. This would mean that major word classes would be divided into sub classes according to the syntactic frames in which they could occur.

2. Selections which are restricted by the other words in the string. Selections which operate in this way were made according to the semantic features of individual words in the lexicon (that is features that exist as a part of the meaning of the word and which determined the combinations of words which were permissible in the string).

One approach to this problem was proposed by Katz and Fodor (1963). They proposed that the features would be represented in a hierarchical form. Chomsky argued that there was no basis for a hierarchical structure because no

one feature was of any more importance than any other, but that the features could be represented as sets. For example, nouns might be represented in the lexicon with a set of semantic features in the following (simplified) way:

John is animate; is human; is male.

Spider is animate; is non-human; is male or female.

Leaf is inanimate; is non-human; is asexual.

A verb may be represented as follows:

Frighten: takes subject which is animate/inanimate; takes object which is animate only.

From this lexicon:

"John frightened the spider" is permissible;

"John frightened the leaf " is not;

and both:

"The spider frightened John "

and "The leaf frightened John" are permissible.

The rules which govern the selection from the lexicon provided information to the semantic component for the operation of the semantic interpretation rules. However "meaning" was still restricted to the 'possible' components of words or words which had properties which would allow them to be combined sensibly. The probability of sentences being "meaningful" in terms of the reader/speaker's experience and knowledge of the world was no part of the model.

The 1965 model differed from the 1957 model because it began to consider problems of speech production that were related to problems other than those of syntax. This led to the creation of a model which contained a semantic component and added the dimension of meaning to the model. Although the model was still concerned with the production of general and abstracted rules the introduction of meaning meant that the abstraction must be at a different level and must allow for some of the properties of spoken language (such as

sensible combinations of grammatical utterances). Chomsky related this model to a restricted aspect of the user's language. That is the model still did not attempt to produce or analyse specific utterances, but instead created a system of rules which would account for the speaker's competence, or his/her potential (or underlying knowledge) for the production of language. Thus the focus of the model remained upon the underlying rules of the language and avoided description of the contribution of the speaker.

#### Hypotheses generated by the 1965 model.

A thorough and detailed review of the psycholinguistic experiments which related to the Chomskian models is presented in Greene (1972) and Lyons (1970). A brief description of some of those experiments which relate to the interaction of the reader/listener with the text (spoken or written) is all that will be presented here.

Some of the experiments and models which followed the 1965 model were concerned with the investigation of the reality of deep structure and the transformational rules by which deep structure became surface structure. The model took no account of the knowledge and experience of the reader/speaker. This allowed for clarity and generality in the model (as has been discussed in relation to lexical selection restrictions). However experiments and models which explored the validity of the model as a psychological process could not ignore these attributes of the human processor.

Katz and Postal (1964) produced a model which anticipated that produced by Chomsky (1965). They attempted to model speech production which relied on rules to transform the deep structure to the surface structure. They found that the processes could not be carried out without some component which explained how the system (which could generate an infinite number of optional sentences) could generate a sentence which expressed the speaker's semantic intentions. In other words the model needed to show how a particular sentence which was apt for the situation could be produced. After all a speaker does

not produce sentences at random and even if his speech production process accords with that stated by Chomsky's model he needs at some point to be able to select words which are relevant and not only those which make permissible and sensible sentences. In short it would seem that a competence model can not be tested because it is too general to model the processes used by individual subjects.

Fodor and Garrett (1967) showed that in comprehension of complex sentences readers used their knowledge and experience of language rather than transformational rules to interpret complex sentences. They failed to find a one-to-one relationship between the number of transformations required and the time taken to read and understand complex sentences. The addition of pronouns or adjectives (which according to the theory should have required extra transformations and therefore extra reading times) actually decreased the time taken to read a complex sentence. For example the sentence "The shot the soldier the mosquito bit fired missed" took more reading time than either "The last shot the tired soldier the mosquito bit fired missed" or "The shot which the soldier that the mosquito bit fired missed". This experiment demonstrated that transformational rules do not adequately account for human text processing. Fodor and Garrett maintained that the difference was due to perceptual mapping processes and that these were of more importance than transformational rules.

Other research showed that the mood or form of a sentence was not memorable. According to the 1965 model this information was contained and stored in the deep structure and was input to the semantic component.

Sachs (1967) showed that subjects who were given recognition tasks could not remember whether the original form was active or passive. Johnson-Laird and Stevenson (1970) showed that sentences which had very different constructions but similar meanings were confused in a recognition task (for example sentences which indicated a buying or selling transaction such as

"Mary bought the car from the garage/ The garage sold the car to Mary").

These experiments would suggest that the syntactic deep structure is not retained in memory, but that a semantic extraction is retained. The work of Herriot (1969) adds strength to this conclusion. He showed that sentences which were semantically improbable required a longer response time than sentences which were semantically probable. For example,

"The lifeguard was rescued by the bather "

required a longer response time than:

"The comedian amused the audience".

The effect was greater than the differences in response time between active and passive versions of the sentences. The task in the Herriot experiment was to indicate the subject and object of the sentence. Chomsky's (1965) model would indicate that this is a decision that is made at the deep structure level. Fillmore (1968) whose work is to be discussed in the next section proposed that the subject/object decision was a surface distinction and that the deep structure had a semantic base in which subject/object was a non-useful distinction.

#### Fillmore's (1968) Case Grammar.

Fillmore (1968) formulated a set of linguistic rules in which the semantic structure was of prime importance. This was his theory of Case Grammar. By "case " he did not refer to the variations of endings attached to nouns which is the traditional usage in Greek, Latin and modern German languages. Rather he referred to a set of relations (which he proposes are universal and maybe innate) which identify the kinds of judgements that human beings are capable of making about events that are going on around them (who did it, who it happened to, what got changed). According to this model the cases could be represented as the relationships between the verb and the noun phrases in the simple, single verb sentence.



The cases which he proposed initially could be considered as a working draft rather than a complete and comprehensive list. He acknowledged that additions would be made to the list as necessary. The cases that he first listed were:

Agentive (A)- "the case of the typically animate perceived instigator of the action identified by the verb."

Instrumental (I) - "the case of the inanimate force or object causally involved in the action or state identified by the verb."

Dative (D) - "the case of the animate being affected by the state or action identified by the verb".

Factitive (F) - "the case of the object or being resulting from the action or state identified by the verb, or understood as a part of the meaning of the verb".

Locative (L) - "the case which identifies the location or spatial orientation of the state or action identified by the verb".

Objective (O) - " the semantically most neutral case, the case of anything representable by a noun whose role in the action or state identified by the verb is identified by the semantic interpretation of the verb itself; conceivably the concept should be limited to things which are affected by the action or state identified by the verb."

(Definitions quoted from Fillmore, 1968, pp 24-25)

The "objective" case as used by Fillmore differs from the notion of direct object as used in traditional grammatical analysis. The objective case (like all cases ) is a constant relationship regardless of the surface form of the sentence, whereas the traditional meaning of "object" depends on the position of the noun in the sentence. For example in Case Grammar analysis both "John broke the window" and "The window was broken by John" can be

represented in the same way:

John - Agent

break - verb

window - objective.

but in the traditional analysis they would have two different representations.

"John broke the window":

John - subject

broke - verb

window -object

and "The window was broken by John":

The window - subject

was broken - verb

by John - ~~adverbial~~ adverbial phrase

Case relationships then remain stable despite changes to the surface form. Fillmore also proposed that case relationships or case structure should be tenseless so that the structure is stable and general for all versions of the stated relationship.

In the model the semantic structure was created by the case grammar rules and was central to the meaning of the sentence. So, in this model the case grammar became the deep structure of language production and comprehension.

The absence of tense form and mood from the grammar required two assumptions:

1. A sentence has two components; a modality component (which provides information about the tense, mood, agreement, form and the syntactic agreements between nouns and verbs) and the proposition or case grammar.
2. The case grammar will relate the noun phrases to the central verb in a heterarchical fashion. This means that there is no implication of sequence at this stage. The grammar creates the deep structure which is a set of

relationships between the verb and the noun phrases. This set is stored and it is invariable regardless of the surface structure of the sentence.

The model would seem to work through the following stages in the production of language (and in reverse in the comprehension process):

1. There is an initial element or the idea that a sentence is to be formed.
2. There is a semantic component in which the deep structure is created from the case grammar and the lexicon.
3. There is a transformational element which creates the sequence and syntax for the surface structure.
4. Finally there is the surface structure or presented form of the sentence.

### 1. The Initial Element.

At this stage the notion that a sentence (S) is to be created and that there are two main components, the modality (M) and the proposition (P) is formalised. It is represented symbolically as :

$$S \rightarrow M + P$$

### 2. The Semantic Component.

The semantic component creates the propositional element from the case grammar and the lexicon. It was represented symbolically as:

$$P \rightarrow V + N1 + N2 + N3 \dots \dots \dots + Nn.$$

Certain rules for the creation of the structure were very clear. For example:

- A. A basic proposition has one verb and one or more noun phrases associated with it in a particular case relationship.
- B. Each relationship only occurs once in each proposition; if the same case occurs more than once then a complex sentence (that is a sentence with more than one proposition) will be formed.
- C. Only the same cases can be conjoined. This rule disallows sentences such as : "John and the hammer broke the window".

However the underlying case relationships (hammer - Instrument and John - Agent) allows each of the cases to be independently associated with the verb and the sentence can be presented as : "John broke the window with the hammer."

D. The agent must be animate. This would prevent the formation of a sentence such as : "The hammer broke the window with a chisel."

In the lexicon both verbs and nouns were marked with restrictions that related to the case structure. Verbs could be marked in relation to the cases with which they could be associated. For example the verb "open" might be marked in the following way: "open" must have Objective, can have Agent, can have Instrument.

Nouns need to be marked in two ways; in relation to the case structure (for example: hammer - can be Instrument, cannot be Agent) and with semantic qualities (for example: dog - is animate, can be agent).

### 3. The Transformational Component.

The transformational component of the sentence deals mainly with modality of the sentence. At this stage there are mechanisms which deal with the transformation of the deep (semantic) structure of the propositions to the surface (presented ) structure of the sentence.

This stage adds the surface structure markers such as mood, tense, agreement and sequence to the deep structure.

### 4. Surface Structure

This is the final stage or the presented form of the sentences which has emerged from the transformational stage.

A Comparison of the Chomskian and Fillmore Models.

Chomsky (1957, 1965) and Fillmore (1968) both attempted to formulate a set of universal linguistic rules which would generate and analyse sentences. Both theories assumed that language had a deep and a surface structure and that the deep structure was the underlying base for the surface structure. Beyond these similarities, their models differed on many levels.

One difference was the universality of the models produced. Chomsky's (1965) model would produce grammatically and semantically acceptable sentences in the English language. It is presumed that this was to be an initial step in the creation of a more generally applicable model. Fillmore's (1968) model was intended to be more universally applicable in the form in which it existed. Although Fillmore illustrated his model with examples from the English language, he demonstrated that it is also applicable in German and suggested that it might be applicable to such diverse languages as American Indian Languages (Paiote, Sioux, and others), Phillipino, Japanese and others.

A further difference was in the relative emphasis which was placed on the syntactic and semantic components of the sentence. In both versions of Chomsky's models, the syntactic component was central to the structure of the sentence. In the 1965 version, the semantic component existed only in relation to the syntax and the lexicon (both of which were a part of the syntactic component). In Fillmore's model the semantic component was central and it created the deep structure of the proposition.

A third difference between the Chomsky and Fillmore models was the importance of the features of mood, form and tense. In the early (1957) Chomsky model these features were added at a stage between the deep structure and the finally presented version of the sentence. In the later model (1965) they were an integral part of the deep structure. In Fillmore's model the "modalities" were peripheral to the proposition or deep structural meaning element and only of importance in the surface structure or the presented

version of the sentence. The case grammar which created the meaning and structure of the sentence in the form of a proposition focussed on the relationships between the verb and the noun phrases regardless of tense and mood of the sentence.

A further difference was in the shape of the deep structure: in both his models Chomsky presented the deep structure as a hierarchical structure in which a series of binary divisions resulted in the production of a string of logically joined words which were then transformed into a surface sequence. The deep structure had the function of resolving any ambiguities at the surface level and the disambiguation process related to the structure of the hierarchy. In Fillmore's model the deep structure was heterarchical and merely indicated the relationship between the verb and the noun phrases, no one case was of more importance than any other at this stage. Importance of a noun phrase in terms of the final (surface) sentence was decided in the transformation stage with ambiguities presumably being resolved in terms of the case relationships between the nouns.

#### The limitations of the linguistic models as models of comprehension.

Both the Chomsky (1965) and the Fillmore (1968) models produced useful theories of text analysis and could generate or analyse sentences which were grammatical and meaningful as isolated single sentences. However their completeness and self sufficiency prevented them from being useful in the analysis and production of larger-than-sentence pieces of text, because their procedures just were not designed to deal with inter-sentential meanings. References which crossed sentences boundaries would only be dealt with as separate propositions in the case of Fillmore and as separate sentences by Chomsky (1965). The meanings of pronominal references might be teased out via this process but the links between the sentences or propositions could not be indicated.

As models of human comprehension and language production processes, both theories lack the pragmatic aspects by which a sentence becomes meaningful in terms of the reader's experience and knowledge of the world. Their concentration on the universal aspects and on a competence model of language results in the theories providing inadequate explanations of both these pragmatic aspects and such performance factors as the figurative use of language (which often violates the semantic feature selection restrictions). These aspects are an integral part of the language user's repertoire.

### CHAPTER THREE

#### Semantic Memory.

Linguistic analyses of comprehension assume that the reader processes text by extracting and subsequently storing the relevant semantic or syntactic features. Other analytic models, which are to be reviewed in this chapter have focussed only on the semantic features.

Tulving (1972) suggested that the structure of long term memory may be divided into two components: episodic and semantic memory. He pointed out that in long term memory the individual has available knowledge of episodes in his/her past as well as knowledge of facts, principles and rules and that these two types of knowledge are represented as separate memory stores. He proposed that there were also organisational differences in the two memories. One difference was that semantic memory (the memory that stores rules, facts and principles) may be productive and generative whereas the memory for episodes is not. It has been argued by others (Cofer 1976; Baddeley 1976; Gregg 1978) that whereas the distinction might be an interesting one, it is not a functional one. Gregg (1978), discusses two examples which highlight the interdependence of the two areas of memory.

1. Children learning the meanings of their first words do so in very context-bound learning episodes because they have no rich underlying base of knowledge to which they can attach the concepts. This example underlines the strong reliance of semantic memory upon episodic memory in the early stages of formation of memory.
2. If one is asked an unlikely question by a stranger, it can not be stored as 'unlikely' unless one has access to semantic memory about the content of the question, or about questions that a stranger might ask. This second example highlights the reliance of episodic memory upon semantic memory in the storage of some episodic information.



Shoben (1980) points out that all information is learned in a context and argues that it is necessary to accept that all stored information must originate as an episodic memory. A broader definition of semantic memory as an internal model of the world which incorporates both personal and general knowledge of the world is adopted by these critics.

Two aspects of semantic memory are relevant to this thesis: (i) The way in which world knowledge is stored in the long term representation; (ii) The way in which input information and knowledge of the world are combined when the information is presented.

The form of the structure by which knowledge is represented varies. One structure that has been proposed is that of the semantic network. Three early network models of semantic memory will be discussed: Quillian (1969); Anderson and Bower (1973); Rumelhart, Lindsay and Norman (1972). All three models assume that semantic memory is based on concepts. These are represented in word form and are separate from, but related to, lexical memory.

#### Early network models of semantic memory.

Quillian's (1969) model.

Quillian (1969) proposed a structure which represented the storage of information in hierarchical categories in long term memory. The structure was a network arranged in a hierarchical fashion. It consisted of units (or nodes) and links. Each unit represented a semantic entity which was connected to a superset. The hierarchy was based upon set membership with superordinate nodes representing sets or categories (for example, "animal"). At the next lower level units represented members of that category (for example, "bird", "fish", "mammal"). And at the next lower level were units which were instances of those categories (for example, "robin", "canary", "lark", "goose"). Units were connected to other units by means of unidirectional links from the set to the superset. Each unit has an unlimited set of properties, or descriptive features. Properties describe the unit to

which they are attached, and also units at subordinate levels in the hierarchy but not units at a superordinate level. If one instance of a category was an exception to the category in terms of a particular property, then the exceptional property was noted at the level where the particular instance was stored. The model carried out searches for relationships between concepts in a sentence verification task. Both concepts were activated and the search continued along the links between them until the relationship was established.

Quillian (1969) and Collins and Quillian (1969, 1970, 1972a) carried out experiments which supported the hypothesis that the network was hierarchically structured by measuring the search time to verify relationships at the hypothesised levels of the hierarchy. It was assumed from the model that when non-recent information had to be accessed, then the number of levels between the concepts (or between the levels at which the concept and the property were stored) would determine the search time.

This aspect of the model was not supported by other research. Rips, Shoben and Smith (1973) demonstrated that it took subjects longer to ascertain that a chimpanzee was a primate, than to ascertain that a chimpanzee was an animal. As "animal" is at a higher level in the hierarchy than "primate", it would be expected (according to the model) that it would take longer to confirm. This raised the question of familiarity, that is, whether a concept which is more familiar to the processor will be more accessible than a concept that is less familiar, and whether familiarity in this intuitive sense is more influential than the hypothetical hierarchical position.

In addition to the question of familiarity, further investigations by Collins and Quillian (1972a), Schaeffer and Wallace (1969) and Wilkins (1971) have shown that the 1969 model is not adequate to explain negative instances (that is statements which cannot be confirmed by stored information). The model had to be expanded to incorporate decision rules which

would explain degrees of semantic relatedness so that time taken to search and refute statements could be explained.

Quillian's (1969) model is limited as a representation of human memory, because it can only adequately account for one process, namely that of confirming information that accords with categories which belong to a hierarchy. It does not explicitly account for the encoding of new information or for the initial process whereby knowledge is structured into a hierarchy. Despite its limitations this is an important model because of the way in which it has influenced investigations into the representation of knowledge in memory. It has been particularly influential in those models which utilised Artificial Intelligence as a means of investigation. Some revision of the model has been made by Collins and Loftus (1975) in their Spreading Activation Model. This model will be discussed later in this chapter.

Collins and Quillian hypothesised that comprehension of texts proceeds in a similar manner to the search for concepts. That is, memory is searched for properties related to the text, all potentially useful units are scanned simultaneously until an overall meaning emerges which allows selection amongst the units. The initial scan produces hypotheses about the meaning of the text and the hypotheses are checked syntactically.

**Anderson and Bower's (1973) HAM model.**

Anderson and Bower (1973) developed a model of memory (Human Associative Memory or HAM) which was based on highly structured associations. The model employs a semantic network to represent linguistic and non-linguistic information stored in long term memory. The basic components of the network are propositions which consist of nodes and associations. Each association is binary and combines or associates two concepts. The node is the point at which two associations are combined. The associations in the network are one of four basic types or classes:

1. Subject and predicate: (S) and (P). The subject is a concept and the predicate tells us something about the subject.
2. Relation and Object: (R) and (O). The relation is a verb-like form and the object a concept. The association of R and O together can form the predicate.
3. Location and Time: (L) and (T).
4. Context and Fact: (C) and (F). The context (C) is a combination of location (L) and time (T). The fact (F) consists of the subject (S) and the predicate (P). The context tells where and when the fact occurred, the fact carries information about what happened in the context.

Propositions are formed from one or more of the above four binary associations. This can be represented as a propositional tree diagram.

The final nodes on the tree (the terminal nodes) do not break down into combinations of associations. These are the basic concepts of semantic memory and can be linguistic or non-linguistic as, for example, the representation of information in a scene. There is no limit on the number of propositional trees that can be connected to any basic concept or terminal node.

The model proposed by Anderson and Bower (1973) describes memory as concepts which are linked together by associations of a specific kind which produce a propositional tree. The network is a superstructure of these propositions that serves to associate the terminal nodes or basic concepts. In this representation, there is no distinction between semantic and episodic memory. The model stores both kinds of information in the same semantic network.

The authors propose that the associations are stored and recalled as independent items rather than as an integrated single gestalt. Their experimental work supports this (Anderson and Bower 1973) but the work of Bransford and Franks (1971) does not, and Foss and Harwood (1975) question the interpretation of their results and propose a gestalt interpretation of

sentence memory.

The process by which new information is put into the structure relies upon a parsing process. Using this HAM encodes the input sentence(s) into a propositional tree. Next, the terminal nodes of the tree are matched with their corresponding locations in the existing memory network. A MATCH process then attempts to find a proposition tree in memory that looks like the input. This involves a search from each memory location that matches the terminal nodes. The search is for paths in the memory that connect terminal nodes in the same way as they are connected in the input. If a matching tree is found in memory, the input sentence is said to be comprehended.

The process then is data driven and information in memory is accessed by the encoded input. New information attaches itself to the stored knowledge of the world. The model provides no explanation of what might happen when new information is introduced for which there is no existing node. It is presumed that the MATCH process can not operate in these conditions. Also, it is not really clear how the 'episodic' content is encoded, nor can the model account for the figurative use of language.

**Rumelhart, Lindsay and Norman's (1972) model.**

Rumelhart, Lindsay and Norman's (1972) model consists of nodes and relations (or links). However, in this model, nodes represent any cluster of information in the memory. The nodes are of two kinds: primary nodes which refer to general concepts and secondary nodes which refer to specific instances of general concepts. For example, the general concept of a cat would be a primary node and the Cheshire cat in "Alice in Wonderland" would be a specific instance of a cat and therefore represented by a secondary node. A relation is an association among sets of nodes. The relationships are not hierarchical (like those of the Quillian model). The links can be uni- or bi directional. The relationships are not binary (like those of the HAM model). Any number of relationships can be attached to a single node.

The Rumelhart et al. model is far more concerned with the encoding of episodic information than Anderson and Bower's HAM. In the Rumelhart model every item of information in memory is encoded in the form of a node plus its relation. The authors distinguish between three classes of information. These are concepts, events and episodes. Concepts refer to a particular idea, for example a cat. Events are action based and involve actors, objects and a general scenario (for example 'the cat grins at Alice'). Episodes consist of a series of events (for example, 'the cat grins and then disappears'). All items of information in these classes are stored as nodes but the relations that connect them differ:

1. Concepts require three kinds of functional relations; the classes to which the concept belongs, the characteristics which define it as a member of that class, and the examples or subsets of that class. Three labels specify these functional relationships:

ISA (is+a) defines set membership;

IS and HAS defines property relationships (HAS for objects, 'the cat has teeth'; IS for qualities, 'the cat is fat').

To specify an example or subclass, the inverse of ISA is used to define the direction of the relationship.

2. Events are encoded by the use of different kinds of relationships. Rumelhart et al (1972) based these relations on the case grammar of Fillmore (1968). In the Rumelhart network events are characterised by the concept of an action. The action is represented as a node and the various 'cases' (Fillmore) or subsidiary features of the action are the relations.

3. Episodes, or clusters of events, are denoted in the model by special relationships. Rumelhart et al named these 'propositional conjunctions'. (Only two were given as examples.) These are : then (which leads to the next event in the sequence) <sup>and</sup> /while (which interconnects events with

unspecified temporal order).

The mechanism for clustering events into episodes allows for flexibility in the representation of information in the network. In the model, time is defined locally within the context of a specific event. The clustering mechanism allows any level of time scale to be represented by an embedding procedure.

New sentences are encoded into the semantic base in a series of stages. The first stage is to convert the information into a set of propositions. This is done according to a set of formal rules. Next the concepts in the information are related to stored general concepts. This input may cause changes in the stored information if characteristics of the specific concept which is being encoded matches characteristics of other encoded examples of that general concept. This leads to the generalisation of information, so that it is no longer related to a specific (or secondary) concept node but instead it becomes related to a general (or primary) node. This process illustrates how new information is not only passively matched to stored information but is active in causing reorganisation of world knowledge. The description however is not very explicit in its explanation of the matching stages.

#### **The LNR Model (1975).**

A more recent version of the model (Norman, Rumelhart and the LNR research group, 1975) produced a series of processes which act on the data base. These include answering questions, solving problems and understanding sentences. The 1975 model also incorporates the notion of primitive meaning structures into the model. The authors argue that the predicates of the propositions could be analysed into a set of semantic primitives which underlie their meaning. For example verbs such as buy, sell, steal, give, lend; are all 'transfer of possession' verbs. 'Transfer of possession' is then the primitive that underlies their meaning. This analysis of verbs can be applied

directly to the structure of events and a similar analysis to the structure of concepts.

The process for understanding sentences which is described in this newer model more clearly outlines the procedures whereby a new sentence is related to the stored information.

1. The sentence is parsed and transformed into a surface proposition. This involves finding the predicate name and listing the arguments.
2. The surface proposition is transformed into the semantic structure, which contains only primitive predicates.
3. The semantic structure is compared with previously stored information. The search is for confirmations, contradictions or partially redundant stored information with which the new information can be integrated. Three outcomes are possible. These depend on the state of the information in the sentence(s):
  - (i) If there is only old information in the sentence, then no new semantic structure is added to the memory.
  - (ii) If there is only new information, then the whole underlying semantic structure of the input sentence is added to memory.
  - (iii) If some information is new and other is old, then the new part of the underlying semantic structure is added to the old information in the memory.

#### **The early semantic network models: a summary.**

The three models (Quillian, HAM and LNR) which have been discussed above all assume that knowledge is stored in a semantic network. Each model describes a different kind of network, and each network has limitations as a model of human memory.

The Quillian model is the most limited model. It can only 'understand' a sentence in relation to prestored knowledge and its pre-stored knowledge is limited to hierarchies of set membership and properties of set members. It



cannot, therefore, understand other relationships between concepts, for example comparisons such as bigger than, or younger than. The model can understand an input such as "A budgerigar has blue feathers" if it already has budgerigar as a part of its subset of birds. It could not understand a figurative representation of the information such as "a budgerigar is a fluffy blue cloud" because clouds are not a property of birds. It is therefore a model which has very restricted understanding. Even if the limited definition of understanding is accepted, there still remains the serious limitation that was discussed earlier. That is, faster reading times resulted from intuitively 'familiar' associations between the concepts than the model predicted on the basis of the hierarchical connections.

The HAM model parses the input according to a hierarchy of associations. This means that concepts within a single input can be associated with each other and can 'understand' at that level. This means that a more complex type of sentence can be input and analysed and understood in terms of the specified associations among the concepts. The model can also 'understand' input according to what it already has stored in the network, that is if the propositional tree of concepts in the input match a proposition tree that already exists in the network. If not the input can be parsed, but not related to what is known.

The Rumelhart (1972) model is more flexible than the two previous mentioned models. It can understand set membership and property relationships between concepts in the particular input; relationships between the action concept and the other concepts in each propositional event and relationships between events which are described in separate sentences or in a single complex sentence. Concepts in sentences are not only understood separately but also within a larger context. At a different level of 'understanding' new input is matched to stored knowledge. This model is the most flexible of the three and goes further than the others in combining new and old information.

It can understand not only in terms of preexisting associations and hierarchies, but also in terms of past events and episodes. Like the other models of semantic networks it cannot understand figurative use of language, nor puns, nor jokes. It is more concerned with formal description of the analysis of the input, than with the use to which the reader puts the information.

Other forms of storage have been proposed. Some of these have been in response to the limitations of the models described above.

#### Alternatives to semantic network models.

A feature comparison model: Smith, Shoben and Rips (1974).

Smith, Shoben and Rips (1974) proposed a feature comparison model which would account in a different way for the classification times used by Collins and Quillian (1970) as support for their hierarchical network structure. The basis of their model was the proposal that the meaning of words in memory can be represented by a list of features. Two kinds of features were proposed:

1. Defining features : these are the most essential features. They define the characteristics necessary for membership of a category. For example, the defining characteristics of birds might be that they are alive and have feathers and wings.
2. Characteristic features : these are features that are usually but not necessarily possessed by category members. For example, 'birds can fly'.

The feature comparison model suggests that a verification task has two stages. First of all the two concepts are compared in terms of all their features. This stage decides how similar one concept is to the other. If the two are very similar or dissimilar an immediate response of true or false can be made. If the degree of similarity is somewhere between the two extremes then it becomes necessary to use the second stage. At this stage the model examines the defining features of the example to see if it possesses the

necessary features of the category.

This model can now account for one item of behaviour that the Quillian model could not: that is, the more typical the example is of the category, the quicker will be the reaction time. The feature comparison model predicted that human reaction time would be quicker when the example was a typical member of the category than when it was a less typical character. The hypothesis was based on the notion that the typical members would be recognised at the first stage of processing whereas the less typical members would require the second stage of processing. Thus, robin, sparrow and thrush would be recognised as members of a bird category more quickly than less typical members such as chicken, goose and ostrich. Experimental work supported this hypothesis (Smith, Shoben and Rips 1974).

The feature comparison model has the same limitations as the Quillian (1969) model. It is concerned only with the representation of knowledge in memory and the process whereby a sentence verification task can be carried out. It does not concern itself with the way in which new information is incorporated into the representation. The representation is on the surface different from the Quillian (1969) hierarchical model which it seems to supercede. However (as is pointed out by Slack, 1978) it would be possible to achieve the same responses with Quillian's model by weighting the properties in the same terms as those of Smith et al (that is, in terms of their typicality) and by initiating a search process which searched the most typical features first. This would then make the assumed form of storage immaterial as it would be possible to obtain the same results from either. However, an essential difference remains between the two models. The hierarchical model assumes that retrieval is a process of direct access to prestored information whereas the feature comparison model assumes that information is computed from the stored concepts.

Set theoretic model: Meyer (1970).

The Smith et al model was similar to the set- theoretic model of Meyer (1970). Meyer also was concerned with the way in which information was stored in memory and how it was retrieved in specific verification tasks. Again the model was not concerned with the problem of how new information was connected with the stored information. In this representation system a concept included both instances and attributes. For example, the concept 'bird' includes such instances as robins, canaries, sparrows and attributes such as 'has feathers', 'can fly', 'has wings'. In order to verify a sentence, Meyer proposed that two stages would be available. He demonstrated that whether one or both of the stages was utilised depended on the task, (whether it was asking if 'all x are y' or if 'some x are y') and on the relationship of the concepts (x and y).

The relationships that he used in his experimental work on sentence verification were:

- (i) The two concepts were identical.
- (ii) The first concept was a subset of the second concept.
- (iii) The first concept was a superset of the second concept.
- (iv) There was an overlap in meaning between the two concepts.
- (v) The two concepts were disjoint.

Meyer's experiments supported the hypothesis that concepts are stored in memory with attributes and instances. Whenever there is any overlap in meaning between the instances and the attributes, memory search will be faster. The greater the overlap, the quicker the verification. When there is no overlap, the attributes and instances of both categories have to be searched to check that there is no overlap. This increases the time taken to verify (or declare as false) the sentence.

More recent models of semantic networks

Spreading Activation Theory: Collins and Loftus (1975).

The Theory of Spreading Activation was a model produced by Collins and Loftus (1975). This was basically a network model which updated the Quillian (1968) model by incorporating a 'defining feature' component and by allowing the storage of properties at more than one node. These changes involved a change from a predominately hierarchical model to a heterarchical network model (however the notion of superordinate categories was retained). The model also specified the processes which search for information within the network.

The defining feature component in the model was conceptualised as semantic distance. This was represented in the model by the length of the line which connected the concept nodes. The shorter the line between the concepts, the stronger the association. The model also included a 'isnota' link. This allows some category exclusions (for example, 'a whale is not a fish') to be made directly. This category exclusion association accounts for the finding of Holyoak and Glass (1975) that many "false" decisions are made in a sentence verification task more rapidly than can be accounted for by a memory search.

The Spreading Activation Theory takes its name from the way in which the model works. A concept node is activated when the processor sees, hears or thinks about the concept. The activation of one node activates adjacent nodes and these in turn activate their adjacent nodes. Activation spreads through the memory but is decreased by time and distance. The result being, the shorter the links between the nodes, the stronger the activation, and the sooner the effect of the activation occurs.

The spread of activation allows for the verification of propositions. In the process of verification, the first concept is activated. This is followed by the activation of the second concept. The activation spreads from both nodes until an intersecting node is accessed. At this point the activation level must be sufficiently high to pass a decision threshold. The activation

level is set as a function of the strength and type of the associations. If the activation level does not reach the threshold, then it is ignored. If the threshold is reached a decision (about the content of the statement and its relationship to the content of the activated pathways) is made. The experimental work (Loftus and Collins, 1975) which supported this model describes a wide variety of data from verification tasks, all of which support the model.

The new model has greater flexibility than the original Quillian(1969) model because of the additional features that have been described. It is this aspect of the model (its flexibility) that has been criticised. Smith (1978) has argued that with the many assumptions of the model it is not surprising that the model can account for so many empirical findings. However the assumptions of the model were made to account for the empirical findings which the Quillian model failed to explain.

This criticism highlights the problem with current models: they have developed in response to empirical evidence that has highlighted weaknesses in previous models. In so doing they have had to continually increase the number of assumptions underlying the model so that it is not possible to make precise predictions, about what is possible and (more importantly ) what is not possible in terms of the model. It would seem that the response to a weakness in a model is to increase its complexity so that there is a structure which will account for the evidence that challenged it. In this way, the model becomes increasingly general and less able to make falsifiable predictions.

#### Anderson's (1976) network model: ACT.

In a similar way to the Quillian model, the HAM model has been updated by Anderson (1976). The new model, ACT, separates the factual or declarative knowledge in the system from the procedural knowledge. The factual knowledge is embodied in the semantic network. This has been modified so that complex sentences can be broken down into several propositions which are represented

separately but combined within an overall propositional tree. The procedural knowledge (about how to perform various cognitive tasks) is embodied in a production system. All productions are condition-action pairs. The condition specifies the features that must be true of memory. The action specifies a sequence of operations which are executed when the conditions are 'satisfied'. At times these operations specify changes that need to be made to the existing memory. The production system employs a parallel form of search that replaces the serial processing utilised by the MATCH procedure in the HAM model.

The procedure used by the production system is similar to that of the spreading activation model. An important difference is that at each node the activation is divided among the alternative links in proportion to their respective strengths of association. Because the model is similar to the spreading activation model and because it has evolved in the same way, the same types of criticism can be levelled at it.

Anderson (1976), recognises this possibility in "Language, Memory and Thought" (the book in which he outlines the ACT model of semantic memory):

"Another remark that needs to be made about the empirical accountability of ACT is that one cannot seriously expect to perform a single experiment and slay ACT. If ACT makes a prediction that proves wrong, that exact version of ACT will have to be abandoned but I am obviously going to propose a slight variant of the theory with slightly changed assumptions that is compatible with those data. It would be nice if the theory could be reduced to a few critical assumptions that could be subject to simple experimental tests. However, things just do not work that way. Act is only going to be rejected by repeated attacks that keep forcing reformulations until the point comes when the theory becomes unmanageable with its patches and bandages." (Anderson, 1976, p.532.)

The problem with Anderson's model is not that the theory becomes unmanageable, but that it is not a theory to begin with. It can more

accurately<sup>be</sup> described as a description, or a model or an analogue. This will allow it to evolve in such a way that it will stay manageable but will be unrecognisable as the same thing as earlier versions. This point is discussed later in relation to the Bradshaw and Anderson (1982) paper on Elaboration of Associations (Chapter Eight).

### Criticisms of ACT.

Wexler (1978) made extensive criticisms of the ACT model and of Anderson's (1976) theory on which the model was based. The criticisms were at three levels: at the level of the model; the theory; and the general approach of cognitive psychology to the problem.

1. Wexler's criticisms of the model are similar to those already made about representations of semantic memory, namely that the model has become too general. Wexler makes the following summary of his criticisms of

".... [the model] is simply so weak that there is no way to find evidence either for or against it. Any phenomena can be represented in ACT, not only the phenomena that turn out to be empirically true, but also those that are false. Sometimes particular models which do constrain the data are stated. But these models constrain only small, tangential aspects of the theory so that the theory ceases to deal with the complex cognitive phenomena which it is supposed to deal with. This is true, for example, of the experimental evidence for particular assumptions. In addition, often the data are so weak (even compared to current standards, as in the case of linguistic phenomena) that even if the theories were stated more stringently the evidence could not be used to sustain the theory." (p.346).

More specifically Wexler criticises Anderson for using only measures of memory and response time. His main criticism of this method is that with only these measures of what is and what is not possible in cognitive processing, important differences in behaviour are being ignored. In addition the use of



production systems allows all behaviour to be described in terms of this system but there is no explanation offered.

2. Wexler's criticism of the theory is on the grounds of its breadth and lack of predictive powers. He says of the theory that

"It is so broad that it is almost untestable." (p.334) and "Anderson's goal is to build a system that will mimic all aspects of a human's cognitive abilities." (p.335.)

Because it is so broad and because the production system can describe any behaviour, Wexler criticises it in the following way:

"There is no explanatory power in ACT because there are no restrictions on human abilities. Suppose, as an analogy, that somebody presented a theory of gene action and of inheritance in which he first claimed that his theory would allow humans to have children who were humans. We would immediately ask, however, if there were any favoured status in the theory to the prediction that humans would have human children, or whether humans could, so far as the theory specified, have other animals as children. If there were no favoured status to the former prediction, we would immediately reject the theory, or consider it vacuous. Such is the case with ACT. Since, to the extent that it is specified, any behaviour is allowable (subject to memory limitations) we should have no confidence that it accurately reflects human abilities." (p.338.)

3. Wexler's criticisms of the approach of Cognitive Psychology in general is mainly focussed on Anderson's work. He is critical of the approach which sets out to account for certain empirical phenomena rather than on the construction of a strong well defined theory with clear, restricted and testable assumptions.

Anderson is quite clear that his aim is to produce an account of certain empirical phenomena. This is mainly because he believes that the same item of behavior can be accounted for by differing models. This makes it impossible

to produce a unique description of the underlying mental structure and organisation. By approaching the problem from the evidence provided in empirical research he believes it is possible to build a complex model which will account for the disparate evidence.

The contrasting viewpoints represented by Wexler and Anderson typify the disparate and apparently unreconcilable approaches to the construction of theory which is current in Psychology.

### Conclusions.

Models of semantic memory are concerned with a description<sup>of</sup> the process by which input is analysed and of the structures by which the analysed text is stored. All input is understood in terms of previously stored information. The models have expanded over time to encompass the analysis and storage of a greater range of associations between concepts. As a result the models have become progressively more complex in their descriptions and have become increasingly flexible in their attempts to encompass all examples of understanding. They do however limit their approach to storage and recall of the concepts in the input. They are not concerned with the ways in which understanding might be used in a constructive manner, for example to create new insights into the solution of a problem or, at a linguistic level, how figurative language might be encoded so that it can construct novel associations between previously unconnected concepts.

CHAPTER FOURConcept-driven Processes.

The semantic network models described in the preceding chapter focussed mainly on sentence verification tasks; that is, on how input information could be checked as true or false against knowledge stored in the structures. Some consideration was also given to the manner in which incoming text was combined with this knowledge structure. For example, the LNR (1972) model showed how incoming data might reorganise existing structures when properties of a specific concept matched those of other encoded examples. In this case, the property of the specific concept became a property of a general concept. Another example is the ACT (1976) model which gave some explanation of how a production system could modify the storage structures to accommodate new knowledge, the action part of the production system having specified the changes that need to be made. In these models it has sometimes been assumed that incoming information is decomposed into more basic underlying meanings before fitting into the structure.

These semantic network models are data driven models. That is they focus on the role of the input text in the process: an approach that in computer processing models is termed "bottom-up" processing. It is a tacit assumption in these models that the processes described can be applied to any incoming information and that the process is initiated only by the data which is introduced into the system. Other models (which are to be discussed in this chapter) have been concerned with the way in which knowledge from outside the text (that is knowledge which is already stored) provides the processor with global hypotheses about what is going on in the text. Such models are concerned with "top-down" processing or the influence of stored knowledge upon the processing of the text.

The conceptual knowledge store described in the semantic network models is not sufficiently abstracted to contain the global structures. The global structures contain information about what is usual or 'normal' in a concept, action or event, and are concerned with the combinations and sequences which give structure and stability to situations. They enable the processor to classify incoming information in a variety of ways, and to use the classifications to anticipate future events and to infer information which is missing.

The structures have largely been the result of work in Artificial Intelligence or in computer simulation models. They have been given a variety of titles: Frames, Schema, Plans, Scripts (which are general purpose memory models and can account for all kinds of stored world knowledge); Story grammars, Narrative structures (which relate specifically to knowledge stored about the structure of stories and narratives) and Scenarios (which relate to general world knowledge but which have been used with specific reference to text comprehension by Sanford and Garrod, 1981).

Beaugrande (1980) distinguished between the general purpose models when he classified them in terms of their perspective on how knowledge is stored. He used these classifications as a framework for the analysis of recall protocols (Beaugrande, 1980) but the distinctions he made between the various kinds of model provide a useful framework for the review of the models.

#### Knowledge based structures.

##### **Frames**

Frame theory was developed by Minsky (1975). He describes a frame as "a data structure for representing a stereotyped situation like being in a certain kind of living room or going to a child's birthday party" (Minsky, 1975, p.355). A frame is represented as a network structure of nodes and relations. In this respect it is similar to semantic network structure. But whereas a semantic network structure is concerned with the conceptual representation of specific items of knowledge, a frame represents knowledge

about things that are always (or generally) true in a situation. The lower nodes or terminals of the frame are 'slots' that contain specific data. The conditions whereby data can be assigned to the terminals can be simple or complex. In the simple condition, the terminals specify by means of a marker whether the item assigned shall (for example) be a person or an object. In a complex condition the relations among items assigned to several terminals are specified.

One of the important assumptions of frame theory is that the terminals are normally filled with a 'default' value. This means that the frame may contain many details whose supposition is not specifically warranted by the situation. These details are valuable in representing general knowledge and for making generalisations. In Minsky's conceptualisation of frames, the default assignments are attached loosely to the terminals so that they can easily be replaced by new items that fit the particular assignment. This allows special cases or variables to be applied to specific situations where the norm does not apply in every detail; for example, a 'buildings' frame would have default values of doors, windows, walls. If a building, such as a lecture theatre, was constructed without windows then the value 'no windows' would replace the default value 'window'. Related frames belong to 'frame-systems'. Each frame in the system differs in the way it represents actions, cause-effect relations or changes in conceptual or visual viewpoint. Within the frame system, the different frames share the same terminals.

When a frame is selected to represent a situation, a matching process tries to match values from the situation to the slots in the terminal nodes. These must meet the required conditions of assignment. If the the frame selected does not match the reality of the situation, that is to say the terminal values assigned by the situation do not meet the conditions specified, then the information retrieval network is required. The function of the network is to provide an alternative frame system to replace the

unsuitable one. According to Minsky, these interframe structures make possible other ways to represent knowledge about facts and analogies.

The theory of frames is not confined to linguistic knowledge, but is applied also to visual processing. However Minsky describes specific frames which are used in the understanding of linguistic input. Initially an individual statement is assigned to a temporary syntactic frame which is mainly concerned with verb and noun structures, the condition for this assignment is complex and involves word order indicator conventions. This type of frame decides whether the input is grammatically acceptable. At the next level the grammatical input is assigned to a semantic framework. The terminals in this frame are similar to those in Fillmore's (1968) case grammar (agent, instrument, actor etc) in relation to an action centred word. This aspect of frames has been elaborated and investigated empirically by Thorndyke (1977) and Thorndyke and Hayes-Roth (1979) and will be discussed more fully under the heading of Story Grammars and Narrative Structures.

Beaugrande distinguished frame models from other 'top-down' models in the following way: in the former models knowledge is viewed as a heterarchical array in which elements are arranged so that there is provision for access to potentially relevant elements. An example might be of a restaurant scenario frame. This would be a network of entries such as parts, uses, substances that restaurants would have. The format for such a frame would be a series of slots which fan out from a control centre with no set sequence for the use of the slots. In this conception, Beaugrande seems to confound the single frame (for example, the parts of a restaurant) with a frame system (for example the parts and uses of the restaurant). A further difference is then relevant: in a frame system, the heterarchical system will 'fan-out' from the terminal nodes and not from a central control system, and it is the selection of a particular frame (from the system) which best fits the terminal node information that becomes the point at issue. However the

essential point (that is, the frame as a static model and the lack of sequence in which to select or utilise the information) is valid for both a single frame or for a frame system. In this important feature Beaugrande's definition is in accord with Minsky's model.

#### Schema.

Beaugrande distinguishes schema models as those in which knowledge is stored as sequences of elements. The schema dictates an order in which it is necessary for the elements to occur. The order is necessary if the elements are to constitute an example of a schema, rather than a frame. For example, a restaurant schema would not contain only the plan of the restaurant layout, but also the sequence in which a customer would have to move through the layout to reach a table. The essential characteristic of Beaugrande's conception of a schema is its commitment to an ordered sequence. Support for the existence of such schema has been supplied by Whyte (1979) in her investigation of the language skills of illiterate subjects. She found that both literate and illiterate subjects were able to generate an ordered schema for changing a wheel, although the amount of detail varied between individual subjects.

Neisser's (1976) definition of schema is more dynamic than that of Beaugrande. Neisser describes a schema as that part of the perceptual cycle which:

"accepts information as it becomes available at sensory surfaces and is changed by that information; it directs movements and exploratory activities that make more information available, by which it is further modified." (Neisser, 1976, p.54.)

Neisser's definition agrees with Beaugrande's only to the extent that the schema specifies the type of information that is acceptable and the sequence in which the events can occur. It differs quite radically in that it allows

the schema to be modified by the information it receives. Another main difference is that Neisser's schema is active and directs the search for information; Beaugrande's definition implies a static model into which information is fed.

According to Neisser's definition the directing function of the schema is termed a planning function. In his view, schema are not fixed but adaptive structures that offer a possibility for development along certain lines, providing that suitable interaction with the environment occurs. This conception of a schema is more closely allied with Rumelhart's (1975) Story Grammar than the static model provided by Beaugrande.

Neisser and Beaugrande differ in their definition of schema in at least two important ways. Beaugrande's definition is of a rigid structure which defines a sequence within a situation, whereas Neisser's schema is interactive, can be modified by the information with which it interacts (that is, it offers a potential sequence should suitable information become available. The other important difference is in terms of the extent of the schema. Beaugrande defines a narrow function which is only one of the functions described by Neisser. Neisser includes a planning aspect in his definition of a schema. Beaugrande lists the planning function under a different descriptive label, because in his conception of stored knowledge, it represents a different perspective: one which includes the goals of the planner.

#### Plans.

Models in which the elements are viewed as an advancement towards a planned goal are classed (by Beaugrande) as plans. For example a restaurant plan would list the actions necessary for achieving the goal of visiting a restaurant and obtaining a meal. Whereas schema (in Beaugrande's classification) are representations of sequences in the structure of the situation (for example, the route through a restaurant is constrained by the



architectural design, the layout of the furniture, and the convention of customer access); plans are shaped by the goals of the character and the choices that the character makes, from stored knowledge, to accomplish that goal. A variety of actions may achieve the same goal. For the user, plans allow unusual goal achievement to be anticipated through the use of knowledge about suitable actions. In this way, each plan is a unique combination of chosen actions which anticipates the achievement of a particular outcome. According to Beaugrande's classification, plans are compiled from elements in a flexible manner whereas schema are rigid and static sequences of elements. He described knowledge stored in this form as "relevant to a person's plan in which elements advance the planner towards a goal" (p.164). The distinguishing features of this perspective, according to Beaugrande's classification, is the flexibility of the plans and the dependency of choice of action upon the end goal.

Schank and Abelson (1975) describe how plans are used in comprehension (that is, in the comprehension or understanding of the content of a discourse) rather than in the planning of actions. They use, as an example, the process by which <sup>a</sup>non-coherent collection of sentences in a discourse can be comprehended. In the following example it is made clear that a "plan" is used as a process rather than as a static framework:

"John knew that his wife's operation would be very expensive. There was always Uncle Harry.... He reached for the suburban phone book."

Schank hypothesises that a reader who reads this text will recognise a general goal for John of "raising money to pay for a legitimate expense". From this it follows that John's actions are moving towards the realisation of his plans to achieve that goal. His choice of actions are those which will allow him to move towards the desired goal.

Plans, then allow the reader to comprehend a discourse which, linguistically, is non-coherent. The procedure used is an identification of

the intentions which have led to the actions which are performed. These intentions are interpreted as the desired goal. Actions are then understood (or comprehended) as the realisation of plans to achieve goals.

It can be seen from the above explanation that plans can be used in two ways: to plan one's own actions to reach a desired goal or to interpret the actions of others in the light of a inferred goal. The emphasis on action is fundamental to Schank and Abelson's more general theory. Schank(1975) expounds this as a Conceptual Dependency Theory. The basic tenet of the theory is that there is a meaning in a statement over and above the words used to express that statement and in order to understand the meaning there is a need to relate the statement to knowledge of the world which is stored in the memory. Knowledge of the world cannot be stored in the surface form in which it is input. In its stored form information is reduced to primitive meanings. The basis of all propositions which are not descriptive is action. Events are made up of a primitive action, an actor and a set of objects that are dependent on that action. These are called conceptualisations.

The theory reduces all actions to eleven primitive acts. Schank (1975) claims that the verbs and the abstract and complex nouns of natural language are constructed from these primitives which are the building blocks of memory. ATRANS is one of these primitive acts and will be described fully by means of an example (Schank, 1975, p173):

"ATRANS. The transfer of an abstract relationship such as possession, ownership, or control. ATRANS requires an actor, an object, and a recipient.

Examples:

give = ATRANS an object from the actor to the recipient.

take = ATRANS an object from someone to the actor.

buy = two ATRANS actions, each causing the other. ATRANS of money from the first actor to the second actor; ATRANS of an object from the second

to the first."

The other primitive acts are similarly specific: For example, PTRANS (the transfer of physical location of an object) requires an object, an actor and a direction (as in the verbs fly or go).

The power of primitive acts is that they can point up similarities in words as well as highlighting their differences. According to the Conceptual Dependency Theory, the basic unit of memory is the conceptualisation. This is made up of the primitive act and the concepts that are dependent on it. Conceptualisations are related to each other with one of four causal links: result causations, enable causations, reason causations and initiation causations.

Plans, then, are reduced to a series of these acts which can be performed to reach the desired state. Schank and Abelson (1975) propose in addition that there are there are groups of actions which they label 'deltacts'. Deltacts constitute subplans for the realisation of specific goals. When deltacts are used often enough they become SCRIPTS. Schank and Abelson (1977) describe scripts as: "a structure that describes appropriate sequences of events in a particular context." (p.41.)

### Scripts.

Beaugrande's description of knowledge stored as scripts is not as inclusive as that of Schank and Abelson described above. According to Beaugrande, elements of knowledge stored as instructions of how to behave and speak in a certain role are classified as scripts. For example, a restaurant script has instructions for the customer, the waiter/waitress, the chef etc. to enact in an established pattern.

Schank and Abelson do not describe scripts as only stored patterns of prescribed role behaviour. Their description is rather more global and centres on a more inclusive stored knowledge of the processor. According to Schank and

Abelson (1975) a script is stored knowledge of the sequence of acts that have taken place within a certain setting, whether it is a restaurant, a dental surgery, a doctor's waiting room or any other place. The processor stores the actions that are performed by all the actors in the script. In this way he/she knows how all the participants behave and in addition what is and what is not an appropriate setting for the action. Thus, Schank and Abelson's conceptualisation of a script, contains 'stage instructions' plus instructions for verbal and physical actions by all the participants. These scripts are based on actions and experience and are not generalisations or abstractions of plausible happenings.

When comprehending events in actuality or in a written account a Schank script allows the comprehender to do two things: to infer those elements that are not explicitly mentioned (for example to infer that the diner in a restaurant has ordered a meal if it happens that a meal arrives without the comprehender seeing or reading that an order has been placed) to note as exceptional (or "place on a weird list") those events or states that do not accord with the setting or actions (for example, the ceiling falling in at a restaurant, or Father Christmas appearing at a Birthday Party). It is this second function which Beaugrande does not utilise in his classification.

The psychological reality of scripts (as described by Schank) as knowledge structures has been investigated by Anderson, Spiro and Anderson (1978) and by Bower, Black and Turner (1979). Both lots of research supported the existence of scripts, but the research by Bower et al raised questions about the form of storage. Bower et al investigated the ability of subjects to generate scripts of routine activities (for example eating in a restaurant or visiting a dentist); to segment the overall activities into their constituent scenes; to recognise and falsely recognise actions from a script which they had previously read; and to order scrambled texts into a canonical order. The results of these investigations supported the idea that scripts of familiar

activities can be generated and that there will be a high level of agreement about the characters, props, actions, the sequence of the actions, and the constituent scenes that make up the script. The memory investigations supported the idea that the input information about a script was incorporated into the stored script in such a way that subjects were unable to distinguish new from stored information in a recognition task and that input information was recalled in the canonical rather than given order of sequence. The work of Gibbs and Tenney (1980) also shows that in a recognition task actions which are usual in an activity (for example, giving an address to the taxi driver) but which are not included in the text, are falsely recognised as being a part of the text when they are explicitly presented later. Changes in unpredictable variables (for example, eating steak or chicken in a restaurant) are not falsely recognised.

In the discussion of their investigation, Bower et al focus on some problems that became apparent in script theory. Their results show that events which occur in scripts about a visit to the doctor and a visit to the dentist become confused in memory tasks. This points to an abstraction and generalisation of these particular scripts into, say, a general 'visit to a health professional' script. Schank and Abelson (1975) had proposed that scripts were rooted in actual experience (repeated deltaxts) rather than abstractions and generalisations from experience. A consequence of the research by Bower et al was a reformulation of the notion of scripts by Schank (1981), so that a script is reconceptualised as a constructive process by which information is gathered from various sections of memory, rather than as knowledge that is stored as a complete entity. This is supported by Galambos and Rips (1982) who showed that both sequential (script knowledge) and centrally based (frame knowledge and story grammars) are compiled as needed, rather than stored as precompiled knowledge. Schank's (1981) reformulation of script theory caused him to re-examine the structure of memory. Memory is now

conceptualised as a multilevel store in which information is stored as either event memory, generalised event memory, situational memory or intentional memory. Details of the levels are theoretically interesting but of little direct relevance to this discussion of scripts. It will suffice to say that according to the theory scripts are constructed from these stores as they are needed. Confusion between the doctor and dentist scripts can now be explained in terms of those memories which derive from events which are common to both activities being stored at the generalised event level.

### Summary.

Beaugrande's comparison or any similar simple comparison between models is made at the expense of the important detail which is essential to the individual models. Each model was devised as a separate conception of knowledge structures and not as a part of a grand theory of structures. It is unrealistic to expect that the separate models will piece together like pieces of a jigsaw puzzle.

Beaugrande's classification while allowing for the differences between the original models to be examined tends to obscure the overlap between them (for example, the overlap between Neisser's Schema and Schank's Plans) and does not explain fully the implications of the original model (for example, the role of experience in Schank's scripts). The essential difference according to Beaugrande's taxonomy is that each classification is progressively more complex than that which precedes it. Thus a schema is a frame with a rigidly ordered sequence of elements; a plan is a flexible, goal directed choice of sequence and a script is a series of instructions for an interactive sequence of behaviours. The definitions of the terms which allows for the ordering of the theories in this way has resulted in an oversimplification of original theories. Beaugrande's taxonomy has been useful as a framework for the discussion of theories and models which use the terms but the very simplification which has allowed clear distinctions to be made

between them has prevented the classifications from being a useful tool for the analysis of other models which are concerned with 'top-down' processing.

The structures in these models are created from the abstracted content of texts and situations but once they are created they are used as permanent structures to encode and recall new input. Story grammars differ from these models because the grammars are concerned with the structures that can be derived from the analysis of specific kinds of texts. They demonstrate how the same structure can be used in the analysis of all examples of these texts regardless of the content.

#### Story grammars and narrative structures in simple stories.

Research into story grammars (Rumelhart, 1975; Bower, 1976; Thorndyke, 1977 and Mandler and Johnson, 1977) has involved the analysis of very specific items of discourse (simple stories) in order to discover either a consistent structure which underlies all simple stories or a generative grammar from which the syntax of all simple stories can be constructed. The research has sought to establish the existence of an abstract structure that is separate from the content of an individual story. Experimental work (which will be described below) has sought to establish that these structures have a psychological reality and are not merely linguistic analyses or computer simulations.

Simple stories are defined by Mandler and Johnson (1977) as those which have a single protagonist in each episode (and by protagonist they mean one character or a number of characters who act as a single group). Their definition specifically states that it is this feature and NOT the length, the number of events, or the number of episodes that constitute a simple story. This feature is not proposed by Bower (1976) but the texts that he uses as examples of the structure and which he quotes in his experimental work are constrained in this way. Rumelhart (1975) states categorically that his

PAGE NUMBERING AS IN THE  
ORIGINAL THESIS



## Erratum

P. 80, 81, 82 are missing.

The content of these pages is printed as P. 83, 84, 85.

program would have difficulty in handling multi-protagonist stories, and Thorndyke (1977) states that his simple stories are those that deal with one main character and his attempts to achieve certain goals. Although only Mandler and Johnson make the specific reference to the single protagonist in each episode, it can be seen that this is an implicit and fundamental feature of each theory.

The three main versions of story grammar which are to be reviewed (Rumelhart, 1975; Bower, 1976 and Thorndyke, 1977; Mandler and Johnson, 1977) are all concerned with the construction of a grammar that has a dual function: that is to provide an adequate account of the production and the comprehension of simple stories. The difficulty of making comparisons between models of stored knowledge and their interactions with incoming information have been discussed above. However each of the models of story grammar can be compared to a greater or lesser extent to the schema model as described by Neisser (1976). They are flexible and the structure can be modified by incoming content. Their structure is hierarchical and this differentiates them from the sequential structure of scripts and the heterarchy of a single frame.

Rumelhart, 1975.

Rumelhart (1975) called his account of the internal structure of a story a schema. He produced two sets of rules, a set of syntactical rules which would generate the structure of all simple stories and a set of semantic interpretation rules which determine which of the syntactic categories are filled by the semantic content of a particular story.

The syntax is constructed as a set of rewrite rules for combining the syntactic categories which could represent the structure of a particular narrative as a hierarchical tree with a binary structure of nodes and links. The nodes represent the elements of the story (the syntactic categories) and the links represent the rule governed relationship between them. The terminal nodes of the structure are filled by the content of the story which is treated

as propositional units.

All the simple stories which Rumelhart selected for analysis could be represented by a hierarchy which combined some or all of eighteen syntactic categories. The syntactic rules define the permissible relationships between the eighteen categories. Some relationships are fixed, others are flexible. For example, a story can only consist of a setting plus an episode; a setting can only consist of one or more states; an episode must contain only an event and a reaction; an event can consist of a choice of mutually exclusive alternatives: either another episode, or a change of state, or an action, or an event followed by another event. When the structure created by these rules is represented as a network of a particular story the hierarchy is one in which the two top levels are fixed and consistent in all stories. The levels below these reflect the individual structure of a particular story and are created by rules which allow more options. The flexibility of these levels of structure allows single and multi-episode stories to be accommodated within the same rule structure.

Rumelhart's model also provides rules for generating summaries of texts. These are based on six semantic relationships by which story content is mapped onto the syntactic rules. Each semantic relationship ( AND; ALLOW; INITIATE; MOTIVATE; CAUSE and THEN ) has one or more summarisation rule.

Rumelhart's model provides a generative grammar for the production and analysis of stories. The rules of the grammar provide a hierarchical branching structure by which the content of single or multi episodic stories can be represented. The model or schema has a structure with optional slots which are arranged initially according to fixed and then according to more optional rules. The flexibility allows the structure to be modified in accordance with incoming information from the story content. This aspect fits neatly with Neisser's conception of a schema. The mixture of fixed and optional rules make this model more passive than the active information

seeking schema proposed by Neisser. Rumelhart's (1977) model develops the basic grammar to include and stress the importance of the goals in the organisation of the text. The goal-attainment organisation of the structure make this accord more generally with the active model described by Neisser.

Thorndyke (1977) and Bower (1976).

Thorndyke's (1977) model has a more rigid structure than that described by Rumelhart above. Thorndyke presents ten rewrite rules. These describe a structured framework which represents the abstract structural components of plots of simple stories. The structure is not a binary tree but is represented by a hierarchy with multiple choices at the second level. Below this level the options are either binary or multi choice. At some levels of the hierarchy in Thorndyke's model of the abstract structure (and this applies particularly at the second level) it is essential that all elements or slots are filled. According to this a story must have a setting; a theme; a plot and a resolution. They must be present in that sequential order. Like Rumelhart's model, the order becomes more optional at the lower levels of the hierarchy, but unlike Rumelhart's model the frame-like component suggests default values for those slots which are not filled by story content. The terminal slots of Thorndyke's model are always states or events.

The method by which an individual story is analysed and slotted into the structure is similar to that proposed by Minsky's (1975) Frame theory. At the bottom level (the terminal nodes) the text is broken down into propositions. These are syntactical units which contain a verb (which describes either a state or an event) and at least one noun. At the next level each event or state has a semantic structure that can be analysed using Fillmore's (1968) Case Grammar. At the next level events and states are grouped to fit the Episodes and States that map onto the fixed structure of Theme, Setting, Plot and Resolution. Thorndyke's experimental work was particularly concerned with

the importance of the Theme.

In his analysis of a text, Thorndyke divides the input text or story content into a series of propositions. Each contains a verb that describes either a state or an event and in this way the content of individual stories maps onto the terminal slots of the generalised structure. The framework includes slots for the goals and subgoals of the protagonist of each episode, the terminal slot for this being a special case of a state: the desired state.

In his initial description of the rewrite rules which outline the narrative structure, Thorndyke presents his model as a multi-level hierarchy with a set sequence of slots at each level. He also specifies which elements of the structure may be repeated in the analysis of a particular story. His experimental work supports the reality of narratives structures as psychological constructs. Two experiments have been carried out (Thorndyke, 1977). In the first of these the structure of stories is controlled in such a way that similar content is presented in four conditions: one in which there is an ideal structure (that is, the theme is at the beginning of the narrative and the structure accords with the structure described above); one in which the theme is not in the specified sequential order, but placed at the end of the story; one in which the narrative has no theme (that is, no causal connections) and one in which the temporal and causative connections have been removed. The results support the idea that the structure proposed by the grammar has an importance for the comprehensibility and recall of stories. Both the mean percentage of material recalled and the mean comprehensibility ratings for the passages declined significantly between each pair of conditions as the structure was progressively disrupted. The content remained the same in the first two conditions (Normal and theme after) but the decline in the percentage mean recall was statistically significant.

This evidence supports the idea that the structure of stories can be described as separable from the content and that an 'ideal' structure story

(that is, the structure described above) will aid recall and comprehension. Related to this it was shown that causal connections aid recall to a greater extent than temporal connections. Analysis of recall showed that the levels of the hierarchy described by the model have relevance for the recall of stories. The analyses showed that those propositions which are placed in the highest level of the hierarchy according to the grammar are more often recalled than those at the lower levels. This result generally supports the results of Kintsch and Keenan (1973) (see Chapter Six). However the methods by which the hierarchy was derived and the theoretical bases of each experiment are so very different that they can only generally support the notion that certain ideas in a story are more memorable than others and that these can be predicted by models of top-down and bottom-up processing and can be identified using differing methods of analysis. This is discussed later in the chapter and also in relation to Sanford and Garrod's (1981) Scenario model in Chapter Seven.

Thorndyke's work also supports the idea that the grammar is utilised as a top down process to aid learning. The results of his second experiment (Thorndyke, 1977) demonstrate that the recall protocols for a second story are facilitated if subjects are given two stories in which the same structure (as constructed by his rewrite rules) is repeated but in which the content and characters are different. There was no facilitation of the recall of a second story for those subjects who were given two stories with different structures but in which the characters were the same. Thorndyke's conclusion is that repetition of general structure aids recall whereas the repetition of specific detail causes confusion. Later work by Thorndyke in conjunction with Hayes-Roth (Thorndyke and Hayes-Roth, 1979) has confirmed this result using single sentences as stimulus material.

Bower (1976) describes work which was planned in conjunction with Thorndyke and which is based on Thorndyke's (1975, 1977) rewrite rules. Bower

emphasises the rules which describe the embedding process and the hierarchies produced as a consequence of the "selection" of these rules. One important factor which influences the embedding process is the goal of the main character. For example, the attainment of a goal may involve the character in the production of a multiplicity of subgoals which eventually lead to the attainment of the main goal. Bower proposes that in a structure such as this, goal attainment produces an overall structure. Without a goal, the structure of the story changes as the subgoals become episodes that are only temporally related. Bower proposes that the goal attainment structure is more tightly knit because the episodes are embedded. His experimental work examines the effect of a goal structure on the recall and on the comprehensibility ratings of a text. Texts were produced in which the same story was used but the structure was changed in such a way that the goal was at the outset of the story (and therefore explicit from the start), or was moved to the end of the story, or was completely removed. A further condition in which the sentences of the story were produced in totally random order was also used. The results showed that the condition in which the goal was the initial part of the structure was rated most comprehensible and had the largest percentage of recall. Next in comprehensibility and recall was the condition in which the goal was present but was placed at the end of the story. The no goal condition and the totally random condition were rated and recalled third and last respectively.

Bower's work supports the idea that structure is important in comprehension and recall. His fourth condition (totally random presentation of the sentences of the story), when compared with the first in which there is a goal structure, provides evidence to support the assertion that story structure has an existence and a relevance that is separate from the actual content of the story.

Bower's work focusses attention on optional elements in the rules. His

description of the hierarchies that are produced in this way and his emphasis on the goal structure as the essential variable would suggest that his conception of structure accords more generally with Neisser's Schema than with Minsky's frames, particularly in its plan driven conception.

Mandler and Johnson (1977).

Mandler and Johnson (1977) specify that a story must have a setting and one or more episodes. Each episode must have a beginning, a development and an end. In a single episode story the beginning will terminate in an event; the development will consist of a simple reaction which will terminate in an event (internal or external) or a complex reaction which incorporates two events, one which is an attempt to attain a goal and one which is the outcome of that event (the outcome may be a state rather than an event). Finally, in the single episode story, the end is a state or an event. If the story is multi-episodic, then the episodes can be embedded at any point within the structure of the initial episode. Mandler and Johnson explicitly detail the points at which the embedding can occur. The overall structure of any story is decided by the number and type of episodes which it contains. On the basis of Bartlett's (1932) work the authors predict that causal connections will be better remembered than temporal connections. On the basis of their model the authors explain that causal episodes are embedded within larger episodes, in the way that is described in this paragraph, whereas temporal episodes (those that are connected by an implicit or explicit "then") will not be embedded but will be placed at the same level in the hierarchy.

The content of stories is mapped onto the syntactical structure by means of two major transformational rules. These are deletions and reorderings. Deletions can be of events or embedded episodes at the beginning of an episode. Deletions other than this will lead to the violation of the story structure. Reorderings will be made if the story is presented in an order which does not agree with the ideal structure. The reorderings can easily be



made if the surface structure of the story contains causal connections.

From the transformational rules, Mandler and Johnson predict that the order of flashbacks in stories will be less well recalled over time than they will immediately following encoding. The authors further hypothesise that different structures can operate at encoding and retrieval and that the structure imposed at recall will be of greater importance than that used at encoding.

The structure used at encoding has three functions: to predict the incoming text; to know when a section of text is complete for storage and to know what must be held until further information is input. The recall structure retrieves the information after the transformational rules have been applied. According to Mandler and Johnson, the recall structures are the more important in the final form which is produced. They quote experimental work by Mandler and Parker (1976) to support this. Mandler and Parker's research demonstrates that when complex pictures are presented in which the organisation is unstable at input, the unstable organisation will be replaced by a more familiar schema at recall.

#### General discussion of the models.

The work of Bower (1976) and Thorndyke (1977) implies that the content of individual stories is mapped onto the structure as the story is encoded. Their experimental work shows evidence that unstructured story content is less comprehensible and less available for recall than content that is structured in a story-like manner. They offer this as evidence for a stable structure which underlies all simple stories and which is used in the comprehension of input text. It is acceptable that the evidence supports the idea of a structure in stories, it may however be argued that other theories would offer different explanations for the existence of structure. For example Kintsch (1974), whose theory will be discussed in the following Chapter, would argue

that the structure is created by the coherence of the text and that each structure is created individually by the input from the text. Kintsch and Keenan (1973) provide experimental evidence which supports Thorndyke's (1977) finding that propositions which are at a high level in the structural hierarchy of a story are most often recalled. However their hierarchical model is based on a "bottom-up" theory of text coherence whereas Thorndyke's hierarchy is based on a "top-down" knowledge structure. Sanford and Garrod (1981) offer an explanation which is built on stored semantic structures and not on syntactic story grammar (see Chapter Seven).

The research and the theories which investigate topdown processing raises a number of questions:

1. Can the structure of information be separated from the content?

Research which has been carried out by Kintsch (1974), Bower (1976), Thorndyke (1977) and the theories and models of Rumelhart (1975), Mandler and Johnson (1977) demonstrate that not only can the structure be separated from the content, but, that an "ideal" structured text is rated as being more comprehensible and has a greater mean percentage recall than less structured and non structured material which has a similar content.

2. Are the structures which are created or imposed at encoding the same as those that are utilised at recall?

Bower and Thorndyke's work implies that they are the same, that the story grammar operates at encoding and at recall. Mandler and Johnson, propose that they can be the same but this is not necessarily so. The research which supports this assumption (Mandler and Parker 1976, Bartlett 1932) would seem to imply that if material which is presented for encoding does not have a familiar structure then the subject will attempt to impose a structure upon it at recall. A phenomenon which Bartlett described as "effort after meaning". Research by Pratt, Luszcz, MacKenzie-Keating and Manning (1982) would support

this implication. They demonstrated that subjects could impose a structure on unstructured material, if they were instructed to do so. Their subjects transformed the material into a story <sup>they</sup> if/were told that they would have to reproduce it as a story for children. Other research by Craik and Lockhart (1972), Craik and Tulving (1975), Reder (1979), Cirilo and Foss (1980) and Bradshaw and Anderson (1982), supports the idea that the same structures are used at encoding and retrieval, but that it is the elaboration of the frame-like structures, given or created at encoding, which determines what is recalled. This research is discussed further in Chapter Eight.

Mandler and Johnson's (1977) work suggests that the encoding and the recall structures may be different but that this is not necessarily so. This unresolved question becomes redundant in the light of the next issue raised by the research.

3. If the recall and encoding structures are different, are they created as needed, or are they stored with the content?

This question was addressed earlier in this discussion to the issue of encoding structures before it was assumed that they may be very different structures. The issue was unresolved. Recent work by Schank (1981) and by Galambos and Rips (1982) has suggested that in recall of material, the creation of scripts or schema is a process that acts on stored material and that the process required by the situation operates on the memory stores when it is needed. This research would suggest that encoding and recall of material are separate processes and that at recall the content is recompiled by the process that is required at the time.

It may be concluded that structure is an important element in the encoding and recall of story like material, and that the the research so far suggests that the structure can be separated from the content. Research into Story Grammars does not demonstrate clearly whether the structures are created at encoding or whether stored structures are responsible for what is encoded.

4. At what point in the comprehension of a story is the structure abstracted?

According to Story Grammars it would seem that the structure is abstracted after the complete text has been read and the rules can be applied to the content. If this is so then the structure is a result of the comprehension of the content and not an aid to it. The process by which the reader abstracts the structure from the text is not discussed in the theories. It is however the focus of the models (discussed in the next three chapters) which were produced by Kintsch and his associates (between 1974 and 1979) and by Sanford and Garrod (1981). The issue will be further explored in the experiments reported in Chapters 10 to 12.

CHAPTER FIVEA Model of Text Comprehension Developed by Kintsch and his Associates.

One of the most detailed models of text comprehension is that proposed by Kintsch and van Dijk (1978). The model was based on the earlier work on propositional analysis (Kintsch 1974) and macrostructure of texts (van Dijk 1977). The 1974 model was one which considered the structure of a specific text and the way in which the structure was created from the input in a "bottom-up" process. The 1978 model considers how the individual structure is created within the constraints of a more general overall structure which is imposed upon the input by the reader's knowledge about the structure of particular kinds of text, such as simple stories or scientific experiments. It is a unique model because it describes in detail and with illustrative examples how the content of particular texts and the readers knowledge about texts in general, are combined to form a text representation in the reader's memory. In this respect it is possibly the most detailed and complex model of text comprehension that has been produced.

The model was further developed by Kintsch and Vipond (1979). The models produced up to and including this version proposed that the comprehension process involved two sequential stages. The first process was the creation of the coherence graph or the microstructural process. This was followed by the construction of the gist or the macrostructural process. An addition to the model (Kintsch, 1979) provided<sup>o</sup>/third process, the creation of the fact structure. This process proceeds in parallel with the microstructural process, creating the "meaning" of the text (that is, relating the input to stored schema of world knowledge) and providing a new basis for the creation of the macrostructural process.

The Microstructural Process: The establishment of a coherent text base  
(Kintsch, 1974).

Kintsch (1974) proposed a method of text analysis which would establish the referential coherence of a text base. The method is based on Fillmore's (1968, 1971) case grammar, and in the same way establishes the semantic relationships between the predicator and the arguments in each sentence.

As with Fillmore's case grammar the representation is the same for either active or passive sentences. Longer sentences may contain more than one proposition, so that the sentence:

"The old man smiled and left the room",

is represented as:

(OLD,MAN)+(SMILE,MAN)+(LEAVE,MAN,ROOM)

The first proposition establishes that there was an old man and the others providing other information.

As the complexity of the sentences is increased so the rules to manage that complexity are expanded and rules are provided to manage the production of embedded sentences, definite and indefinite description of nouns, quantification of an argument, modality of a sentence, logical implications and to indicate location, time and tense. For example (Kintsch, 1976, p91):

"Turbulence forms at the edge of a wing and grows in strength over its surface, contributing to the lift of a supersonic aircraft."

(FORM, TURBULENCE)

(LOC: AT, 1, EDGE)

(PART OF, WING, EDGE)

(GROW, TURBULENCE, STRENGTH)

(LOC:OVER, 4, SURFACE)

(PART OF, WING, SURFACE)

(CONTRIBUTE, TURBULENCE, LIFT, AIRCRAFT)

(SUPERSONIC, AIRCRAFT) [Example from Kintsch (1976) p91.]

Kintsch illustrates his procedures with a text base (an ordered list of propositions) derived from a sample text. For example:

TEXT: "The subjects were 20 female students. All subjects were volunteers."

TEXT BASE: (STUDENT, SUBJECT.)

(FEMALE, STUDENT)

(VOLUNTEER, STUDENT)

(NUMBER, SUBJECT, TWENTY)

This point demonstrates Kintsch's departure from Fillmore's analysis as he combines information from more than one sentence. This combination forms the beginning of the text coherence graph in which propositions are connected in a linear and hierarchical manner in terms of their shared arguments. The above text graph becomes the following coherence graph:

```
(STUDENT, SUBJECT)
  1 _____ (FEMALE, STUDENT)
  1 _____ (VOLUNTEER, STUDENT)
  1----- (NUMBER, SUBJECT, TWENTY)
```

The Kintsch and van Dijk (1978) process model of text analysis assumes that the listing, or parsing, of the text into the propositional text base is an automatic process that needs no further explanation. The (1978) model describes two processes: the establishment of a coherence graph (the microstructure of the text) and the derivation of the gist (the macrostructure of the text). For convenience of explanation these are described as two separate processes but it is acknowledged that the two may be interdependent. The text base described above is the basis of the microstructural model. The research on which the macrostructural model was based is reviewed below.

The Macrostructural process.

## a. The construction of the gist (Kintsch, 1974).

In his 1974 work Kintsch addressed the problem of texts which were longer than one paragraph in length. It is possible to produce a text base or a coherence graph of the micro-connections in any number of paragraphs. However, such representations may be too detailed and too self contained to describe the more global connections that outline the coherence between paragraphs. This pattern of coherence represents the gist of the total text. Kintsch introduced van Dijk's (1972) notion of macropropositions into his theory, arguing that when they are coherently connected they produce a macrostructure or global representation of the text. Van Dijk(1972) proposed that lists of propositions which were well connected, could be summarised and given a name by a single proposition. These resultant propositions could then be connected together as a linear sequence and the summarisation process repeated until the desired level of organisation was attained. The final abstraction ~~was~~ the macrostructure or global representation of the text.

## b. Macrorules: A revision of the process for the establishment of macrostructure (van Dijk, 1977).

Kintsch (1974) formalised this procedure and produced a hierarchical representation of an abstract text:

$$T = A+B+C$$

$$A = A1+A2+A3+A4$$

$$A1 = prop1+prop2+prop3... +propN$$



He illustrated this with an example text from an experimental report:

Text = Introduction+Method+Results+Discussion

Method = Subjects+design+procedure

Subjects = Propositions7+8+9+10

Prop7 = (STUDENT, SUBJECT)

Prop8 = (FEMALE, STUDENT)

Prop9 = (VOLUNTEER, SUBJECT)

Prop10 = (NUMBER, SUBJECT, TWENTY)

Van Dijk (1977, 1977a) refined this procedure for deriving the macrostructure of a text and produced a set of macrorules which would operate on the microstructure and produce the macrostructure. To describe the macrorules and the way in which they would operate, van Dijk (1977) redefined the microstructure upon which they worked. In the same way as Kintsch (1974), he proposed that the input text is transformed into its constituent propositional list. This, the implicit text base, has a theoretical underlying text base in which all inferences are made explicit. Thus all propositions necessary to give relative interpretations of each proposition in the implicit text base become available. It is on this hypothetical explicit text base that the macrorules operate. Their purpose is to produce the semantic macrostructure of the text in which each macroproposition is defined in relation to the others.

Like the micropropositions of a coherent text base, the macropropositions should also have a linear coherence which connects them in terms of their arguments. In order to demonstrate the operation of macrorules, van Dijk (1977) assumes an initial idealised text base which is fully coherent, that is one in which there is complete overlap of arguments between the propositions. Since there can be several levels of description of any event, macrorules must be recursive. It must be possible to apply them to organise global meanings into still higher level global meanings. From these

recursive properties it follows that the level of abstraction of a macrostructure is relative to the underlying level of propositional representation. This in turn may be a macrostructure with respect to a still more specific level of representation. So a macrostructure is typically more general than its corresponding microstructure. Van Dijk's (1977) four macrorules operate on the microstructure to produce the level of generality required. The Macro rules are :

(i). GENERALISATION

(ii). DELETION

(iii). INTEGRATION

(iv). CONSTRUCTION

(i). Generalisation.

The generalisation rule deletes certain general properties, which at some macrolevel have become irrelevant, and replaces them with a more general concept. It can operate on predicates in general and so applies to both nouns and verbs. It relies on a certain amount of world knowledge for its successful operation.

(ii). Deletion.

The deletion rule deletes from a given text base those propositions which contain irrelevant information. Each proposition expressed by a discourse is considered as relatively unimportant if it is not a condition for the interpretation of another proposition. However, a proposition will not be deleted if it has consequence for an event denoted by a macroproposition.

A characteristic of the Deletion and Generalisation Rules is that information which has been removed in the process of abstraction cannot be recovered from the macrostructure unless other cues are utilised.

(iii). Integration.

The integration rule covers the possibility that macroinformation may be directly expressed in the discourse. The more specific information of the

discourse may be deleted because its global information has already been expressed in the text by a proposition that serves as a macroproposition. That is, all detailed information may be deleted that somehow has been integrated into another proposition of the discourse.

(iv).Construction.

The construction rule directly replaces a sequence of propositions at the microlevel with a macroproposition. This is a similar operation to that of integration except that there is no macroproposition already present at the microlevel. The new macroproposition is constructed on the basis of information present in the text base. The micropropositions in the text base represent conditions, components and consequences of the global fact denoted by the macroproposition. For example:

"I bought building materials, laid foundations , erected walls and made a roof."

becomes:

"I built (a house)."

A macroproposition is sometimes constructed on the basis of incomplete knowledge. The missing information is created using frame-knowledge to supply default values (Minsky, 1975). For example, if someone buys building materials, lays foundations and makes a roof, the construction rule operates upon frame knowledge to infer that the person has also built walls. If this default value was invalid then conventionally the exceptional information would be supplied (for example, "The building had an open front", or " There were floor to ceiling windows throughout").

**Summary of the operations performed by the macrorules.**

Macrorules (i) and (ii) operate on information that is related by consequences and conditions. Information that is deleted in the process of their operation is that which is irrelevant to the argument of the discourse presented in the text. Once this has been deleted by the operation of these

Rules it cannot be recovered from the information left in the macrostructure. For example, the type of pet that is kept by a man cannot be specified from the generalised information that he has three pets; that a child's ball is red in colour cannot be deduced from the information that a child broke the window with the ball.

In contrast, Macrorules (iii) and (iv) operate on information which is coherently organised (that is related by the referential coherence between the propositions). The information that is summarised by the integration or the construction of macropropositions is recoverable by inductive reasoning. If a house is built, it can be reasoned that it is necessary to have built walls, foundations etc. and that this will be preceded by the provision of building materials.

#### c. Semantic Macrostructure and Superstructure: van Dijk (1977)

The macrorules now specified (van Dijk, 1977) produce a different global structure from that outlined by van Dijk (1972) and formalised by Kintsch (1974). In order to distinguish between the global structure which was produced by the earlier heuristic procedures and that produced by the macrorules, the 1972 and 1974 macrostructural process is referred to by van Dijk (1977) as the super-structure while the global structure produced by the current model is referred to as the semantic macro-structure. Van Dijk elaborates the notion of super-structures with the example of the super-structure of a story: the narrative superstructure. A tentative example of the formation rules of a narrative superstructure are:

NARRATIVE = ACCOUNT +MORAL

ACCOUNT = SETTING +EPISODE

EPISODE = HAPPENING +EVALUATION

HAPPENING = COMPLICATION + RESOLUTION

These rules apply initially to simple stories and some categories are

recursive. According to van Dijk, these rules are important to the macrorules procedure because they define the global syntax of the discourse. In this role they may place constraints on the macropropositions derived from the content of the text by the macrorules.

Kintsch and van Dijk's (1978) model.

The basic notions of the two global structures described form the basis of the macrostructural processes in Kintsch and van Dijk's (1978) process model of text comprehension. As stated earlier, for simplicity and convenience of description Kintsch and van Dijk considered the two processes (that is, the construction of microstructure and the construction of macrostructure) as separate procedures.

**a. Procedures which establish the microstructure (Kintsch and van Dijk, 1978).**

Kintsch and van Dijk (1978) model the procedures by which a processor with limited storage capacity can extract the microstructure and the macrostructure from a text. The procedures by which the microstructure is extracted will be described first. To produce the ideal and explicit text base described by Kintsch (1974) the input text must be transformed into a propositional list and all the inferences necessary to provide all the consequences and conditions for the propositions in the surface structure must be made. At this stage the model does not contain procedures to carry out these two stages of the process. However, the authors acknowledge the need for a parser to transform the input into propositions and a means of indicating where inferences need to be made.

The fully coherent, explicit text is established as an initial starting point. The process model describes how the processor attains this representation beginning from an ordered list of propositions which is arranged in the same order as the input sentences. The first step is to establish referential coherence. The text base is checked to see if this

exists. If there is some argument overlap among all of the propositions, then the text is accepted for further processing. If gaps are found then inference processes are initiated to close them. The model assumes that these operations cannot be carried out on the text base as a whole because of limitations on the capacity of working memory.

The processing, therefore, proceeds in cycles. The text is processed from right to left sequentially in chunks of several propositions at a time. The chunks may not all contain the same number of propositions. In order for the model to work, a maximum number of propositions must be specified for the input (again these may vary according to the characteristics of the text and the processor).

If texts are to be processed in cycles it becomes necessary to make some provision in the model for connecting each chunk to the ones already processed. The model makes the following assumptions with reference to this issue:

1. There is a working memory and a storage memory. Part of working memory is a buffer with a limited capacity.
2. When the propositions in the cycle are being processed a number of propositions, not exceeding this limited capacity, is selected and stored in the buffer.
3. Only those propositions stored in the buffer are available for connecting incoming material to already processed material.
  - 3.a. If there is a connection between current input and the propositions in the buffer, then the input is accepted as being coherent with the already processed text.
  - 3.b. If there is no connection, a resource consuming search is made of all previously processed propositions. In reading comprehension this may require that the text is reread.
    - (i). If the search process is successful and a proposition is

located that shares at least one argument with the input proposition then the set is accepted and the processing continues.

(ii). If a search process is unsuccessful an inference process is initiated to add one or more propositions to connect the input set to the already processed propositions.

The model proceeds in cyclic fashion through the whole text constructing a network of coherent propositions. This network is represented as a graph with the propositions as nodes and the shared referents as the connecting links. The graph can be arranged in levels: the top level being those propositions which would give the simplest graph structure (that is, those which have the greatest number of connections from them); the second level would be all the propositions that are connected to the top level. The third level would be all those propositions that are connected to the second level but not to the top level, and so on.

In this idealised coherence graph, the topmost propositions are not the most important or the most relevant, in the intuitive sense of the words. Their position at the head of the graph is only important in the sense that these propositions introduce concepts that are referents for propositions which appear later in the text. That is, they are important because they introduce presuppositions of their subordinate propositions.

The model assumes that establishing coherence demands few resources when there is argument overlap. However, memory searches and making inferences make relatively heavy demands on the available resources and contribute significantly to the difficulty of comprehension. If the memory searches and the inference making procedures were not initiated then the results would not be a coherent graph of the text base but unconnected clusters of ideas.

When the text is processed in this cyclic fashion, certain propositions may be processed more than once by being stored in the buffer and processed again in the next cycle. The reprocessing will increase their chances of being

recalled from long term storage and thus increase their importance in the text. Which propositions will be favoured in this way depends on the nature of the process which selects the propositions to be held over from one processing cycle to the next. The selection strategies adopted for the model are those of Importance and Recency:

(i). Importance.

Those propositions which are well connected in the graph already constructed are likely to be more important than those which are less well connected. Therefore those propositions that are at the top level in the coherence graph will be given preference.

(ii). Recency.

If there needs to be a choice between two propositions that are equally important on the basis of the definition given above, then the most recently processed of the two will be selected because it is more likely to be connected to the next input cycle.

The model proceeds with a maximum number of four propositions in the memory buffer.

b. Procedures which establish the macrostructure (Kintsch and van Dijk, 1978).

The coherence graph, formed in this way with the constraints on propositional input and on buffer storage, is operated on by the macrorules. In this model (Kintsch and van Dijk, 1978) only three of the macrorules are operative: Deletion, Generalisation and Construction. The Integration Rule has itself been integrated into the Construction Rule. Further clarification of the Deletion Rule is offered :

"'Delete' here does not mean 'delete from memory ' but 'delete from the macrostructure.' Thus, a given text proposition - a microproposition - may be deleted from the text's macrostructure but, nevertheless, be stored in memory and subsequently recalled as a microproposition." ( Kintsch & van Dijk, 1978, p372).



The macrorules are applied under the constraints of a schema. The schema can be of two kinds: it can either be the superstructure (as described above in this chapter and by van Dijk 1977) or the reading purpose of the processor (the special purpose overriding the text structure). The macrorules proceed in the following order. The first macro-operation (Generalisation) is to form all generalisations of the micropropositions (as the theory has no formal inference making component, the required generalisations are supplied intuitively). The second macro-operation (Deletion) takes those propositions that are not generalised and deletes them if they are irrelevant. If they are relevant they become macropropositions. Relevance is decided by the schema (in the example text that illustrates this model the schema is an experimental report schema). The schema picks out those generalised and those remaining micropropositions that are relevant to its particular subsections plus those that are conditions and consequences for the selected relevant propositions. The third macro-operation (Construction) operates as necessary in the manner described above.

The rules are now applied recursively directed by an increasingly strict relevance criterion, until only the most relevant macropropositions for each subsection of the schema remain along with the most important antecedents and consequences. This then constitutes the macrostructure of the text.

### c. Limitations of the 1978 model.

The limitations of this model were acknowledged by the authors as being:

1. The input and output to the text deal with a representation of the text and not with the text itself;
2. The model does not contain a theory of inference. It merely indicates that there is a need for an inference making procedure to be initiated;
3. The model stops short of comprehension because it does not deal with the organisation of the propositional text base into facts, which would

utilise stored world knowledge;

4. The construction of the microstructure and of the macrostructure have been considered as separate processes when in fact they may be interdependent.

Further work by Kintsch (1979) and Kintsch and Vipond (1979) has reconsidered some of these issues.

One purpose of the model was to enable the authors to predict the difficulties of comprehending a text. It was hypothesised that, if long term memory searches and inference making were required to establish a coherence graph, they would make demands on processing resources. This could be measured in one of two ways. If the processes were carried out, each memory search and inference made would increase comprehension time. If they were not carried out the text would be imperfectly understood (this could be established by testing). If no long term searches or inferences were required then only those processes that made relatively small demands on the resources (those that were necessary for establishing coherence amongst existing arguments) would be used and the text would be easy to comprehend. In general, the number of long term memory searches made will vary according to the numbers of propositions in the input cycle and in the buffer.

#### A Modification of the 1978 model (Kintsch and Vipond, 1979).

Kintsch and Vipond (1979) explored further the implications of varying the number of propositions in the input and in the memory store. They produced Coherence Graphs for four texts. Each text was processed in four different ways:

- (i). The input size was held between 6 and 10 propositions and the buffer size was fixed at 4 propositions;
- (ii). The input size was held between 6 and 10 propositions and the buffersize was fixed at 7 propositions;

(iii). The buffer size was held at 5 propositions and the input fixed between 2 and 5 propositions;

(iv). The buffer size was held at 5 propositions and the input fixed between 11 and 16 propositions.

(The input size is a range rather than a fixed number of propositions because, in accordance with research by Jarvella (1971), the size of input of text is constrained by the boundaries of the phrase or sentence. This makes it impossible to input a fixed number of propositions when using actual rather than theoretical texts in an example).

The coherence graphs produced in this way were compared with the ideal text graph and the number of reinstatements, memory searches and inferences which were required in each condition were calculated. When a small buffer was used extra searches were required whereas when small inputs were used an extra number of processing cycles were needed. Because of this the smaller input had reprocessed many more propositions far more frequently than the larger input. If the most frequently processed propositions were the basis of recall of the paragraph, as Kintsch and Vipond now imply, this would mean that the recall of the text would differ from the ideal text base. To attain the same coherence as the text base the text must be reprocessed and the propositional graph reorganised. If carried out, these extra processes would add to the processing time. If they were not carried out, the representational text graph produced would be incoherent. This in turn would affect subsequent retrieval of the text. Kintsch and Vipond considered that the interaction of these two factors (time taken to read, and a retrieval measure) would be indicators of the difficulty of comprehending the text.

Kintsch and Vipond proposed further that input size would be affected by the familiarity of the text and that familiar text would be processed in larger units than unfamiliar text. The difficulty of a text would now depend not so much on the coherence of the ideal text graph but on the familiarity of the information and the processing capacity of the reader in creating the

coherence graph. A further difference that is suggested in the discussion of the model is that perhaps the reader's purpose also intervenes at this stage (the construction of the coherence graph) and that this is responsible for the selection of the propositions to be retained in the buffer. If this strategy were employed rather than the importance and recency strategies the result would be a selective rather than a coherent representation. This strategy along with the larger input of propositions would account for the skimming strategy of reading.

In this model the process which would produce the macrostructure was still described separately from the process which produced the microstructure. The microstructure was encoded and then processed further to produce the macrostructure. It was proposed that a cyclic operation similar to the microstructural process would occur to establish the coherence of the macrostructure (once the macropropositions had been derived by the operation of macrorules). The difficulty of comprehending the global outline would be decided by the number of reinstatement searches, reorganisations and inferences that needed to be made. As in the previous models the macrostructure would establish the gist of the complete text.

#### Problems connected with the 1979 model.

The expansion of the microstructure comprehension process in this model has created a problem concerning the relationship between the recall of a paragraph and the gist of that paragraph in the wider context of the complete text. The recycling of propositions in the production of coherence leads to a recall based on the frequency of cycling of certain propositions. The operation of macrorules on the hypothetical explicit text base then produces macropropositions which are in themselves coherent in terms of reference (that is in connectedness of arguments between propositions), antecedents and consequences and which also relate to an overruling schema (either the superstructure or the reader's purpose). It seems that there is not

necessarily any overlap between the contents of the two structures.

The Fact Structure: Kintsch (1979).

Kintsch (1979) provided a solution by creating an additional process in the model which he named a 'fact structure'. The process which was necessary to create the fact structure operated in synchrony with the process which created the coherence graph. Both operated on the propositional list which had been derived from the incoming text. Kintsch's notion of a fact structure is based on Minsky's (1975) concept of frames (see Chapter Four). A frame would be established by information in the initial input and the slots created by the frame would be available for incoming propositions or as a basis for making inferences in the absence of information. Subsequent inputs would be attached to the slots in the frame. If information was introduced which was not relevant to the frame, a new frame would be created.

The fact structure would be concerned with the sense of the input so that information which was referentially coherent but nonsensical would be interrupted until a new fact structure or frame was created. For example, in the text below a new frame would be created either from the reader's knowledge of anthropological customs (in this case particularly about the ghost marriages of elder sons who died without wife or heir, in which the next eldest son married and begot a heir in the spirit of the deceased) or by elaborate inferences which would create a meaningful link between two inputs which on the surface appeared to be incompatible.

"Among the warriors were two unmarried men, Kakra and his younger brother Gum. Kakra was killed in battle. Subsequently, according to tribal custom, Kakra was married to the woman Ami".

The process which establishes referential coherence would not note these inconsistencies in meaning but would merely establish a link which made 'Kakra

was married' and 'Kakra was killed' coherent with the propositions:

(THERE, WERE)

(TWO, MEN)

(UNMARRIED, MEN)

(MEN, KAKRA, GUM)

The coherence process and the fact structure differ in the way they would treat the following input:

"The Swazi tribe was at war with a neighbouring tribe because of a dispute over some cattle. Among the warriors were two unmarried men.....".

The coherence process would register that there was incoherence between the propositions of the sentences, and would indicate that some inference would need to be made to link the tribes, the war, the cattle and the warriors in the inputs. On the other hand, the fact structure would make the inference easily by filling in one of the frame slots. The initial input (i.e. "The Swazi tribe was at war with the a neighbouring tribe because of a dispute over some cattle") might create this frame:

WAR: (was at war with )

actor: (the Swazi tribe)

opponent: (a neighbouring tribe)

cause: (because of) (a dispute over some cattle)

outcome:

The next input (i.e. "Among the warriors, were two unmarried men, Kakra and his younger brother Gum.") would add to the frame in the following way:

WAR: (was at war with)

actor: (among the warriors)

(of (the Swazi tribe))

(were two (unmarried men))

(Kakra and (his younger brother Gum))

opponent: (a neighbouring tribe)

cause: (because of) (a dispute over some cattle)

outcome:

In this fact process, when knowledge of the world is employed (in the shape of frames), a warrior is easily inferred to be a member of the tribe who is at war with an opponent. It is the process of creating the fact structure which notes inconsistencies in meaning. This creates new frames to encompass the discrepant information and allows inferences to be made which rely on similarity of meaning rather than referential coherence. The fact structure, when created, has a further function: it serves as a base for the creation of the macrostructure. Kintsch hypothesises that the propositions, when grouped together as facts, are the structure on which the macrorules will operate and that only those which are relevant to the schema are retained. The schema determines what is relevant, sets up expectations, calls for certain facts and infers them if they are not directly represented in the input set.

In this latest version of Kintsch's model, inference making ceases to be a concern of the coherence process. Instead it is incorporated in the process of creating a fact structure. This re-location of the inference making process accounts for the experimental data which will be discussed in Chapter Six (Vipond, 1980) which suggests that inference making has no effect on the time taken to comprehend text (when this is defined in terms of establishing coherence). The experimental data is discussed in the following chapter.

CHAPTER SIXPsychological evidence for Kintsch's models.

The various versions of the comprehension model of Kintsch and his associates (Kintsch and van Dijk 1978, Kintsch and Vipond 1979, Kintsch 1979) rest upon assumptions about the 'processor' of the text. One of the major assumptions is that this processor transforms the surface structure of the sentence into the underlying text base. By describing the text base according to its theoretical characteristics, it is possible to empirically test some of the effects that these characteristics would have on the reading process if they had a reality. By using human processors (as opposed, for example, to a computer implementation of the processor) it is possible to test whether these characteristics have a psychological reality in the reader.

Text characteristics and the empirical evidence which supports the existence of microstructural processes related to them.

The proposed characteristics of the transformed text base are these:

- a) It consists of an ordered list of propositions;
- b) Each proposition varies in complexity according to the number of arguments that are linked to the predicate (this in its turn is dependent on the number of cases linked to the predicate, as in the case grammar of Fillmore, 1971);
- c) Text bases having the same word length and number of propositions vary in complexity according to the number of different arguments employed;
- d) The text base has a hierarchical structure in which the propositions having the greatest number of links with other propositions are superordinate.



a. Sentences are coded as propositions.

Kintsch and Keenan (1973) devised an experiment to investigate the psychological reality of the propositional base structure of a sentence. They proposed that if the propositional base had a reality then the time taken to read a sentence should increase as more propositions are processed, and the number of propositions recalled should be related to processing time in a lawful way. These hypotheses were tested using the same stimulus material<sup>in</sup> two conditions: free reading time and fixed reading time.

In the free reading condition the average reading time was 0.97 secs per word. Mean reading times as a function of the number of propositions in the base structure of the sentences showed that one extra second of reading time was required for each proposition read. When the reading time was based on the number of propositions recalled correctly (that is, the number of propositions assumed to be processed) instead of the number of propositions presented, the extra reading time per proposition was increased to one and a half seconds.

In the fixed reading time condition reading time was restricted to three words per second. This choice of rate was based on results from previous experimental work (Kintsch and Monk, 1972) which had indicated that this was sufficient time for a subject to read through a text, but sufficiently restrictive to produce noticeable behavioural effects. The results showed that the recall of material was not as good as when reading time was free. This difference was greatest when a large number of propositions was presented.

The results of the experiment demonstrated that when reading time was unrestricted the presentation of extra propositions led to increases in reading time, while restricting reading time led to a reduction in the number of propositions recalled. This provided support for the notion that the content of a text is stored in a propositional form.

b. A proposition consists of a predicate and n-tuples of word concepts. These form the basis of semantic memory.

Kintsch proposed that the number of word concepts related as arguments to the predicator of the proposition could be one or many and (in accordance with Fillmore, 1971) that the number of related arguments in a proposition depended upon the number of cases utilised by the predicator. The composition of the proposition, in terms of word concepts, is a fundamental element of the theory. The whole theory rests upon the assumption that propositions do consist of multi-argument units and are of variable size. Without this basic tenet, the propositional base of a text would be very different in size and number of propositions from that utilised in the process model of Kintsch and van Dijk (1978).

Kintsch and Glass (1974) ran two connected experiments which used sentences having the same number of word concepts, but which differed in their propositional structure. One set of sentences consisted of one complex, multi-argument proposition per sentence, the other had two or three simple propositions per sentence. Subjects were asked to read and then recall these sentences. The results demonstrated that recall was in terms of propositional units not sentences: complete recall for the one proposition sentences whether there was one, two or three arguments per predicator was 93%, 91% and 92% respectively. With two propositions the complete sentence recall was 84% and with three propositions it was 74%. In both the two and three proposition conditions single propositions (i.e. partial sentences) were recalled. Kintsch and Glass used this evidence as support for the proposal that propositions were variable in the number of word concepts they employed as arguments but, despite the complexity involved, each proposition was processed as a complete unit: these units forming the basis of semantic memory.

c. The number of different arguments employed increases the difficulty of a text.

Kintsch and Keenan (1973) based the results of their investigation into the propositional base of texts on mean reading times, and treated all propositions as if they were alike. This assumption allowed a general statement to be made about the time taken to read propositions vis a vis time taken to read words, but it ignored idiosyncratic differences within propositions. For example, a proposition may contain a predicate and one or more arguments (see Kintsch and Glass above) and the arguments may contain new or repeated word concepts. Kintsch, Kozminsky, Streby, McKoon and Keenan (1975) investigated the number of different arguments as a variable.

Kintsch et al used the same procedure as Kintsch and Keenan (1973): subjects read a text at their own pace, immediately afterwards they recalled as much of the text as possible in writing. A two-by-two design was used: the text base was short or long (a short text was 20-23 words or 8 propositions in length, a long text was 60-63 words or 23-25 propositions long); the number of different arguments in the text base was few or many ('few' in the short text was 3 arguments, in the long text 7-8; 'many' in the short text was 7-8, in the long text 16-23). In each of the 4 conditions 3 texts were used.

The results showed that with unlimited reading time, the mean time taken to read texts which had few different arguments was less than mean time taken to read texts which had many arguments. This result was the same for both long and short texts.

The experiment was replicated using scientific material instead of the history texts used in the first experiments. The same pattern of results was obtained, although the actual times were longer in every condition. From these two experiments Kintsch et al. concluded that a longer reading time would result from a larger number of arguments in a proposition. It also provided further support for the psychological reality of a propositional text base.

d. The coherence graph has a hierarchical structure.

Kintsch and Keenan (1973) and Kintsch et al (1975) offered experimental evidence to support the assumption that a text base has a hierarchical structure in which the propositions which are most connected to other propositions in terms of shared arguments (or referential coherence) hold a superordinate position.

Kintsch and Keenan, in the experiment discussed above, found that when reading time was fixed and restricted, the number of propositions recalled was less than when reading time was unrestricted. Their analysis of those propositions that were recalled demonstrated that recall of propositions was not random, but that some propositions were consistently recalled and others were consistently neglected. Those that were recalled were those that occupied a superordinate position in the text base when the sentences were transformed to propositions according to Kintsch's theory. These findings were incidental to the designed study but were so consistent that subsequent division into theoretical superordinate and subordinate propositions resulted in recall scores of 86% for superordinate propositions and 74% for subordinate propositions.

Kintsch et al (1975) tested the importance of the hierarchical position in the text base for recall in a more systematic manner as an adjunct to the experiments described above. The main purpose of the experiment was to test the effect of the number of different arguments on free reading time and on the recall in free and fixed reading time. The texts used were paragraphs rather than the sentences used in the Kintsch and Keenan (1973) experiment. As in the Kintsch and Keenan experiment, it was found that in restricted reading time the number of propositions recalled was reduced. Again this was not random. Kintsch et al (1975) predicted which propositions would be recalled. The basis for the prediction was the superordinate and subordinate position of propositions. According to the theory the superordinate

propositions would be recalled better than subordinate ones. This was found to be the case for all superordinate propositions regardless of the serial position of the corresponding sentence in the text base.

The four sets of evidence discussed provide the psychological evidence from the work of Kintsch and his associates for the reality of sentence transformation into a propositional base with multi argument propositions and a hierarchical structure. The process models (Kintsch and van Dijk, 1978; Kintsch and Vipond, 1979; Kintsch, 1979) are based on these assumptions and this evidence.

### The processing capacity of the reader.

The process model of Kintsch and van Dijk (1978) makes several assumptions about the processor. These assumptions are initially about the processing capacities of the reader:

"This checking of the text base cannot be performed on the text base as a whole because of the capacity limitations of working memory." (Kintsch and van Dijk, 1978, p367.)

#### a. The role of working memory.

The model of working memory utilised is attributed by Kintsch to Atkinson and Shiffrin (1968). Incoming text that has already been transformed into a list of propositions, ordered according to their order in the surface text is fed into the short term memory: some items are held over in a buffer store to match up with new incoming information, other is fed into a long term memory for that particular text base. The model differs from that of Atkinson and Shiffrin in several important ways one of them being that in the Atkinson and Shiffrin model the buffer store is a sensory register that initially holds information in a relatively raw state until it can be processed. According to Atkinson and Shiffrin:

"The choice of which items to enter into the buffer is based on momentary characteristics of the current string of input items and may appear at times to be essentially random." (Atkinson and Shiffrin, 1968, p45.)

Kintsch and van Dijk saw the process of selection as being far from random. Rather it involved the sophisticated selection procedures that have been outlined above. These selection procedures would make heavy demands on the resources of the working memory. Atkinson and Shiffrin's model has been criticised by Baddeley (1976) because of its complexity, which allows for an indefinite number of control processes. Baddeley suggests that it is only because the model has been restricted to rote learning processes that it has avoided the sterility of merely postulating additional processes to account for any result that does not fit the existing system. The additional processes proposed by Kintsch and van Dijk would seem to make even more demands and call for the addition of new control processes to explain them.

Atkinson and Shiffrin's model was restricted to the processes in working memory that were appropriate to a rote learning task. Kintsch's model describes the setting up of a coherence graph from incoming text which is being processed in continuous cycles. This involves different processes. These include procedures for setting in motion memory searches for propositions that have been processed, reinstating propositions in the buffer, noting that inferences need to be made, processing rehearsed and incoming propositions for inclusion in the rehearsal loop.

In the Kintsch model the function of working memory appears to be restricted to establishing coherence between the current and the previous input, and the function of the buffer is that of a rehearsal loop. An outcome of this process is that those propositions not held over in the buffer for matching with the input are passed into a long term memory store for this text base. This representation is held until a coherent graph of the text base is formed.

b. The limitations on the size of input.

Kintsch and van Dijk's (1978) model assumes that the limitations imposed on the amount of text input is partly a function of the size of the working memory and partly a function of the surface structure of the text. It is implicitly stated that there will be an upper limit to the amount of input that can be processed at any one time but this will vary from individual to individual. The actual number of propositions within this limitation will be decided by the structure of the text. Kintsch supports this argument with evidence from Jarvella (1971) and Aaronson and Scarborough (1977) whose work demonstrates that sentence and phrase boundaries determine the chunking of text in short term memory. The upper limit will not be exceeded, but if the written sentence contains too many propositions for the capacity of the processor, then the input will be the nearest phrase boundary that contains fewer propositions than the processor's limit. A further limitation on the processing of the text imposed by Kintsch's concept of working memory is that the number of propositions which are held over for establishing coherence with the new input will be limited by the size of the buffer in the working memory. These limitations are in accordance with those outlined by Atkinson and Shiffrin who proposed that competition for resources in the working memory would mean that the buffer size would necessarily be restricted when other procedures were being carried out. They suggested that it was possible to maintain six items in the rehearsal loop but that three was a more usual number when other resources were being used.

Evidence that the reader processes the text as proposed by the Kintsch and Vipond (1979) model.

Kintsch and Vipond (1979) produced a model which predicted the effects on the reading process of different sizes of input and buffer capacity. A prediction generated by this model when the upper limit of the input was fixed at 10 propositions and the buffer size at three propositions was tested by

Vipond (1980). The model assumed that the processes of inference making and reinstatement of propositions from the stored text representation would require longer reading times if carried out. It was assumed that imperfect understanding would result if they were not carried out.

a. Existing evidence which links reading time to inference making and to the reinstatement of propositions.

(i). Inference making.

Evidence to support the assumption that inference making increases reading time is provided by the work of Haviland and Clark (1974). They tested the time taken to comprehend the second of a pair of sentences under two conditions. A condition in which no inference had to be made:

"Ed was given an alligator for his birthday. The alligator was his favourite present."

A condition in which an implicit inference was required to establish the coherence: e.g.

"Ed was given lots of presents for his birthday. The alligator was his favourite present."

Subjects indicated by pressing a button that they had understood the second sentence. Significantly less time was required to understand the sentences in which the same idea (e.g. alligator) was presented both times and no inference had to be made.

(ii). Reinstatement of propositions.

Evidence that reinstatement searches increase reading time was provided by an experiment carried out by Lesgold, Roth and Curtis (1979). Two sentences which were coherent and needed no inferences were produced:

"A thick cloud of smoke hung over the forest. The forest was on fire."

The two sentences were interspersed with additional material two sentences in length. In one condition these changed the topic:



"Glancing to one side, Carol could see a bee flying around the back seat. Both of the kids were jumping around but made no attempt to free the insect."

In the other condition they were consistent with the topic:

"The smoke was thick and black, and began to fill the clear sky. Up ahead Carol could see a ranger directing traffic to slow down." (Examples from Lesgold et al., 1979, p295.)

It was found that texts in the second condition (in which no reinstatements were needed) required a shorter reading time than the texts in the first condition.

b. Evidence that the reader will try to establish a coherent text base from the input.

The process model also assumes that the goal of the processor will be to attempt to establish an ideal coherent text base :

"The first step in forming a coherent text base consists in checking out its referential coherence; if a text base is found to be referentially coherent, that is if there is some argument overlap amongst all its propositions, it is accepted for further processing." (Kintsch and van Dijk, 1979, p367.)

Evidence which supported the existence of a coherence-establishing process was provided by Kieras (1978). Kieras demonstrated that when seven, simple sentences were presented in an order that would establish maximum coherence, they were easier to integrate and recall than the same sentences presented in a meaningful, but less coherent order. The time taken to comprehend the last sentence when the sentences were presented one by one was longer for the sentence preceded by the irrelevant material, the condition in which a reinstatement search was necessary.

Vipond (1980) made further investigations of the process. He selected 25 texts which ranged in length from 246 to 266 words. These were adapted from school text books and encyclopedias. There were five texts on each of five topics. Within each topic the texts varied in content and in difficulty. Topics which were familiar to the subjects were chosen. How difficulty was initially estimated is not made clear. A list of micropropositions was derived by using the method devised by Turner and Greene (1977) - a revised version of Kintsch's original method. The list was ordered according to the appearance of the predicates in the text. Propositions that were to remain in the buffer were selected according to the 'leading edge' rule of importance and recency.

Coherence graphs were drawn from these parameters which showed the predicted memory network at each cycle of the comprehension process. The model then showed the points at which memory searches, reinstatements and inferences would be required. These were tallied for each text. Two other measures were calculated: the number of reorganisations that would be required to equate the derived coherence graph with the "ideal" graph (derived without any limitations on input or buffer) and the number of propositions in each level of the hierarchy. These measures were used to predict the difficulty of each text. The five texts were read and then the subject was asked to recall the text in writing. Reading and writing times were unlimited. The recall protocols were scored for the number of propositions recalled. Reading time was recorded. Average percentage recall was divided into reading time score to produce a microcomprehension efficiency score. Intercorrelations between the variables and the efficiency scores were calculated. Significant correlations between the scores and three microvariables were found: reinstatements (that is, the number of propositions that are reinstated from the long term memory representation); reinstatement searches (that is, the number of times the memory would need to be searched to make a reinstatement) and breadth of

processing (that is, the breadth of the hierarchy of propositions).

Reinstatement /microcomprehension:  $r = 0.437, p < .05$

Number of reinstatements/ microcomprehension:  $r = 0.521, p < .01$

Breadth of processing /microcomprehension:  $r = 0.503, p < .01$

The number of inferences to be made was found not to correlate significantly with microcomprehension efficiency.

c. Evidence that a separate process will be used to establish the microstructure.

Kintsch and van Dijk (1978) assumed for convenience that the processes involved in the construction of the microstructure and the processes involved in the construction of the macrostructure were separate. Kintsch and van Dijk had considered a process whereby the macrostructure was constructed from the microstructure by the application of recursive macrorules these being repeated until a desired level of abstraction was reached. The Kintsch and Vipond (1979) model concerned itself with longer texts. In this model the macroprocess began at the point where the macropropositions had organised the text into subunits, each macroproposition serving as a label for these units. The macropropositions now needed to be organised into a coherent whole in the same way as the micropropositions were organised. According to the model the same kind of cyclic process continued to operate on the incoming macropropositions. The same limitations on input and buffer size were assumed with the consequence that there would be a need for reinstatements, reinstatement searches, inferences and reorganisations as were needed in the processing of the microstructure.

Vipond (1980) investigated two aspects of the macrostructural process:

(i). The same variables would be predictors of processing difficulty at macrolevel as at microlevel.

(ii). That the construction of the macrostructure was an independent process.

(i). Macro-variables and their correlation with macro efficiency measures.

Vipond used the same material and subjects that he used in his investigation of microstructural processes to concurrently check the text for macrovariable correlations. An ideal macrostructure of the texts was created. From this it was predicted how many reinstatements needed to be made, and how many reinstatement searches, inferences, and reorganisations of the text would be necessary and what would be the breadth of the hierarchy, when an input of approximately seven macropropositions (paragraph size) and a buffer size of four to seven macropropositions was used in a cyclic fashion. A macrocomprehension efficiency score was derived from the number of macropropositions and the time taken to read the material. This was correlated with the macrovariables. Two were found to have significant correlations with the score. These were:

Reinstatements:  $r = 0.535, p < .01$

Number of reinstatements:  $r = 0.576, p < .001$

Further analysis of the data showed two distinct clusters of variables, those related to macrostructural processes and those related to microstructural processes. This suggested that the macrovariables and the microvariables were measuring unique components of text structure. This led to the second investigation.

(ii). Processes that are necessary to establish the macrostructure are independent of processes which establish the microstructure.

Vipond (1980) reasoned that in one sense the two processes could not be defined as independent: that is, in the sense that macrostructures are derived from the microstructure. However, if the two processes were independent in the sense that they use separate resources, an extra load imposed on one of the processing tasks should not interfere with the other process.

To test this hypothesis the following experiment was carried out. A macroprocessing task was set under two conditions, normal and difficult. In the normal condition a text was presented with the paragraphs in the normal order. In the difficult condition the order of the paragraphs was scrambled. The order of the sentences within the paragraph remained the same. The microprocessing task within these texts was made easy or difficult by varying the number of words from the Thorndike-Lorge (1959) list of familiar words. If the processes were not independent increasing the difficulty at one level of processing should prevent the reader from taking advantage of the easy conditions at the other level. It was expected that interdependence would be indicated by an interaction effect between word and macroprocessing. The tests were administered with a fixed reading and recall time. Recalls were analysed. No significant interaction was found. It was therefore concluded that the reader used two different processes to create the microstructure and the macrostructure.

Summary of the evidence for the existence of the processes described by the models.

The model developed by Kintsch and his associates has been developed in stages over time. The stages of development coincide with the order of the processes in the model: the transformation of the text to the text base; the creation of a coherence graph; the development of a macrostructure and the introduction of a fact structure.

a. The transformation of the text to the text base.

No explanation is offered of the process by which input is transformed into a text base. It is only assumed that this happens. However differences in reading time show that there is a relationship between the number of propositions in texts which have the same number of words, and the time taken to read them (Kintsch and Keenan, 1973).

b. The creation of a coherence graph or the microstructural process.

The creation of a coherence graph has been modelled in detail. Kintsch, Kosminsky, Streby, McKoon and Keenan (1975) have supported the model by showing that texts which closely model an 'ideal' coherence graph (that is texts which have a lot of overlap between the arguments) are read more quickly than texts which have more arguments and less overlap. This has been demonstrated using texts in both conditions which have the same number of propositions.

Both stages of the model have used differences in reading times between the texts as empirical evidence for the existence of propositions and the coherence between them. This assumes that the differences in reading time can only be attributed to the processing of propositions. Whilst this is a plausible explanation, it is not conclusive proof.

c. The Macrostructural processes.

The third stage of the model is the development of macrostructure. This stage models the process by which readers derive the gist from the text. Two alternative models have been produced and a third tentatively offered with the introduction of the fact structure.

The first version (Kintsch, 1974) proposed that macrorules operate on and summarise clusters of propositions within the coherence graph. These produce a macrostructure which is a description of the structure of the text and can be compared to the story grammars that were discussed in an earlier chapter. The difference being that the structure in this model is created by a data driven, bottom-up process and is created from the content of the story.

The second version (Kintsch and van Dijk, 1978) proposes that the macrorules operate on the complete coherence graph and that the summarisation process is constrained by the reader's knowledge about the structure of texts. Only those propositions which are relevant to the stored structure are retained as text gist or macrostructure. However information which is rejected

as being irrelevant to the gist can be retained and stored separately from the memory for the text.

Experimental work was based on the assumption that the creation of micro and macrostructures are separate processes. Again it is assumed that variables which would create difficulty in the production of the coherent structures would increase the reading time. Efficiency scores were calculated from reading times divided by the number of propositions recalled. Texts were assessed for the factors which would increase processing time in the creation of micro and macrostructures and measures were taken of the correlations between the difficulty factors predicted and the efficiency scores obtained. Statistically significant correlations were obtained separately for the macrostructure and the microstructure of the text. Factor analysis of the data revealed separate clusters of variables, those related to the microprocess and those that related to the macroprocess. Further research which was analysed using an analysis of variance showed that there was no interaction between the macrostructural processing of the whole text and the processing of easy and hard word versions of the text. This supported the notion that macroprocessing was independent of one other process, but did not test the separation of micro and macroprocesses.

d. The fact structure.

The existence of the process by which the fact structure (Kintsch, 1979) is established has not been supported by experimental evidence. The plausibility of the model in relation to the Sanford and Garrod (1981) Scenario model is discussed in the next chapter.

### Conclusions

The early models produced by Kintsch and his associates (Kintsch, 1974; Kintsch and Van Dijk, 1978) were concerned with the effects of the referential coherence of input texts on the formation of gist and summary (the latter in

relation to the macrostructures which were assumed to be a part of the stored knowledge of the text processor). The early models describe a text driven (or bottom-up) process for the creation of structure: the process being constrained by stored knowledge applied in a top-down fashion by the processor. The result of the process is a unique structure which is encoded for each text on the basis of its coherence. These early models posed two problems for Kintsch: firstly how to resolve the problem of coherence when the referent for a proposition was implicit rather than explicit in the previous text; secondly how to represent "sense" in a representation (for as it has been shown earlier the coherence model did not distinguish between sensible and nonsensical texts, it merely distinguished between those which were coherent and those which were not).

A solution to both these problems was proposed by the introduction of an additional process to the model by Kintsch in 1979. The new process was the creation of a fact structure. The process assumes the existence of a knowledge structure in the processor which is similar to Minsky's (1975) frames. An appropriate frame is activated by the semantic input and incoming information is related to the slots in the frame. This structure accounts for implicit reference by referring the incoming text to the frame slots when coherence can not be established. This explains the inference making procedure in a way that the coherence model had been unable to do. (It also might account for the experimental results of Vipond, 1980, which suggested that making text-driven inferences was not an important factor in the microcomprehension or establishment of the coherence of a passage.)

Kintsch (1979) also demonstrated how seemingly nonsensical input could be comprehended by the selection of a suitable frame. This is illustrated by the Kakra example in the previous chapter. The rather unusual example which Kintsch selects demonstrates that the frame is created from stored information which is activated by a semantic input and that the resultant frame can be



idiosyncratic or esoteric. The example used relies on the assumption that the information which the slots can contain is determined by the topic of the frame. How the topic is determined is not specified but it would seem from the example that is given that the reader has begun by reading the text, has decided upon a topic which would fit the story and has created a frame which will accommodate the topic. As the topic cannot be decided until the initial input has been understood without a frame it would seem that initial bottom-up processing creates the frame from stored pragmatic knowledge of the world. Kintsch's model assumes this process but does not explain it. From this point the frame is used in a top-down fashion, providing potential values for the interpretation of incoming information. The underlying assumption would seem to be that the values of the slots are defined by stored concepts which are associated with the frame topic. Kintsch also states that a new frame is created when the current one is inappropriate. Again the process by which this is determined is not made clear but relies on the example as illustration.

CHAPTER SEVENSCENARIOS.The model of text comprehension produced by Sanford and Garrod.(1981).

Sanford and Garrod (1981) produced a model which accounts for the processing and comprehension of text in terms of explicit and implicit referents in the text. They have carried out a series of experiments (Garrod and Sanford 1977, 1978; Sanford and Garrod 1978, 1981) to investigate the processes by which the referents for input text are established. Like Kintsch (1979) their primary aim was to create a model which would account for the establishment of coherence when the referent was implicit in the text. Their initial starting point was a re-examination of the assumptions and implications of the existing referential models (Haviland and Clark, 1974; Kintsch, 1974). These models assume that explicit prior presentation of referents is essential for coherence and hence ease of comprehension. They also assume that failure to locate a referent in the text which has been processed immediately prior to the input will lead to searches of stored memory of the text to locate the referent. Failure to do this will lead to a chain or bridge of inferences being established which will link the input to the processed text. In cases where such a bridge cannot be built a new coherence graph is created (as described by Kintsch, 1974) or a different topic is assumed (as described by Haviland and Clark, 1974). The implications of the models are that if an explicit referent is not available, then time taken to process the input will be greater than if it is (the time being taken by the search and by the inference making [Kintsch] or bridging [Haviland and Clark]).

The model provides an alternative to text-driven bridging as an explanation of how input text locates implicit referents. In terms of its functions, the model is comparable to that of Kintsch (1979). Both propose a

frame like structure which is activated by information in the input text. Once activated the frame is then available to accommodate new input. Implicit information in the text is available as slots in the frame. New text can be related to the available slots as it is input. One difference in the models is that in the Sanford and Garrod model explicit information is a part of the same frame or Scenario whereas the Kintsch (1979) model incorporates two separate parallel processes for the construction of a coherence graph and a fact structure. A further difference in the models lies in the level of detail which describes their activation and function. As stated, the Kintsch model is largely intuitive and descriptive. The Sanford and Garrod model is supported by experimental investigation. This took place in progressive stages which will be described in some detail.

#### The development of the Scenario model.

##### a. Bridging as a process initiated by the linguistic knowledge.

Garrod and Sanford (1977) challenged the basis of the Haviland and Clark (1974) theory of bridging. Bridging, according to Haviland and Clark, is a class of inference which links information in an input of text to information that has been input earlier. It occurs when there is no explicit referent for the concept that is being input, but there are linguistic indications (such as pronominalisation or the use of a definite article) that the input is related to previous input. According to this theory, all bridges are made on the basis of linguistic knowledge.

Haviland and Clark's experimental work relied on the Given-New Contract theory (Halliday, 1967; Grice, 1975). The theory assumes that the speaker and listener have a contract to make every message meaningful. According to the contract the speaker's obligation is to relate each new input of information to information which is already established, or which he believes that the listener already knows, so that each sentence contains some information that

is new and some information which links it to that which has gone before (or has been given). The listener's obligation is to distinguish between the input which introduces new information and input which refers to established or "given" information and to link the given information to previously processed discourse. One of the ways in which given information is distinguished from new information is by the use of definite or indefinite article. The use of a definite article "the" indicates that the next word in the input is given information: the use of the indefinite article "a" or "an" indicates that the next word is new information. Pronominalisation of an input also indicates that a referent has previously been established. But as the previous sentence demonstrates, the theory ignores the additional functions of articles particularly that the indefinite article introduces a general concept and that the definite article is used to introduce specific concepts. The examples chosen by Garrod and Sanford are not affected by other usages of the articles, but the use of the rule as if it were unambiguous does weaken the argument for the exclusive use of definite and indefinite articles as referents.

It was hypothesised by Haviland and Clark (1974) that if no explicit referent was given for an input then the reader/listener would construct a bridge or chain of inferences to link the information which was given (indicated by a definite article) to that which had gone before. Haviland and Clark (1974) carried out two experiments to investigate whether listeners/readers did this. Their hypothesis was that if readers did make bridges to link the given information in an input to existing information then processing time would be longer for sentences which did not have an explicit antecedent referent than for sentences which did. Their method was to present the same sentence to two groups of subjects under the two conditions (with explicit and with implicit antecedents referents). One sentence presented was:

"The beer was warm."

The use of the definite article "the" is an indication from the speaker to the

listener that "beer" is already given. When the antecedent sentence was:

"Mary unpacked the beer"

it contained an explicit referent for the input. The results showed that the target sentence took significantly less time to comprehend in this case than when it was preceded by the antecedent sentence:

"Mary unpacked the picnic things."

in which the referent was not explicit. This, along with a similar experiment which controlled for the priming effect of presenting the actual word twice, supported their hypothesis and gave strength to the assumptions that bridging inferences must be made to establish the referent for given information and that establishing an explicitly coherent link was the process by which this was achieved.

Two assumptions are implicit in this experiment. One is the notion that the bridging process is initiated by linguistic features in the input and the other is that the process only takes place as and when it is required. Garrod and Sanford (1977) challenged these assumptions.

#### b. Semantic Bridging (Garrod and Sanford, 1977)

Garrod and Sanford (1977) showed in a series of timed reading experiments that under some circumstances there was no significant difference between the time taken to process new information in an input (indicated by the use of an indefinite article) and given information (indicated by the use of a definite article) for which a referent had to be created by the use of an inferential bridge. This effect occurred when words from the same semantic category were used in the antecedent sentence and in the target sentence. The experiment suggested that the Given-New strategy may not be the only one which the processor was using. The following example will make this clear. The sentence:

"A pedestrian was killed by the vehicle"

would indicate that "vehicle" had previously been explicitly or implicitly

given. The preceding sentence:

"A bus came trundling down the hill"

would allow a bridge to be made from "bus" to "vehicle". Garrod and Sanford found however that similar processing time was taken when:

"A bus came trundling down the hill "

was followed by

"It nearly smashed into a vehicle."

"Bus" is explicitly the referent for "it" and therefore if the process of establishing linguistic referents is used the sentence should involve less processing time than the sentence in which the "bus-vehicle" bridge needs to be made. Garrod and Sanford attribute the processing time to a referential search which was initiated by the new semantic input of "vehicle". In other words a false search was set up for a referent on the basis of the semanticity of the words and not on the basis of explicit antecedent referents.

As a result of this (and other research including that which showed that the phrase "a horse-drawn vehicle" did not produce the same effect as the phrase "a vehicle") Garrod and Sanford concluded that there may be a context driven, semantically controlled bridging process which operates alongside the Given-New operation and that the reader goes through the initial stages of semantic bridging even in the presence of the Given-New syntactic cues which establish referential links.

c. The decomposition of input into primitive meaning (Sanford and Garrod, 1980).

To account for the phenomenon of semantic bridging, Sanford and Garrod (1980) proposed that the input may not in fact be encoded in the form in which it is initially presented (with the addition of the necessary referential bridges), as is implied by the Kintsch (1974) and Haviland and Clark (1974). Instead they proposed in accordance with Schank and Abelson's (1977) theory, that the text is decomposed on the basis of its verb content and stored as

primitive acts. Primitive acts are discussed in Chapter Four of this review, but their importance in this context is this: as a verb input is decomposed into primitives, the meaning is captured and restrictions on semantically acceptable selections are specified. This in effect sets up a framework which allows further input to be parsed into a meaning structure. If the text is decomposed in this way, then the meaning structure will provide a slot for implicit referents in the same way as Kintsch (1979) suggests. If this meaning structure is created at input, then it will be possible to make semantic inferences as quickly as explicit referential connections.

Sanford and Garrod (1980) describe an experimental investigation of this. If the initial input is decomposed in the way which is suggested, then the sentence:

"Mary dressed the baby."

would have the concept of clothing associated with it and a conceptually driven bridge for the input of "clothes" from the text will be created. If this is so then the implicit reference to clothing should be as acceptable to the processor as one that is explicitly stated, as for example in the sentence:

"Mary put the baby's clothes on."

To test whether this was the case, Sanford and Garrod (1980) gave subjects a self paced reading task during which time to read the target sentence:

"The clothes were made of pink wool."

was recorded. If the conceptually driven bridge was set up which would anticipate the input of clothing then the time taken to read the target sentence should be similar whichever of the priming sentences preceded it. If however text-driven bridging was initiated, then the target sentence should take longer to process when

"Mary dressed the baby ."

preceded it because the bridge would have to be constructed at the time when

"clothes " was input.

Clark and Haviland had assumed from their results that the bridging process would take two hundred millisecond to operate. Sanford and Garrod found a non-significant difference of only seven milliseconds between the conditions. They concluded that this was insufficient time for a text driven bridging operation to be carried out and that a conceptually driven bridging process (initiated by the decomposition of the semantic content of the first sentence) was carried out prior to the input of the target sentence.

#### d. The operations of the conceptually driven framework.

Sanford and Garrod (1980) next addressed the problem of how the conceptually driven framework might operate on the input. Two possibilities exist. One is, as Minsky's (1975 ) model of Frames proposed, that when the framework is activated it will have terminal nodes which are supplied with default values. The default values will supply information about what is normally associated with the situation depicted in the frame. These values will exist except when it is explicit in the text that some other value fills the slot.

The other possibility is that the framework exists in relation to a decomposed verb as is stated by Schank and Abelson's Conceptual Dependency Theory (see Chapter Four). In this case slots for explicit input values will be created by the constraints which are imposed by the primitive meaning of the verb. This would mean that the slots would remain empty unless a suitable value is explicitly supplied by the input.

If the Frame supplies default values, then the values will be instantiated as a part of the frame even though they are not explicitly referred to in the input. If the framework supplies only constraints on the values, then the frame will have only an interpretive function which will incorporate explicit input.



If the conceptually driven process of the Sanford and Garrod model utilises a framework which has an interpretive function (that is if it accords with the Schank and Abelson, 1977 model) it will rely on explicitly stated semantic values to fill the slots in the representation. If this is the case and there is an explicit input of "clothes" (as in the sentence "Mary put the baby's clothes on") followed by a reference to "the material" it will be possible (according to Sanford and Garrod) to relate "the material" to "clothes" (they do not specify the nature of this link). But if there is no explicit mention of clothes (as in "Mary dressed the baby") there will be no concept to which "the material" can be related. In this case the input

"The material was made of pink wool."

would require a semantic bridge to be made at the time of processing. This would involve two stages of processing:

- 1.The baby was dressed in specific clothes;
- 2.The clothes are made of material.

If "clothes " has been explicitly mentioned as an antecedent, then only stage two of the bridging operation needs to be carried out.

Sanford and Garrod (1980) tested the reading time for the target sentence with the explicit and nonexplicit antecedent sentences. They found that reading time for the target sentence when presented with the explicit antecedent was quicker on average by seventy milliseconds. From this result, it was concluded that an explicit input was needed to provide values for the slots in the the framework and that default values were not supplied. This supported the Schank and Abelson model and it was concluded that conceptually driven text representations depended on the decomposition of input at encoding.

### The Scenario Model.

From the preceding four stages of investigation Sanford and Garrod have concluded that the basic model or Scenario by which a text is comprehended is a framework which is created by the decomposition of the initial input into a semantic primitive meaning. The slots in the framework are filled by explicit inputs from the text, and the possible inputs which are acceptable are constrained by the primitive meanings of the decomposed verbs.

A scenario can be described in the following way. Input units from the text are phrases or sentences. The initial input is important because the decomposition of the verb into its primitive meaning forms the initial frame or scenario. In addition, the entity in the initial noun phrase is the topic of the scenario. This is in accordance with the theory of Lyons (1968) and supported by the experimental work of Hornby (1972). Hornby demonstrated that subjects chose pictures which featured the first mentioned entity in the input, as being those that the story was about (for example, subjects given the sentence:

"The cat was stroked by the boy "

selected a picture of a cat as being what the story was about. Subjects who were given the sentence

"The boy stroked the cat"

selected a picture of a boy).

#### a. Primary Processing.

##### (i). Verb decomposition.

If the initial input is "Mary dressed the baby." first of all the verb "dressed" is decomposed into its primitive meaning. Sanford and Garrod do not describe in detail this stage of the process, but in accordance with the theory of Schank and Abelson (1977), the verb "dressed" can be decomposed into two ATRANS primitives as follows:

ATRANS something to someone;

ATRANS something from someone.

Sanford and Garrod use the terms acts, roles and props. These would seem to map onto the primitives terminology: ATRANS = acts; someone = role; something = props (although again this is not made explicit by Sanford and Garrod). The primitive meaning of dressed could be represented thus:

Acts: ATRANS 1  
 ATRANS 2  
 Roles: 1. Transferred from  
 2. Transferred to  
 Props: Thing transferred.

(ii). Topic identification.

After the verb has been decomposed the topic is identified. In the given sentence, the first mentioned entity is Mary. MARY is therefore identified as being the topic of the input.

(iii). Mapping of the explicit input onto the framework.

Next the explicit input is mapped onto the framework to form the Scenario.

MARY.....Role 1.....>

Dressed ATRANS

baby .....role 2.....>

Those slots which are not filled with explicit input (in this example: Props = thing transferred) are also a part of the Scenario. Sanford and Garrod distinguish the two areas as Explicit Focus (the filled slots) and Implicit Focus (the anticipatory, unfilled slots). The term "Focus" is used in preference to the equivalent "Foregrounding" (Chafe 1972) because Sanford and Garrod conceptualise the Focus as the backdrop for subsequent input and in this way seek to avoid confusion. The next input may map directly onto the

current scenario. This presumably will be the case if the input was:

"She put his clothes on."

She will map onto Role 1 (in explicit focus);

put on will map onto ATRANS;

his will map onto baby;

clothes will map onto prop (in implicit focus).

"Clothes" will now be moved from implicit to explicit focus.

This kind of processing in which input maps directly onto the Scenario which is in focus is Primary processing (Sanford and Garrod, 1981). Secondary processing occurs when direct mapping cannot take place.

#### b. Secondary processing.

Secondary processing involves carrying out a bridging operation or initiating a new Scenario. If a bridging operation takes place then the original Scenario remains in focus. It is assumed from the experimental work (although this is not made explicit in the model) that a semantic bridging operation will be initiated if a further sentence such as:

"The material was made of pink wool."

is input. There are no slots for the role, the action or the props in the current Scenario. Although a text driven bridge might be

"Clothes are made of material."

it is not clear how this bridge can be made within the Scenario model.

If the secondary processing involves the creation of a new Scenario then the current Scenario moves out of focus and a change of focus occurs. The change of focus is signalled in two ways.

#### (i). The input can not map onto the framework of the Scenario.

In this case both topic and Scenario dependent roles move out of focus if inputs do not refer to them. Topic roles are more quickly restored to focus than Scenario dependents if they are subsequently introduced either by a noun or pronominal reference. (This was demonstrated by Purkiss (1978) who showed

that reading time for a target sentence [which contained noun or pronominal references to topic or Scenario dependent roles] increased as a function of the number of intervening sentences which had no reference to the roles. The increase in reading time was attributed to the extent to which the Scenario had moved out of focus).

(ii). There is a time change described in the text and the expected time length of the Scenario is violated.

The reality of this signal has been investigated by Anderson (1980) who worked closely with Sanford and Garrod. She ascertained the maximum time associated with a number of activities (for example, having a haircut or seeing a film ) by asking a number of judges to estimate the normal duration time. Scenarios with a topic role ( which was independent of the Scenario but filled a role in it) and a Scenario dependent character were created. A time change which was within the time range for the Scenario and one which was well beyond the maximum range were prepared as alternative final sentences for the input. Two target sentences were devised, one which referred to the topic character and one which referred to the scenario bound character. Reading time for the target sentence which featured the Scenario dependent role was longer when it was preceded by the input which violated the time span than when the antecedent sentence time span was within the normal range. The reading time for target sentences about the topic role was not affected by the violation of Scenario time span. The conclusions drawn from this experiment are that violation of the normal time span moves Scenario dependent roles out of focus and that to recover them requires a further process. If the roles have moved out of focus, then it can be inferred that the Scenario on which they are dependent is also out of focus. When the focus changes in this way it is only the Scenario which moves out of focus. Topic roles remain in Focus awaiting further input.

### c. Storage and Long Term Representations.

As entities and Scenarios move out of focus they are encoded in a long term representation of the text. This representation is Scenario based, the inferences having been derived from the framework rather than constructed as a text driven bridge. This means that the reader now has a meaning based memory of the input. Entities and Scenarios which have remained in focus or been called back into focus for a large number of inputs have a stronger representation in memory and are therefore more likely to be recalled than entities and Scenarios which have been in focus for less time. The product of the processing is an individual memory structure for the text: that is, a memory of what the story is about. The structure will have a hierarchy of scenarios and roles. Over time this memory will gradually decay, but the meaning represented in the scenarios will become a part of the long term semantic memory which will be stored separately partitioned from the particular individual memory of the story. This long term semantic representation will remain.

Sanford and Garrod proposed that the stored representation of the text would be hierarchical. The hierarchy would be created by processing. Those scenarios which had been in focus often (either by remaining in focus or by moving back into focus) would be the most important in the hierarchy. These would most easily be recalled.

Being "in focus" depends on the initial decomposition of the sentence or phrase into a primitive meaning. If it is the first input it will determine what values will be acceptable for the Scenario. The relevance of subsequent input to these values will determine how long or how often the particular Scenario is in focus. Input which is subsequently processed will either fit into the existing Scenario or create another Scenario. The length of time any Scenario will remain in focus will then depend on how relevant it is to the text.

A comparison of the Kintsch (1979) model with the Scenario model.

Sanford and Garrod's (1981) Scenario model is an analysis of text which depends upon the decomposition of the input into deep semantic meanings of acts so that referential coherence can be established. The authors have compared their model with linguistic models which are based on the referential coherence established by the usage of pronouns and of definite and indefinite articles which precede nouns. Their experimental work has established that in some situations readers do use processes which are based on semantic meaning rather than on linguistic features.

The model can also be compared with the Kintsch (1979) model which was described in the previous chapter. Both models are semantic models and in both analysis is centred on the verb content of the input (the rest of the input being related to the meaning defined by the verb). There are however, important differences:-

1. Kintsch's (1979) model describes two parallel processes: one which establishes the semantic, referential coherence of the text and one which extracts the meaning. In the Sanford and Garrod model these two procedures are combined into one operation. This is possible because the Scenario model decomposes the input into primitive meaning and coherence is established from this base.
2. Kintsch is concerned about the effect of knowledge of the overall structure of texts on what is stored. His model takes this knowledge into account and describes a third process which operates under the constraints of knowledge about text structure. The Sanford and Garrod model takes no account of such constraints and focusses only on the transformation of incoming text with no overall framework other than that created by the decomposed input.
3. According to Kintsch's model, all the concepts in each propositional input are related to the verb or predicate of the proposition. In theory

the Sanford and Garrod model incorporates only those concepts which are compatible with the primitive meaning of the verb. The model is only concerned with those inputs which fit into the anticipatory slots.

Both models lack a mechanism which would incorporate novel uses of words, that is figurative use of language such as metaphor, irony or satire. The models are also inadequate to account for the significance of the choice of particular words in particular sentences. For example, the use of "pink wool" (earlier in this chapter) to describe the material of the baby's clothes is only significant because it is a suitable description of material. Lemon silk, white cotton, blue nylon, would be equally suitable just as grey cobweb would not. On the other hand the choice of "pink" may well have rather more significance for the reader who is concerned about the sex of the baby. Neither the Kintsch (1979) nor the Sanford and Garrod (1981) models make provision for such speculations as they are incidental to the main Scenario or Frame.

The next chapter is concerned with models which examine these more pragmatic aspects of comprehension.



CHAPTER EIGHT.Elaborations: Associations and Constructions.Introduction

The models and theories which have been discussed up to this point have primarily been concerned with the underlying structure of texts and the way in which text is encoded and retrieved in terms of these structures.

Models of semantic networks (see Chapter Three) are only concerned with the surface representation of a text. They see the input text as being encoded as a sequence of propositions corresponding in a one-to-one manner with the surface structure of the text. The more recent models (Norman, Rumelhart and the LNR group, 1975; Loftus and Collins, 1975; Anderson, 1976) provided more sophisticated and complex procedures for the storage and access of input. However the models provide only for the storage, accessing and reconstruction of the specific input and do not approach the question of the spontaneous activation of information from beyond the input.

Similarly, the Smith, Shoben and Rips (1974) feature extraction model (which differed from the network model in that it did not initiate a search of the stored associations but instead computed the similarity of the features of two concepts) acted only on the given concepts and the attributes which were stored with them and did not produce information about concepts other than those that were input.

These models can only work on given concepts and their stored associations. Knowledge is limited to associations (or properties, or attributes) which are already associated with the input concept. The procedure matches input concepts to stored concepts and in so doing activates the associated concepts. These models assume that meaning is wholly carried in the text.

Kintsch's (1974) model constructs a coherent representation of the surface form of larger units of text and Kintsch and van Dijk's (1978) macrostructure model is produced from this representation by the application of macrorules which operate in relation to a stored knowledge about text structure. Story grammars allow all simple stories to be described as having a comparable underlying syntactic structure regardless of the actual content of each individual story.

These models (despite their differences in approach) are concerned with the process by which coherently joined units of text (the unit being larger than a single concept) are transformed into a global representation which is a typical rather than a unique model of text. In general the Kintsch and van Dijk (1978) model is concerned with the process by which a representation is formed and the story grammars are concerned with the structure by which the story is represented.

Input material which does not fit these structures is largely ignored. The models are only concerned with the encoding and retrieval of content which maps onto the structures and are the meaning that the structures impose on the text.

The recent models which have been discussed (Kintsch, 1979 and Sanford and Garrod, 1981) are more concerned with ways in which specific and more complex texts are incorporated into the reader's existing knowledge. Bartlett (1932) called this process "effort after meaning" and "simply the attempt to connect something that is given with something other than itself." (p.227). The two models (Sanford and Garrod, 1981 and Kintsch, 1979) view effort after meaning through differing theoretical frameworks but both are concerned with the way in which the information in the input interacts with pre-existing knowledge to form a deep semantic representation of the input.

Kintsch's (1979) Fact Structure attempts to create a single frame of the whole text (comparable to Minsky's, 1975, frames). This involves ongoing

revision of the frame to encompass new facts and their associated concepts. Underlying Kintsch's model is the assumption that all concepts have pre-existing associations which are activated when the concepts are encountered in the text. These associations furnish default values for the framework, and along with the input are responsible for the extension or revision of the framework as the text is input and the existing framework becomes inadequate to encompass the new values.

Sanford and Garrod's (1981) Scenario model builds up a representation by creating a series of Scenarios which are linked together by means of the topic to provide a complete representation. As has been discussed in the previous chapter this model rests on the assumption that a framework or Scenario is created from the relationships between concepts in the initial input. The decomposition of the verb into its particular primitive meaning depends on the way it is related to certain concepts in that specific context and not on a general pre-existing meaning of the verb only. Nonetheless it is limited and constrains what might be comprehended from the input.

This fundamental difference in assumptions (that there is a 'meaning' that is dependent on pre-existing associations [Kintsch,1979] or that meaning is context dependent [Sanford and Garrod, 1981]) is similar to the basic difference which separates models of elaborative processing. In general, models of elaborative processing (some of which are to be described below) propose that concepts in the input are elaborated by a process whose function is to aid the retention of the verbatim message. However some of these models ( Craik and Tulving, 1975; Reder, 1976, 1979; Anderson and Reder, 1979) assume that there is a general elaborative processing of all concepts as they are encountered in the input and that the elaborative process is one of reactivation of pre-existing associations. Other models (Stein, Morris and Bransford, 1978; Stein and Bransford, 1979; Masson, 1979) assume that elaborative processing is a specific process which constructs associations

between particular concepts in the input: associations are constructed that are meaningful within the context of the input in which the concepts are being used and may be created for that particular input.

All the elaborative models are concerned with the retention of input, but all posit the use of information which is not presented in the text (either pre-existing or specifically constructed) as an aid to the retention of the surface form of the input. Stein et al and Stein and Bransford propose that in addition to embedding the input into the memory, elaborative processing anticipates future possible inputs because it questions the potential significance of inputs. The approach expounded by Stein and his associates has developed from prior work by themselves and others on the construction of meaning.

#### The construction of meaning.

A theory of the construction of meaning has been proposed and developed by Bransford and his associates (Bransford and Franks, 1971; Bransford, Barclay and Franks, 1972; Bransford and Johnson, 1972; Johnson, Bransford and Solomon, 1973; Bransford and Johnson, 1973; Bransford and McCarrell, 1977). Bransford defines the constructive process as one in which something new is constructed out of components of information acquired at different times. In addition to being constructive and creating aids for the retention of text, the process is productive and creates meaning and although it is initiated by a linguistic input it is linked to non-linguistic elements of memory.

In their initial work in the area Bransford and Franks (1971) proposed that subjects do not merely retain the information in each proposition or sentence of an input and recover this verbatim in a recognition task. Rather they spontaneously integrate information from semantically related acquisition sentences, and base their recognition accuracy and confidence on the wholistic, integrated representation rather than on the exact sentences which are input.

They demonstrated this by giving subjects a recognition task in which they were asked to listen to a text and then judge whether presented sentences had formed a part of the original text. The experimenters restricted their recall stimuli to sentences that had been a part of the text or could only be inferred from an integration of the the sentences of the input. For example, subjects heard sentences which told them separately that jelly is sweet, and that ants ate the jelly. Subsequently subjects falsely recognised a sentence such as 'the ants ate the sweet jelly 'as one which they had formed a part of the text.

The recognition of such sentences allowed the authors to conclude that information is retained in an integrated form, rather than as linked but separate sentences (as the semantic network models propose). They next proposed (Bransford, Barclay and Franks 1972) that although linguistic information might be used to remember the input it is not sufficient to explain what is retained. According to the authors

"...sentences are not viewed as linguistic objects to be remembered.

Instead they are viewed as information which subjects can use to construct semantic descriptions of situations." (p.194).

They proposed that the constructed description which is based on a linguistic input contains more information than that which is in the linguistic input.

Using the false recognition paradigm from the 1970 integration experiment, Bransford et al (1972) demonstrated that subjects construct a description which integrates their knowledge of spatial information with the linguistic input. They presented two groups of subjects with sets of sentences. The sentences given differed in the same way as the two given in the following example:

- A. "Three turtles rested on a floating log, and a fish swam beneath them."
- B. "Three turtles rested beside a floating log, and a fish swam beneath them."

Bransford et al hypothesised that subjects who were given sentence A would falsely recognise the following as the sentence which had been presented initially:

"Three turtles rested on a floating log and a fish swam beneath it."

Subjects who were given sentence B would not however recognise the sentence (though they would detect the pronoun change).

The experimental results confirmed this hypothesis and the authors concluded that subjects generate and store a description of the situation rather than store only the linguistic input. Spatial relationships which can be inferred from the information in the input, but which are not a part of the information given linguistically, form a part of the retained construction. The authors carried out two connected experiments to support the assertion that subjects base their recognition on the description of spatial relationships which they construct from a given input.

In a similar experimental task to that described, subjects failed to recognise conceptually similar sentences which violated the implied spatial relationships, but falsely recognised sentences which were conceptually different, but did not violate the given spatial relationships. These further experiments give additional support to the hypothesis that understanding meaning involves more than remembering linguistic input.

Bransford and Johnson (1972) focussed their experiment on the subjects' need to have more information available than that provided in the linguistic input. They demonstrated that linguistic input in which the lexical and syntactic structure were familiar, was not comprehensible or memorable unless related to a context. Subjects in five experimental conditions were given identical texts to learn for recall and to rate for comprehensibility on a

seven point scale. The five conditions were:

1. The passage alone (without context);
2. The passage with a context preceding the presentation;
3. The passage with the context after presentation;
4. The passage with a partial context preceding it;
5. The passage read twice but with no context.

The context was a picture that provided a informational base for the concepts in the text.

The results of the experiment showed that when the context was given before the text (Condition Two) subjects recalled more idea units and rated the comprehensibility of the passage higher than subjects in the other four conditions. The main area of interest was the difference in retention and comprehension between Conditions One and Two in which there was context or no context. The other conditions provided controls for other variables. For example, subjects who were given the partial context had retrieval cues for the concepts in the input, but they were not aware of the relationship between the concepts in the particular context. Their comparatively worse recall and comprehension of the passage demonstrated further that it is not the association between the concepts in the input and the context (such as a semantic network would provide) which makes the passage comprehensible and memorable but the links constructed on the basis of the relationships between the concepts in the particular input.

In the task described above, subjects needed a very specific context in order to comprehend and remember the text. The context described an extraordinary situation which subjects were unlikely to have encountered previously and therefore would be unlikely to have constructed without the context. It could therefore be argued that the results of the experiment were due to relevant associations being created by the concepts in the context (even though the results for the partial context condition argued against

this). Bransford and Johnson repeated the experiment using texts which described familiar situations about which most subjects would have pre-experimental knowledge. In two experiments they tested the effects of presenting a descriptive topic before and after a passage which described procedures for doing the laundry. In one experiment the passage of text was read to the subjects, in the other it was presented in written form. The results of both experiments demonstrated that knowledge of the topic prior to hearing/reading the passage increased both the comprehension (as rated) and the recall of idea units. From this it was concluded that prior knowledge or pre-existing associations per se do not increase the comprehension and recall of a passage, but that the appropriate knowledge must be activated beforehand to allow the reader to construct an appropriate situation from the input. The role of the topic is to allow the reader to create an appropriate context.

Johnson, Bransford and Solomon (1973) investigated further the kinds of construction (in addition to spatial descriptions) that readers or listeners make in order to create a context. On the basis of work by Fillmore (1968) and Kintsch (1972), they proposed that subjects would construct a description which included an instrument inferred by a verb. For example, if subjects were told that "John shot the intruder", they would infer and therefore falsely recognise the statement "John used a gun". They used a false recognition paradigm to investigate this hypothesis.

An experimental group were given twenty sentences each embedded in a short context. Eight of these contained verbs which would allow the subject to make inferences on the basis of the instrument which the verb implied. The other sentences were eight filler sentences and six which would allow the subjects to make inferences on the basis of spatial relationships. One of the eight "instrument" sentences was: "He was pounding a nail". A control group was given texts in which only the key verb was changed (or in the "spatial" sentences, the preposition was changed as in the turtle experiment) so that



the control group would read the sentence: " He was looking for a nail."

The subjects were told that they would be asked questions about the texts they were about to be given. They were then allowed to listen to the texts with a two second interval between each. They were allowed a three minute interval and then required to write yes/no in response to a set of statements ('Yes' meaning that they had already heard the statements in the text). The results showed that the experimental subjects falsely recognised the inferred statements in both the "instrumental" and "spatial" texts more often than the subjects in the control condition. The difference was statistically highly significant. Bransford and Johnson concluded that the results further supported the idea that subjects construct meaning using the linguistic input rather than that the linguistic input itself contains the meaning. This series of experiments demonstrated that very small changes in the linguistic input, can result in large changes in the meaning communicated.

Bransford and Johnson (1973) further supported their view on the construction of meaning by quoting a series of unpublished experiments, which they and their associates had carried out. One of these was an exploratory study carried out by Bransford and McCarrell (see Bransford and Johnson ,1973) which investigated (in a false recognition paradigm) the differences in construction when subjects were given different conjunctions to join two sentences. Differences in recognition were scored between subjects who were given a recognition task following the acquisition of "John missed the bus because/so he knew he would have to walk to school". Subjects in the "because" condition falsely recognised "John wanted to walk to school" more often than subjects in the "so" condition. These results suggest that subjects construct meaning on the basis of pragmatic knowledge when other sources of knowledge are not appropriate.

This exploration led Bransford and McCarrell to investigate further the creation of contexts in "because" conditions. They investigated the idea that

the creation of a context on a pragmatic basis might be hard or easy depending on the pre-existing relations between the concepts in the two joined sentences. They proposed that a sentence such as:

"The account was low because Sally went to the bank"

would be easier to comprehend than a sentence such as:

"The haystack was important because the cloth ripped".

They found that subjects, given the topic of the sentence as a cue were able to recall the "easy" sentences more often than the "hard " sentences. However a control group who were given context cues ('withdrawal' preceeded the "easy" statement in the example and 'parachute' preceeded the "hard" example) showed no difference in recall. From these results they concluded that it is the construction of a suitable context that makes the content hard or easy to learn and if a context is created, then the difference between "hard " and "easy" disappears, regardless of pre-existing associations.

Cues as aids to recall.

Doll, Lapinsky, Bransford and Johnson (see Bransford and Johnson, 1973) demonstrated that relevant cues aid recall, but irrelevant cues lead to less recall than no cue. From this they argue that when no cues are given subjects attempt to create a context for the input (with varying success) but the provision of irrelevant cues suppresses this and the context that created from the irrelevant cue retards the comprehension process.

Fenrick (see Bransford and Johnson, 1973) carried out a pilot study which investigated the use of cues after recall. The purpose of the investigation was to see whether the constructed meaning could be restructured when given cues suggested that a new context could accomodate the information. Two groups of subjects were given the same text. The text content centred round a trip to the countryside. The "topic before" group were told that the passage was about an escaped convict. The other group were not given a topic. Afterwards both groups were asked to recall the passage. Following recall, both groups were

given cues which emphasised certain aspects of the text (these were sections of the text which could emphasise the 'convict' aspect, but which without a convict context were not particularly relevant). Subjects were then asked to recall anything more about the text which was not included in the original recall. They were told that the cues would help them in this way. Those who had received the topic before added nothing. Subjects who had not received a topic had created a recall which centered around a trip into the countryside. Of these 56% changed their interpretation and reconstructed their story to fit an 'escaping' framework. Bransford and Johnson interpret these results as further support for their view of the construction of meaning. They suggest that further processing can occur and that meaning can be restructured in a compatible context. They contrast this result with the picture context results of the Bransford and Johnson (1972) experiment which showed very poor recall of material when the context was presented afterwards. They argue that the story given in the Fenrick experiment allowed a context to be constructed and that this could be used as a framework for the previously unrecalled information. In this way it provided a structure for the recall of the input which became newly relevant. In the picture context, the text alone had not permitted the construction of a framework and therefore there was no foundation for the recall of newly relevant information.

Bransford and Johnson (1973) conclude that subjects do not merely integrate presented information but that they spontaneously make assumptions about extra linguistic circumstances and draw on a wide range of prior knowledge to do so. The experimental work (both that which has been published separately and that which is quoted in Bransford and Johnson, 1973) clearly supports this conclusion.

Other research which has investigated the construction of meaning.

Research other than that carried out by Bransford and his associates has supported the theory of construction and restructuring of meaning in accordance with a given or experiential context.

Sulin and Doolin (1974) gave subjects brief biographical passages. In one condition the main character in the text was given the name of a well known person (for example Helen Keller). In the other condition the main character was given a fictitious name. In all other respects the texts were identical. In the recognition task which followed subjects in the "famous person" condition falsely recognised well known facts about the character (for example, in the case of Helen Keller: "she was deaf, dumb and blind"). Subjects in the "fictitious " condition were less likely to recognise such statements.

Anderson, Reynolds, Goetz and Schallert (1976) used ambiguous passages to show the effect of the reader's interest and experience on the interpretation of a text. They read an ambiguous passage which could be interpreted as describing either a card game or a musical quartet. The authors found that students of music more often recalled the passage as a musical quartet than did physical education students.

Brown, Smiley, Day, Townsend, and Lawton (1977) told children a story about an escape. One group were told that the story was about a convict, the other group were told that the story was about a character from the then current television serial "The Planet of the Apes ". Children in the second condition falsely recognised incidents from episodes of "Planet of the Apes" as being part of the text whereas children in the "convict" group did this less often.

Pichert and Anderson (1977) demonstrated the effect of the reader's perspective on the recall and rated importance of idea units in a text. Subjects were given a perspective by the experimenters. For example, they were

told to read a passage of text either as if they were a home buyer, or as if they were a potential burglar. Subjects read identical texts. The text related to the example was about two boys who played truant from school and went to play at the home of one of the boys. It contained information which would be of interest to a homebuyer (about a musty basement or a fine fireplace or damp) and information that would be useful to a potential burglar (the location of a television, or of expensive bicycles). Following the reading the subjects rated the idea units in the passage on a five point scale of importance. There were low correlations between the group ratings which supports the idea that each group constructed a different meaning for the passage. Subsequent recall of the passages showed that subjects recalled the ideas which were relevant to the perspective that they were given. Significant differences were found between the groups in terms of the idea units recalled from the passage.

#### Reconstruction of meaning.

Anderson and Pichert (1978) supported the results of Fenrick's pilot study (Bransford and Johnson, 1973) when they demonstrated that the constructed meaning of a passage could be restructured in response to subsequent input. In their experiment, subjects read and recalled stories in the same perspective conditions as in Pichert and Anderson's (1977) experiment. After recalling the passage, half the subjects in each condition were told of the alternative perspective. Following this all subjects were asked to recall the story once more and to include all the information which they could remember. Those subjects given the 'new' perspective were able to recall information that was relevant to their new perspective and which they had failed to recall in their initial attempt. This was more information than that recalled by the control group who made a second recall but were not given the 'new' perspective. These results support the idea that meaning can be restructured to incorporate given information which was irrelevant when the

text was read, but that the information will not be recalled without a perspective shift.

These results were also supported by Spiro (1980) who carried out a complex experiment which involved placing subjects in one of thirty-two conditions. Only the conditions and issues that are relevant to the current discussion will be described. Subjects were given a text to read which had two alternative endings. The core story concerned an engaged couple. In the story the man was anxious to discuss the fact that he did not wish to ever have children. The alternative endings were that the woman was delighted or horrified. Later some subjects were given additional verbal information as they were leaving. This was either that the couple married or that they didn't marry after all. This information contradicted the expectations of some of the subjects. For example, those who had been told that the woman was delighted would not expect the couple not to marry: those who had been told that the woman was horrified would not expect that the couple would marry. After variable time delays, which ranged between two days and six weeks, the subjects were asked to recall the written text. Those who were given the contradictory information incorporated this into the text in such statements as "although they were in agreement about babies, there were other things that they did not agree upon". Or, "They separated, but realised after discussing the matter that their love mattered more."

Spiro concludes that memory of information which has been received in the past is restructured at recall to accord with the subject's current cognitive state, and that distortions of the information may be caused in the attempt to produce an accurate account. When this happens, the distortions cannot be distinguished from accurate memories.

The reconstructive view of remembering information is of an active process in which specific memories and knowledge of the world are reconstructed to produce a memory which accords with more recent information,

and (as suggested by Spiro) the reader's current cognitive state. According to this view, memory structures are flexible and are built or restructured for the situations in which they are required.

Constructive and reconstructive models of memory differ from the association models of memory such as Reder's (1976) Elaboration model (to be discussed in the next section of this chapter) in which remembering is a process involving the retrieval of stored traces. They also differ from the concept of rigid knowledge structures (such as Schank and Abelson's, 1977 scripts) which would need an infinity of stored contexts for their storage.

The flexible structures which are proposed by this model (and incidentally, also by Schank, 1981) require less storage and permit greater flexibility in their storage. In addition they allow a way of relating comprehension processes to the reader's needs and purposes.

### Elaborations

The final part of this chapter will examine the differences between elaborations when they are described by "association of meaning" models and as "construction of meaning" models. Briefly, the difference is that the association models view meaning as an analysis of the text whereas the construction models take the view that meaning is created by the reader from the text. This difference was initially discussed in Chapter One when Educational perspectives were reviewed.

Reder (1976) in particular is an exponent of the elaborations as association view. She makes a clear distinction between elaborations and inferences in terms of the kind of information that they supply in the creation of a text representation.

Stein's work (Stein, Morris and Bransford, 1978; Stein and Bransford, 1979) supports the elaborations as "construction of meaning" view. This particular view makes no overt distinction between elaborations and

inferences. Elaborations would seem to be a specific type of inference which which are created as links between specific concepts in a sentence.

#### Elaborations as semantic associations.

Craik and Lockhart (1972) proposed that the retention of a linguistic input depended on the depth of processing at encoding. They proposed a model of encoding in which there were different levels of processing. An initial shallow encoding established the structure of the word (that is the shapes and spaces by which the letters of the words were represented). A further stage involved phonemic processing which established the sound of the words (enabled the input to be processed at a deeper level). The deepest level of encoding involved the processing of the meaning of the word: semantic processing.

Craik and Tulving (1975) tested the reality of the model. Subjects were given a series of single words and one of three tasks (each of which would ensure a different level of processing) by which to encode them. After the task had been completed the subjects were given an unexpected recognition task. According to the levels of processing hypothesis, it had been predicted that most words would be recognised when the word had been processed to the deepest level and least when the word had been processed at the shallowest level. These predictions were supported by the results and a similar pattern of results were obtained from subjects who were given a recall task.

Although the results supported the prediction of the depth of processing hypothesis, Craik and Tulving also found that the time taken to process the input varied in a consistent manner so that the condition which produced the best recall results also involved the longest time for processing, and similarly the least well recalled involved least processing time. This caused some doubt as to whether the superior recall was due to the time taken or to the level of processing. A further experiment was devised which involved a lengthy processing of the structural properties of words (subjects were asked



to check vowel/consonant sequences in words against a given pattern). It was found that under these conditions the stimulus words were recalled with the same frequency as those which had had similar lengths of time for semantic processing. Craik and Tulving concluded that the amount of time taken reflected the amount of elaborative processing which was carried out and that this accounted for the retention of words rather than the type or depth of processing. Klein and Saltz (1976) provided support for the elaboration of processing hypothesis by showing that retention of semantically processed words could be varied by controlling the amount of elaborative processing at encoding. Subjects were asked to rate words along one or two dimensions: either pleasant/unpleasant (only) or pleasant/unpleasant and big/little. They found that words which had been rated on two dimensions and therefore processed twice were recalled more often than words which were processed on only one dimension.

In summary this initial research suggests that subjects will carry out elaborative processes for any input if they are instructed to do so, or if they are given a task which directs them to do so. However in a normal (non-experimental) reading situation, the reader seeks meaning and uses semantic processing to a greater extent than structural or phonemic processing.

Reder (1976) proposed that during the process of encoding, semantic elaborations are made automatically as each concept is encoded. She defined elaborations from within the framework of Anderson and Bower's (1973) HAM model of semantic memory. According to the Anderson and Bower model, an input is encoded into a hierarchy of associations. Reder proposes that in addition to this framework of associations (which provides a representation of the exact input) inferences and elaborations are also made and stored with the concepts. Inferences are essential links that are made to bind the text together and which are commonly made by all skilled readers. Elaborations are a special subset of inferences which relate each concept at input to the

idiosyncratic associations which the reader has stored when the concept has previously been encoded.

It seems that inferences operate at text level and create links between the separate propositional trees of the input so that a coherent representation can be formed. Elaborations enrich the input with general plausible but non-essential information about the concepts .

Reder (1976) points out that we sometimes become aware of these elaborations. For example, when we see a film which has been made from a book that we have read, we may become aware that locations, room-settings and characters are not exactly as we had expected. Our expectations had been formed by elaboration of the text and were quite different from those formed by the film producer and presented on the screen. This example illustrates the idiosyncratic nature of associations and raises an additional problem: that is which of our many stored associations do we use as elaborations and by what process is the selection made? For example, if the book describes the female character as blonde and blue eyed, which of our blue eyed blonde associations do we choose to elaborate our understanding? Reder does not discuss this aspect of the process.

According to Reder elaborations have two main functions:

1. They provide connections between statements which at first seem unconnected. (This seems to be at a different level from the inference making process: elaborations making the associations between propositional trees which map onto the same concept).
2. They help to detect anomalies and contradictions. (Again this seems to be in terms of pre-stored associated knowledge, rather than at the coherence level of inferences.)

Reder (1976, 1979) argues that associations that are made by elaborative processing can be used for a specific purpose (in addition to the more general purpose of making 'real' or 'fleshing out' the settings and characters in the

text). That is they are used in making plausibility judgements. She carried out experimental work which supported the hypothesis that elaborative associations (rather than actual inputs) are accessed when subjects are asked to make a judgement about the plausibility of a statement in relation to a just read text. She proposed that separate processes exist for the making of plausibility judgements and judgements about whether a probe statement had been presented verbatim. In the latter case the actual representation of the input is activated and retrieved, but in plausibility judgements, all associations of each concept in the input are activated in order to uncover a subset of relationships. Reder hypothesised that the relationships that have been activated at input by the process of elaboration will require less judgement time. She provided evidence for a two stage process in making plausibility judgements: retrieval followed by judgement. From this, she argued that prior activation by the elaboration process aids the retrieval stage.

Reder's model proposed and supported the notion that understanding the input involves going beyond the surface representation. It requires the activation of associations of concepts and in addition requires processing at a different level than that which is necessary to encode the concepts and provide a coherent text representation. It is however a very general process in that it involves automatic activation of all concepts on the basis of their pre-existing associations. Other research (Stein, Morris and Bransford, 1978; Stein and Bransford, 1979; Masson, 1979) has suggested that elaborative associations which are made in a general and non deliberate manner will not be sufficient to explain the way in which information which is salient for a specific purpose might be recalled.

## Elaborations generated in context

Stein et al (1978) carried out experimental work to explore the conditions under which the elaboration of semantic associations is effective. An initial experiment demonstrated that when subjects were asked to recall missing words from a sentence following a task which involved semantic processing (a rating judgement of the comprehensibility of the sentence) they produced a smaller number of recalls than subjects in a comparable phonemic processing task. From this they concluded that semantic processing per se was not a successful recall strategy. Craik and Tulving (1975) had predicted that the complexity of a sentence would increase the amount of elaboration. Stein et al hypothesised that complexity of sentence in itself would not aid retention of specific information. They devised an experiment in which two groups of subjects were given the same amount of elaborative text in conjunction with a target word. Some were given information which emphasised the distinctive properties of the target word, others were given information that was congruent with the text. Those who carried out the task which precisely clarified the importance of the target words recalled more of them than those given the imprecise elaboration. Stein et al concluded that to be of value in aiding retention elaborations can not be general but must be focussed on some specific aspect of the meaning so that relevant links can be established to ensure successful recall. The experiments carried out by Stein et al depended upon the target word and the elaboration being generated by the experimenters. Stein and Bransford (1979) set up conditions in<sup>which</sup> to investigate whether elaborations are useful for recall when readers generate their own. In the experiment some subjects were given a base sentence with continuation material that directed them to make inferential elaborations about the topic of the sentence. These subjects recalled the connected target word more often than a group which was given general associated information about the sentence. Stein and Bransford argued that the information provided was not

directly responsible for the differences in the results but that the differences existed because the subjects were able to use the information to make relevant elaborations. In the condition where the information was not made available, the reactivation of general preexisting semantic associations was of no help.

They tested the relative usefulness of significant and congruous information in a further experiment. The results of this experiment showed that retention scores were highest when subjects queried the significance of certain concepts and generated elaborations about the relationships between them. These scores were superior to those gained by subjects who were encouraged to generate semantically associated information, but were no better than those of subjects <sup>who</sup> were given information about the topic. Subjects who were encouraged to make congruous associations rather than inferential elaborations retained less than those subjects who were given associative, semantically congruent elaborations.

These results then support the proposal that elaborations which are created to link the significance of concepts in a particular input are more useful as an aid to retention than generally congruous but not significant associations. This is the case regardless of the source of the elaboration (that is whether it is generated from the input or from stored knowledge in response to a directed question). Bransford, Stein, Shelton and Owings (1981) have pursued the argument that learning involves the clarification of the significance of relationships between the concepts. They report the work of Stein, Bransford, Owings and McCraw (see Bransford et al 1981) who asked subjects to generate additional information which would help them to remember stimulus sentences. The elaborations were recorded and judged for precision. The results showed that the more precise the elaboration, the more accurate the recall. This gave additional support to the argument for the construction of elaborations as an aid to recall.

Masson (1979) agreed with the conclusions of Stein and Bransford (1979) that elaborations are generated in response to the relationships between the concepts in a particular input rather than created as an automatic process for all concepts. But he questioned the methodology used in the experiments. He believed that there could be an inhibitory effect caused by the provision of a cue or direction towards non-specific associations. He proposed that the difference in retention scores gained over non-directed subjects by subjects who had been given information (or questions which would direct them to make precise elaborations) may have been inflated by the method used. He maintained that the instructions which had been given to the non-directed group had encouraged them to make elaborations that were not related to the topic and in so doing had suppressed any spontaneously generated elaborations. He proposed that if the effect was to be tested efficiently, then the control group should be allowed to make a free recall and the precise elaborations should be compared with this for a true measure of difference.

He devised a set of experiments in which subjects given cued recall were compared to those given free recall. His results showed that cued recall is superior to free recall even when given the same elaboration of the target word at the outset. He further demonstrated (Masson, 1979) that the cue was just as effective when it was inferentially derived from the context as when it was a verbatim cue. Again the scores were superior to free recall. Masson concluded from these results that cueing at the outset provides a context within which the reader processes the sentence and that presenting the cue at recall recreates the context.

Masson concluded that the creation of context for a to-be-remembered sentence is an important aid in the retention of meaning. However the created context can only usefully be employed in a recall task if the same context is available or can be inferred at recall. This substantiates the results of Stein and Bransford in that it supports the idea that cues must be specific to

the target word and that cues must provide a suitable context at encoding and recall.

Bradshaw and Anderson (1982) reiterated the elaboration of associations position. They restated the view that the generation of a memory episode which encodes a set of multiple, partially redundant propositions with the to-be-remembered information will better enable the to-be-remembered information to be recalled. They made an additional assumption that the ability to make elaborations is a function of what the reader knows in that domain. From this they reason that the provision (at encoding) of information which extends the domain of knowledge around a to-be remembered sentence will aid the retention and recall of the information. They also argued (on the basis of Anderson's, 1977, ACT model of memory networks) that irrelevant information provided at encoding and recall will interfere with recall of the information. They provided support for these assumptions in an experiment in which subjects were given a base sentence to remember. The sentence was presented in one of four conditions:

1. Alone;
2. With additional sentences which contained information which was supportive of the target sentence, but not relevant to it;
3. With additional sentences which contained information of events that resulted from the base sentence;
4. With additional sentences which contained information about the cause of events in the base sentence.

The authors hypothesised that the base sentence would be recalled more often (in cued recall) by subjects in Conditions Three and Four because in these conditions subjects would have relevant knowledge about the base sentence and could therefore make elaborations which would provide extra and relevant links in the semantic network. Subjects in Condition Three would recall the base sentence least often because of interference. The experimental

results, which were based on measured recall obtained two weeks later, supported the predictions.

While the experimental results support the theory advanced by Bradshaw and Anderson, they could equally well be explained by reference to the construction of context. The provision of extra information by Bradshaw and Anderson would seem to contradict the theory of preexisting associations which are automatically activated when information is input. The stimulus material used provides an explanation which depends on integration of information in the input and shifts the emphasis from the integration of information into a preexisting network. In doing this the authors appear to be losing the distinctive approach which distinguished the elaborations-as-associations position from that of Masson and of Stein and his associates.

#### Conclusions.

"Association" and "construction" views of elaborations present two differing explanations for the elaboration of input. Although it may seem that the two explanations conflict, in reality the two models are using the same term to describe two quite different processes. The views are not necessarily conflicting but describe complementary processes in the comprehension of information. The two processes described support the idea (introduced in the introduction to this thesis) that text (or any other input) can be processed in more than one way and that how it is processed depends on the demands of the text and the purpose of the reader.

However the notion of multiple processes each of which initiates inference making procedures (whether they are automatic associations or strategic constructions) raises problems about the processing capacity of the reader. Theories and models which have proposed ways in which inference making might be constrained are discussed in Chapters 10 to 12.



CHAPTER NINEAdequacy of the models and measurements reviewed  
for the analysis and comprehension of Complex Texts.

According to the models and theories which have been reviewed, comprehension of text is a process by which the input from the text is in some way combined with some aspects of the reader's stored world knowledge. Three main areas of discussion arise from the review. These are:

1. The relationship between the content and the structure of the text;
2. The representation of the text in memory ;
3. The processes utilised by the reader in the comprehension of text.

These three issues will be discussed in relation to the texts used in the development of the particular models in the first part of this chapter, and then the implications of the issues/models for the comprehension of more complex texts will be discussed using a particular text as an example.

The relationship between the content and the structure in the comprehension of  
a text.

The theories and models which have been reviewed in the previous chapters suggest that the structured representation and the content of the text might be related to each other in three different ways each of which has different implications for the comprehension process.

a. The structure is separate from the content. The content acts as input for structural analysis.

This approach is adopted by the generative-transformational models: at sentence level by Chomsky (1957, 1965) and at story level by Bower (1976), Thorndyke (1977), Rumelhart (1975) and Mandler and Johnson (1977). The Chomskian model of generative transformational grammar (Chomsky, 1957) proposed that sentences could be generated by the rules of the grammar, and

that the same set of rules could generate all grammatical sentences regardless of their content. Story Grammars maintained this approach with transformational generative rules which would generate a story. According to the generative models the starting point of the process by which a story or a sentence is generated is an initial abstract concept of a story or a sentence. The rules then produce a hierarchical framework with terminal nodes into which the content of a particular story or sentence can be slotted in the prescribed syntactical order. Comprehension involves the reversal of the rules so that the process begins with the content which is analysed to discern the constituent syntactical parts and ends when the analysis is complete and the content has been recognised as a sentence (or as a story).

According to these models the rules are abstract (not content dependent) and can be applied to all examples of the types they describe. The rules include options that allow alternative structuring within a basic grammatical framework so that an infinite variety of surface forms with many possible syntactic combinations can be created. This is the main strength of the grammars (that the rules can account for the production and analysis of all grammatical sentences or all simple, single protagonist stories).

There are three main limitations to the grammar so far as comprehension is concerned. The first is that the rules which produce all grammatical sentences cannot explain the production and analysis of utterances which are comprehensible but not grammatical (for example colloquial and elliptical speech). Secondly the rules which will produce all simple single protagonist stories are inadequate to generate or analyse complex stories in which there is interaction between the characters or in which the plot moves forward and backward in time and space.

The other main limitation is that the very flexibility of the rules which allows all examples of their type to be generated means that the structure of a particular story cannot be generated in an abstract manner unless the

content has first been made available. (In other words unless speakers know what they want to say, they cannot generate a framework from the available options in which to say it). This means that whilst a set of abstract rules is flexible enough to create or understand the structure of all sentences or simple stories, the generation or comprehension of the structure of any specific story is content dependent.

Chomsky (1965) conceded this point to some extent when he introduced the semantic element into his model. However the revised model which he proposed was primarily a syntactic structure onto which a semantic element was added. The problem of how speakers actually said what they intended was avoided, the main reason being that Chomsky was intent on modelling the competence of speakers and not their actual performance in specific examples.

Bower (1976) and Thorndyke (1977) have demonstrated that the structure of a story is separable from the content (that is, that the same structure is discernible when repeated in more than one story). However this means that the structure is not useful in the comprehension of the first story until after the story has been read and the structure has been analysed. These studies avoid the problem of how the story content is processed initially so that analysis can begin, and the issue of at what point in the story the structure is discernible. The assumption would seem to be that comprehension of the structure takes place after the whole of the content has been processed in some way and therefore the structure of the text is a result of and not an aid to the comprehension of the content of the text. It seems therefore reasonable to assume that the models which describe the structure of the text in abstract terms are not concerned with the comprehension of the content of individual text. In other words they are concerned with competence and not with actual performance and its consequent problems of meaning and representation of individual texts.

Although the experiments by Bower and Thorndyke have shown that the

awareness of a story structure decreases the reading time for stories with the same structure (embedded or conjoined), it is very unusual (outside an experimental or a classroom situation) for a reader to read simple stories in rapid succession or to read a succession of unrelated sentences. What is more usual is that the reader will read narratives which have a more complex structure (with interactions between the characters and with shifts of time and place in the setting). In this situation (where they are reading for example "best sellers", detective stories or romantic fiction) the reader is more likely to be interested in the content of the narrative. The author may manipulate the structure of the narrative so that some aspects of the content are emphasised, and so that the readers can make predictions about the kinds of events and actions that might follow. In such examples the speed at which the structure of the text might be comprehended is not an important issue, what is more important is the effect that the variation in the structure will have on the reader's comprehension of the content. Clearly, this has not been the concern of structural models.

According to the generative grammar models, comprehension would seem to be the construction of an ideal structure from the given input: the structuring having resulted from application of rules to a text the content of which has already been comprehended in some way.

It can be concluded that generative grammars are inadequate to account for the generation or comprehension of the structure of a particular text because of their flexibility to describe the general and abstract aspects of all typical texts. This flexibility prevents them from constructing or analysing the structure of any specific text unless there is some existing representation of the content. And yet the models do not incorporate procedures which will account for the inclusion of the content of particular texts. The contribution of the models to an understanding of the comprehension of texts is that they describe the structural constraints which

surround the production or analysis of a type of text and that they demonstrate by use of examples that these constraints exist when the reader is recalling the content of actual texts. However an assumption that a simple story structure has primacy in models of text structure can be misleading. This will be discussed further in Chapter Ten.

b. The structure is created from the content of the text.

Case grammar (Fillmore, 1968), coherence graphs (Kintsch, 1974; Kintsch and van Dijk, 1978 and Kintsch and Vipond, 1979) and Scenario (Sanford and Garrod, 1981) models describe the process by which the structure of a narrative is created from the content of the text.

Case grammar describes the semantic structure of text content by means of rules which focus initially on the constraints imposed by the semantic content of the verb. This means that the rules are working on small inputs of text (that is, the initial input is as large as the point at which the first verb is encountered). Although general categories can be imposed upon individual verbs because they have semantic properties in common (and in some respects the semantic qualities of the verbs are comparable to the "primitives" of Schank and Abelson, 1975) and also upon the noun phrases that are permissible in conjunction with the verb on the basis of these categories, it is knowledge about semantic properties of individual verbs that allows the reader (or the text creator) to decide that two or more verbs constitute a particular category. It is not that categories pre-exist and that the reader/writer fits the verbs into them. Thus the categories are loosely defined and the number and type of related noun phrase cases is tentative. Propositions, which are the basic units of Case Grammar, are made up of the verb and its connected noun cases. Fillmore was concerned with uncovering universal rules by which text could be produced and analysed. As a result his work looked for commonalities within and between the structure of languages.

Kintsch was not concerned with the underlying structures of language meaning and the universal aspects of the Case grammar. He assumed as his starting point that texts do accord with these rules and in his model he used case grammar as a means of breaking down the text into small units or propositions. He then used the cases within the propositions to demonstrate the coherence between propositions. At this point Kintsch (1974) was not looking for deep underlying meaning or structure but the surface coherence of the text.

In his 1974 model Kintsch analysed texts into propositional units to produce a coherence graph. This gave an idealised analysis of the text showing the referential connections that held the text together as a complete unit. In this version, Kintsch's (1974) coherence graph described the process as it applied to a complete text with no constraints upon the size and number of units processed. Like the generative grammars described in the previous section it required some total representation of the complete text so that it could make all the possible connections in the content. However, Kintsch and van Dijk's (1978) cyclic model described the process by which the reader might produce this structure from a word by word left to right input of text. The process by which the text was analysed (proposition by proposition) was constrained by the limitations of the capacity of working memory. The resultant structure was a graph of the specific text, its shape being dependent on the connectedness of concepts within and between the propositions and on the limitations of the reader's working memory which constrained the possible connections which could be made.

Sanford and Garrod's (1981) Scenario model also describes the process by which the structure of an individual text is created from the content. As with Fillmore (1968) the analysis of the text is initiated by a semantic decomposition of the verb, and the structure of the sentence is created from the semantic links between the verb and the rest of the sentence. The model

does differ from case grammar in that there is a fixed number of primitive meanings to which the verb is decomposed. Unlike Kintsch (1974) Sanford and Garrod use the decomposition of the verb and its deep semantic meaning as the basis of the model. The process and the similarities and differences between the Scenario and coherence models are described in some detail in Chapter Seven. What is relevant to this discussion is that although they deal with different levels of meaning both models view the relationship between content and structure in a similar manner: that is, according to these models the structure is derived from the content and is based on the semanticity or "meaning" of the text.

Kintsch and van Dijk (1978) have proposed a further stage in the coherence model. This describes the process whereby the reader's knowledge about story structures constrains the production of a summary of the story. Propositions which form the coherence graph are further processed to produce a structure which is relevant to the stored ideal structure. Like the generative grammar models this assumes two stages of comprehension, but it differs from generative models in that it focusses on the process by which content is shaped in accordance with these structures.

The models which have been described in this section are concerned with the process by which structure of a text is derived from the content by the reader. In these models the content and its semanticity are of prime interest and the structure is created from and is reliant upon the content.

### c. Knowledge-based structures.

A third approach to the relationship between the content and structure in the comprehension of text is that the text representation is structured according to stored knowledge based structures. According to these models (Scripts: Schank and Abelson 1975; Frames: Minsky, 1975 ; Schema: Neisser, 1976; Fact structure: Kintsch, 1979) the reader integrates the content of the script into his/her pre-existing stored knowledge. The models each assume

different kinds of structures for the knowledge and differing degrees of flexibility in their operation (these differences have been discussed in Chapter Four) but they agree that the structures are based on experience that has been gained from interactions with the world and that the knowledge from this has been formalised into structures. The structures are used and modified in the comprehension of newly input information.

Schank (1981) has recently modified his notion of Scripts. Initially they were conceptualised as plans for direct experience (say, a visit to a restaurant) which had been repeated and modified by continual experience until they were stored compiled as a structure for the planning or comprehending of that experience. The interaction between the compiled knowledge and new experience (either direct experience or experience gained from texts) was two way: new information could be added to the existing structure or the knowledge which was stored in the existing structures could fill in omissions in the input (for example, the knowledge that a waiter/waitress would have brought the food in between ordering and eating).

His modified version (1981) is more economical for the storage of information. Instead of there being precompiled scripts for each kind of experience (for example, "visit to French Restaurant"; "visit to snack bar" or "visit to Pizzeria") knowledge is stored in general (for example, all eating places have food, tables, crockery) and specific categories (for example, French Restaurants have waiter service). In this way general information gained in one experience can be applied to other experiences for retrieval or comprehension processes.

These models assume that the content of the text is structured during the comprehension process and that pragmatic or world knowledge structures are applied to and changed by the comprehension process.



**Summary.**

The three approaches which have been discussed perceive the relationship between the content of the text and the structure of the text representation in quite different ways. All agree that a structured representation is an outcome of the comprehension process. It is the formation of the structure that differs. According to the different models it is seen primarily as a syntactic, a semantic or a pragmatic structure. The syntactic structure is the result of stored rules about structure imposing a structure upon the content; the semantic structure is a result of stored rules about semantic relationships creating a structure from the content and the pragmatic structure is a result of interaction between stored structures of pragmatic knowledge and the content of the text.

Each of these approaches tacitly accepts the existence of the other structures (Kintsch's final model, 1979, incorporates all three aspects) but stresses the competence of their particular model to deal with aspects of comprehension which one or another model is inadequate to explain. All depend for their explanation upon texts which are in some way restricted or are specially constructed to support their argument.

In all these models it is assumed that the comprehension process relies on the content of the text and when the text does not comply with the rules of the model then it is rated as being incomprehensible. For example, according to the Kintsch models if the text did not cohere then it would be incomprehensible and according to Schank and Abelson's Scripts input which did not relate to a script would not be comprehended. It is in this respect that the models differ from the content based models (for example see Bransford and Johnson's, 1973 work in Chapter Seven). These show comprehension to be a contribution which is made by the reader, that is in situations where the text is not coherent or sensible, the reader will create elaborate inferences to link the text input with stored knowledge about the situation described.

### Representation of the Content of Text in Memory.

All the models agree that as the content of the input of a text is encoded it is transformed in some way. The hypothetical form it takes differs according to the perceived structural changes of the particular model. The various models see the content of the text as being encoded according to their particular structural perspective. However they differ basically in that some models propose that there is a separate representation of the text content whilst others assume that text is stored as a part of a more general memory.

#### a. Models which have a separate text representation.

There are two basic kinds of model which see the text as a separate representation in memory, Kintsch and van Dijk (1978) and Sanford and Garrod (1981). The Kintsch model proposes that there is a representation which is held for further processing and is then stored as a structured text. Kintsch's (1979) model sees the process as one in which there are parallel processes of coherence formation and meaning extraction, followed by the macroprocess. During the macroprocessing the macrorules extract an overall gist which accords with some overall structure. Information which is present in the coherence graph but which is not relevant to this structure is stored in a separate and more general memory. Story grammars conform to this second (macrostructural) stage of the model. Implicit in the story grammar models is some earlier stage in which a temporary representation of the story is formed and the structural rules operate on this.

The Scenario model of Sanford and Garrod (1981) also proposes that there is a separate text representation. However this is stored directly from the encoding without reprocessing further into a more general structure. The text is stored for a length of time as an encoded structure in a long term text representation. Over time the structure fades and the informational content is absorbed into the general long term memory, where it serves as a base for future scenarios. Both types of model are based on the linguistic features of

the text.

b. Storage of text as a part of a more general memory.

Models which propose that text input is stored as a part of a more general memory are those which see text as general information input in which the content of the input is encoded according to general conceptual semantic and/or pragmatic knowledge and not primarily according to linguistic distinctions and grammars. The three main types of models are the semantic network models, particularly the Quillian (1969) and the Anderson and Bower (1973) HAM models; and the "constructivist" models of Bransford et al (see Chapter Seven) and the "primitive" or decomposed models of Schank and Abelson (1975).

The main difference between the models is that the semantic network models imply that knowledge is stored in a propositional form and can be retrieved entirely when (or if) accessed. The other models assume that the information within the story is integrated (Bransford) or disintegrated (Schank) and then combined with other stored knowledge such as spatial and temporal knowledge about events and concepts.

The form in which the information is stored is only one of the issues that separate the semantic network models from the pragmatically based "constructivist" models. A further issue which is connected to the process by which information input from text (or from other sources) is integrated into stored knowledge will be discussed in the next section.

**Summary.**

The two main differences between the models on the issue of the storage the context of a text representation centre around the problems of:

- (i). Whether the text is stored as a separate representation or a part of a general memory representation and
- (ii). Whether the content is decomposed into primitives or not.

The first issue is the one which divides linguistically based models from the world knowledge models; the second issue is one which separates world knowledge models into those which are based on the representation of conceptual information in the text and those that are based on a representation which incorporates pragmatic knowledge activated by the concepts in the input.

The processes which may be involved in the comprehension of text.

It has already been said that the comprehension of text is a process in which the input is in some way connected with the stored knowledge of the reader. Two of the issues that arise from the models which describe this process have been discussed and these have highlighted the differences which arise from the perspectives which primarily consider the syntactic, the semantic and the pragmatic structural aspects of the process. Further discussion has considered the differences which separate the linguistic approach from the world knowledge approach on the issue of storage of the content of the input. More specifically the differences have been examined which arise between the models which assume that the content is stored in a recoverable propositional form and those which assume that the informational content of the input is transformed either by expansion or decomposition.

The third issue which is to be discussed is the actual processes by which the information in the text is integrated with the stored knowledge of the reader. In order to do this the issue will be discussed in relation to the three perspectives of text representation: that is, as an encoded propositional representation; as a decomposition into a primitive underlying semantic meaning and as a representation in which input is linked to stored knowledge by links which indicate the significance of the input to that which is already known.

**a. The Propositional Storage Models.**

These models are the semantic network models, in particular HAM (Anderson and Bower, 1973) which has been extended by the work of Reder (1979). This particular model assumes that content of the input is stored in the terminal nodes of a network of associations. As the input is encoded into this network the associations with which the input concepts have previously been encoded are activated. A consequence of this (according to Reder's theory) is that the input is elaborated by these activated concepts and they automatically become a part of the comprehension process. This means that the input is automatically enriched by pre-existing knowledge of the concepts and the representation is tagged with these additional associations. The elaborations are utilised if the reader has to make plausibility judgements about the text. He/she also becomes aware of these elaborations if subsequent input contradicts them. Other than in these circumstances the reader remains unaware of them. The implications of this process are that unless there has been prior connections between concepts then the connections are not made and the input remains isolated from other nodes in the network and is not therefore integrated with existing world knowledge.

**b. The Decomposition of Meaning Models.**

These models (Schank's primitives, Sanford and Garrod's Scenarios) assume that input is decomposed into a primitive deep meaning which allows a framework of expectations to be created for the incoming text. If the input does not fill the slots in this framework then a new framework is created into which the input will fit. This creation of frameworks continues until the text has been processed. Frameworks which have been created are accessible and can be recalled if required in the process. The implications of these models for comprehension is that unless there is a pre-existing action based connection between items of information in the input, they will not be linked. New links will not be constructed to create new insights into the material and

figurative or metaphorical use of language will only be comprehended if the verb content implies the same primitive action between two concepts. This model is not creative and could only interpret text in terms of what is already known.

c. Construction of meaning models.

These models are concerned with the process by which the reader makes the input meaningful. The meaning can be constructed between items in a text, so that certain aspects of the text become meaningful because they can be linked to other items in the text. For example, in a sentence such as "The man who wore brown shoes had to stoop to get through the doorway", the stoop would be significant because it meant that the man could not easily get through the doorway if he was upright. It is likely that recall would contain some mention of his height and that the colour of his shoes would not be recalled unless cued or prompted and thus made significant.

The meaning can also be constructed by the links that are created between the content of the text and the reader's general knowledge about the situation. Bransford and Johnson (1972) have shown that it is necessary to activate world knowledge in some way before the text is input so that the reader can create the necessary links. It is this need to activate the appropriate context that separates "Associationist" models from the "Constructivist" ones. In Associationist models previous associations are automatically activated; in Constructivist ones specific contexts have to be constructed to accommodate the input. The differences between the models has been discussed in Chapter Eight. The same kinds of differences have also been discussed in relation to the linguistic models of Kintsch (1979) and Sanford and Garrod (1981). In the discussion about the meaning structure of texts, it was stated that Kintsch offered no explanation for the selection of particular frames for the input of text to the meanings but that the frames were related

to non linguistic knowledge about the total situation that was being described. Sanford and Garrod, on the other hand, derived their frame from the primitive meaning of the verb content of the text. It would seem that the Kintsch (1979) meaning process relies on the same construction from pragmatic knowledge as that described by Bransford and Johnson but in neither case is the process by which the information is activated made clear.

The implications of the constructivist models are that linguistic knowledge on its own is not sufficient for successful comprehension. Unless the reader has the appropriate pragmatic knowledge to form a framework with which to link the text and prestored knowledge then comprehension cannot be effective. A further assumption is that specific and appropriate non-linguistic knowledge must be activated if the comprehension process is to be effective. The framework is not one that consists of prior associations between certain concepts which are automatically activated but one that is significant and relevant to the combination of concepts which together make up the text. This is so even if there has been no previous association between the actual concepts and the activated knowledge. The process by which this creative aspect of text comprehension is effected has not been explored. Research which has been carried out in this area has involved methods which instruct the reader to make links by which to remember a sentence (Masson, 1979) or has used cued recalls to test whether the reader has made the links (Stein and Bransford, 1979). Goetz (1977) has shown the conditions under which certain links are made. The question remains whether the reader will do this spontaneously and if so how this can be investigated and measured.

#### **Measurements of Comprehension.**

The extent to which a text is comprehended can not be measured directly. All that can be measured is what can be retrieved, and what is retrieved is a function of the retrieval instructions that are given to the reader as well as

what has been comprehended from the text. Pichert and Anderson (1977, see Chapter Seven) and Pratt, Luszcz, MacKenzie-Keating and Manning (1982, see Chapter Four) have demonstrated adequately how recall instructions can influence both the content and the structure of what is recalled. Both the instructions and the text are influenced by the design of the experiment. For example, one cannot measure the reality of a proposed hierarchical storage structure unless one creates a text which includes concepts which would appear at various levels in the hierarchy. And certainly one can not decide whether there is any reality in the model unless the instructions present a specific task by which those distances can be measured in some way. In this example the instructions and the text present the reader with a very specific task from which it is possible to make certain statements about the differences in response time and to correlate this with hierarchical structures. This design does not allow any spontaneous comprehension processes to be expressed or measured. At the other end of the scale Educational tests of comprehension have based their measurement on items which can be scored from information in the text. These tests allow for the expression of spontaneous and creative constructions but such expressions can easily not be measured because they do not accord with a standardised marking scale. False recognition paradigms allow for the measurement of constructed responses but do not allow for the free expression of constructions. The plausibility of their content may allow readers to "recognise" constructions which they may have never made spontaneously.

What is measured as comprehension then cannot be freed from the experimental design, the text and the instructions and to measure spontaneous constructions in the comprehension of text in isolation from these variables is impossible.

To measure what the the subject produces when asked open ended questions about a text is possible although such measures will perhaps influence the



subject to construct information which he may not have otherwise have constructed because it was not relevant to his own personal interpretation of the significances in the text. In this situation such a problem seems intractable.

#### Implications of the Issues/Models for the Comprehension of Complex Texts.

The models of text analysis and construction which have been discussed so far have mainly used simple texts as examples and models. These have been either simple sentences or propositions or simple stories. Whilst these have been suitable to elicit the information that was required by the experimental design, they do not resemble the texts that are regularly read outside an experimental situation.

The use of single word and single sentence texts as stimulus material for the study of discourse processes has frequently been criticised on the grounds that their simplicity is inadequate for models of complex processes and that the structure and content of each sentence interacts with that of other sentences within a continuous text and does not only add to them. Because of this it is not possible to construct a model of the reading process from single isolated words or sentences and then transpose it to a longer text. It is in response to such criticisms that models of continuous text have been built. However the same criticisms may be levelled at models of simple texts :that is they are inadequate to describe either the structure or content of complex texts.

Simple stories have been studied because they have a clearly definable structure which can be described by a set of rules. If the structure of a text can not be described by these rules then by definition it is not a well structured simple story. The content of the simple story relates to one main protagonist and describes either a series of events which either occur in sequential order or which occur in relation to a well defined goal. Because

both the content and the structure are simple they are predictable. Characters are not complex because they follow a well structured sequence and their goals do not conflict with those of other important characters. But because of their simplicity, their predictability and the lack of complications simple stories are incredibly boring and are rarely read for pleasure.

Narrative stories which sell well and are by definition read by a lot of people may or may not have a well defined overall structure. Within the overall structure of the total text there are Chapters and even paragraphs in which the structure is not simple and straightforward; in which there is more than one important character and in which the goals of the characters are not well defined or in accord. In such texts the structure and the content are not predictable. It may be that tentative hypotheses (such as those proposed by Neisser 1976, see Chapter Two ) are constructed and then confirmed or denied as the text progresses. This statement can be supported by the use of the opening paragraph from a "best seller" which will illustrate these points.

"Jenny Fields was arrested in Boston in 1942 for wounding a man in a movie theater. It was shortly after the Japanese had bombed Pearl Harbour and people were being tolerant of soldiers because suddenly everyone was a soldier, but Jenny Fields was quite firm in her intolerance of the behaviour of men in general and soldiers in particular. In the movie theater she had to move three times, but each time the soldier moved closer to her until she was sitting against the musty wall, her view of the newsreel almost blocked by some silly colonnade and she resolved she would not get up and move again. The soldier moved once more and sat beside her."

(Extract from John Irving, "The World according to Garp", 1978, p.1.)

a. The Structure.

The surface structure of the text can be divided quite neatly into three parts: the Outcome, the Background (or setting) and the Events.

(i). The Outcome:

"Jenny Fields was arrested in Boston in 1942 for wounding a man in a movie theatre."

(ii). The Background:

"It was shortly after the Japanese had bombed Pearl Harbour and people were being tolerant of soldiers because suddenly everyone was a soldier, but Jenny Fields was quite firm in her intolerance of men in general and of soldiers in particular."

(iii). The Events:

"In the movie theater she had to move three times, but each time the soldier moved closer to her until she was sitting against the musty wall, her view blocked by some silly colonnade and she resolved she would not get up and move again. The soldier moved once more and sat beside her."

However the time structure of the paragraph (when presented in this way) is not simple and straightforward. The initial statement (i) follows in time after some missing event which follows (iii). If the structure of the text is changed to create a sequential time structure - for example, (ii) then (iii) then (i) - then the content is also changed. For example, the "It" at the beginning of section (ii) refers to the arrest in section (i). If section (i) is removed or put into a time sequence then "It" becomes a general non-specific introduction. As a further example tentative hypotheses about the goal of Jenny Fields (iii) which could be made on the basis of the information about her arrest, would change if this information was not given until last. Therefore one of the ways in which this paragraph has become complex is that the time sequence of the events clashes with the structure of the plot: the arrest in section (i) is in effect a setting for the interpretation of events

which follow it in the text, but precede it in the time sequence.

How the structure and content in the Garp paragraph relates to the theories in the review.

The previous section of this chapter showed that the surface structure of a paragraph from complex text can be neatly divided into three sections. Theories such as that of Kintsch (1974) and Kintsch and Vipond (1979) which are concerned with the coherence of text would be able to produce a representation of the text. These theories are however not concerned with the meaning structure of the text. Sanford and Garrod's (1981) model would be able to produce a meaning structure for the separate sections but would not connect the Scenarios except by means of the characters and the roles. An overall meaning structure of the total paragraph would not be created if the links are not explicit in the text. The text representations produced by these models would not reflect the complexity of the meaning structure.

Simple story grammars would be able to describe either the sequential structure of the events or the causal structure of the events according to the goals of either Jenny Fields or the soldier. The previous section described the inadequacy of these self contained descriptions for a complex text and this inadequacy was emphasised as the complexity of this particular text was explored.

Eventually one must ask why the author has structured the opening paragraph of his text in such a complex way and more specifically how such complexity relates to the general goals of a writer. But perhaps such a discussion is best left until after theories relating to the process by which the reader links the content of the text to his knowledge of the world have been discussed in relation to this particular paragraph.

The main distinction between the "Associationist" theory of Reder and the "Constructivist" theory of Stein and Bransford is that Reder proposes that the links between the text and knowledge of the world are made automatically on

the basis of prior associations and that the current context has no significance in this process whereas Stein and Bransford would see the links being made to highlight the significance of concepts within a specific context. In the comprehension of the Garç text a construction which is made on the basis of what is or is not significant may be a very individualistic and idiosyncratic decision. For example, the colour of Jenny's hair, her height may be generally of very little significance and therefore any elaborations which the reader makes about these qualities may be made in a very automatic manner. On the other hand elaborations about why the soldier moved closer or why she moved away should generally be made to clarify the reasons for the characters moving about the cinema. Other elaborations such as her /his age and appearance may be made automatically but not seem to be significant unless the reader is asked about them in a test of comprehension. Asking direct questions about these factors may lead the reader to construct a link to make them important or significant as a response to the question and not as a part of his/her own construction of the meaning of the text. Again this leads to an intractable problem if comprehension is tested by means of direct questioning.

To return to the question of why the author might construct the text in this particular form, it is first necessary to ask why the author has written the text at all. At a very broad level of explanation it can be said that the author would want his text to be read and that he would write it in such a way that it would appeal to as large a population as possible. If a novel or a narrative is to appeal to a reader then the reader must want to read more about it and it would seem that one way to ensure that the reader wanted to read more is to ensure that he/she becomes involved. One way in which this involvement might be achieved is by the creation of a text which is sufficiently complex in structure and content for the reader to create a variety of predictions at several different levels about the characters and the events and the overall structure. However it must be coherent and have

sufficient predictability for the reader to confirm some of his hypotheses at one level whilst at the same time leaving ambiguities and uncertainties at other levels for the creation of new speculations and hypotheses.

It is not possible to measure the intentions of the author from the text. It is however possible to obtain responses of subjects to the text or to various sections of the text. An analysis of these responses will indicate whether the reader does become involved in this way; that is whether the kinds of predictions made by subjects in response to sections of the text will vary in a reliable way.

Ways in which these responses might be elicited and analysed are explored in a pilot study described in the next chapter.

CHAPTER TENExperiment One: a methodological exploration.Introduction.

## Structural Models of Narratives.

Experimental work which has supported the structural models of text comprehension has shown that the structure of a text affects the ease with which it is comprehended and recalled. Bower (1976) showed that disruption of the structure of a story which originally conformed to an ideal structure interfered with what was comprehended and recalled. Stebbing and Raban (1982) have shown that 7 year old children who are beginning to read will comprehend and recall a goal structured story more easily than a temporally structured story. They have also demonstrated that a goal structured story is better recalled in temporal order than a temporally ordered narrative which has no overall goal structure. Bower (1976) has demonstrated that a goal structured story is more easily recalled than a story that has only a temporal structure. From these limited examples it can be concluded that structure and order are important for the comprehension of a story and that the structure of a goal based story is of more importance in comprehension and recall than a simple list of events in which the order of the sequence has no particular significance. It could be argued that one reason for this is that the reader will become more involved in a story which has a goal based plot because she/he anticipates the structure and on that basis is able to make tentative predictions about the outcome. Further it can be argued that in a simple story the accuracy of these predictions is aided by the regular structure because the apparent predictability of the structure leaves the reader free to speculate about the content. These arguments assume however that the reader is certain about the structure and that the structure is predictable and

regular from the outset.

Research into the structure of simple stories (Rumelhart, 1975; Thorndyke, 1976; Bower, 1976; Mandler and Johnson, 1977; and Stein and Glenn, 1979) has shown that a structure can be created on the basis of optional, regular rules. When these rules are reversed for the comprehension of a story it is not possible to be certain about the final structure of a story until the final event has been processed. The research (which has been concerned with the structure rather than the process of text analysis) has not established the point at which the reader is sufficiently certain of the regularity of the structure to be able to proceed on the basis of this. Because these studies have been limited to the grammars of simple stories research has proceeded as if this is the only structure of which the reader is aware and it seems to have been assumed therefore that he/she does not have to make a decision about the kind of structure being used.

Most stories do not have a simple structure. However if the simple story perspective is adhered to then it would be easy to conclude that the optimal conditions for comprehension are those in which the development of the text follows a well structured form in which there is a clearly defined goal structure. Even more beneficial would be a situation in which all stories were presented in this way so that the structure could be taken for granted because there were no alternatives to consider. Such conclusions would be misguided for several reasons. One of the reasons (discussed in the introduction to this thesis) is that a text which is easy to comprehend may quickly become boring because it demands no major contribution from the reader. Another reason is the more tangible evidence offered by Bower (1976) that repetition of the structure eventually leads to confusion of detail to such an extent that the reader would be unable to distinguish the content of one story from the content of another.



It would seem that the process involved in the comprehension of narratives may be considerably more complex than is assumed by structural models of simple stories. This would be because there are alternative structures for text and the reader cannot confidently decide until the story is read which of these structures has been used. Thus arguments which depend on the assumption that the structure of simple stories can be taken for granted at the outset is invalid because the reader cannot be sure until the story is read that the simple structure is what has been used or (as will be discussed later in this introduction in relation to experimental work carried out by Goetz, 1977) which simple structure has been used.

If readers cannot be sure about structure they may process content in at least two ways: they may make tentative hypotheses about prior and future events and about the characters which they will seek to confirm in the incoming text (as is suggested, for example by Neisser, 1976); or they may make no ongoing predictions but instead make any necessary inferences when the reading is complete and the reader has become certain about the significance of information in the preceding text to the outcome (as is suggested, for example, by the work of Stein and Bransford, 1979 and Goetz, 1977).

#### **The Effect of the Outcome on the comprehension process: some speculations.**

The strength of producing stories according to the rules of generative grammar is that both the content and the structure lead up to an outcome such as the achievement of a goal or the solution of a problem. In a simple story there is always a highly predictable outcome, particularly when the goal of the protagonist is presented at the outset, for example "Once upon a time there was a prince who wanted a wife." or "There was a farmer who wanted to get his cow into a barn." This predictability is often exploited as in short stories or in jokes when the reader is led towards a predictable outcome and then is presented with a divergent solution or outcome.

Goetz (1977) demonstrated that the outcome of a story affected the

inferences made by the reader after reading the complete text. He demonstrated that readers make inferences about those events which are significant or important to the outcome. The text he used described the preparations made by a business woman for a journey by aeroplane. All subjects were given the same information about the character and her background. This included information about her divorce and her career and about a rather general problem which she had promised to discuss with her son. Then they were told that she arrived half an hour late at the airport. No reason was given for her late arrival. One set of subjects were told that her tardiness saved her life by causing her to miss a flight that crashed. Others were told that she had to wait for a delayed flight. Subjects were given a multichoice questionnaire which included the target question "she was late arriving at the airport because....". Those given the first version made the inference that she was late because she had stopped to talk to her son significantly more often than those given the second version. Subjects in the second version based their inferences about her lateness on more general possible causes such as talking on the telephone, traffic conditions or running into a friend. Goetz concluded that the importance of the outcome led to the inference being constructed and that the prior event was given importance in this context.

The conclusions which were reached in this experiment were based upon measures which were taken after the outcome of the story had been read. They were therefore based upon a complete structure of the text. Those subjects who were given the first outcome read a goal structured story. Those given the second outcome read a list of events which were temporally structured but in which the events had no particular relevance to the outcome. An alternative explanation for Goetz's results might be that, as was argued earlier, the subjects having reached the end of the text were retracing the structure and relating the content to the structure and making those inferences which they thought were necessary. Those subjects who were given the goal structured

story would be seeking a causal relationship which had some significance in relation to the outcome. Whereas those who were reading a list of events would be looking for a plausible association. All intervening events are plausible whereas only the woman's intentions to discuss some problem with her son was specifically mentioned and might, on reflection, constitute a cause that would be significant in the light of the outcome.

Whatever the process, in Goetz's experiment it is the information in the outcome which makes her lateness of significance or irrelevant. This conclusion is the same whether it is argued that this is because of the content of the outcome itself or because the content of the outcome confirms the structure of the text.

Goetz did not test his subjects until they had read the complete text. Therefore it is not possible to decide whether the results are attributable to the content or the structure of the text. What the experiment did show was that changes in the structure and the content made by the use of alternative outcomes do influence the kind of inferences made about preceding events. Goetz used his data to conclude that inferences are made after reading and that inference making is constrained by the relevance of the outcome to the preceding information. The assumption was that inferences are made at recall or on reflection and not as an ongoing part of the process of comprehension. The conclusions, based on measures taken after the text had been read, in no way explored the possibility that more tentative hypotheses might have been formulated during the comprehension process.

It is proposed that tentative hypotheses are made in an ongoing fashion but that the information in the outcome will confirm or change them. If the outcome is responsible for the construction of inferences in a story which has a regular sequence, then changing the order of the structure so that the outcome is given at the beginning of the story, should mean that these inferences will be made as the relevant input is encountered. Evidence is

presented by Mandler and Johnson (1977) that the order is an important factor in the story structure and that a reader will recall a story in canonical order even if <sup>it</sup> is presented as a flashback (that is with the outcome at the start and the events which led up to it presented afterwards). Whilst this is valid for the conclusion that the structural rules are used in retelling and recalling simple stories (and the experimental work of Pratt, Luszcz, MacKenzie-Keating and Manning, 1982, supports this notion), it gives no indication of the procedures that are used in the processing of the text for comprehension. Baker (1977) showed that flashback texts are more quickly recognised when presented in their flashback structure than when presented in canonical order, which suggests that the originally presented order of the text can be more important in the process than the canonical order per se.

Bower's (1976) research showed that disrupted text is less well comprehended and recalled than text which is presented in the order in which it was written, demonstrating clearly that ordered text is better remembered and understood than disordered text. However his research was concerned with the disruption of simple stories which had been written in canonical order according to the rules of story grammar and it says nothing about texts which are deliberately written and presented in a different order from that prescribed by a simple story grammar. It is possible (as shown by the Garp extract in the previous chapter) to produce a well structured, coherent and comprehensible story which does not comply with simple story rules.

It is proposed that if readers are given a text in which the order is structured in such a way that they are given the outcome at the beginning the inferences that they make will differ from those which are made in a tentative manner (about the same events) when the outcome is not known. If the tentative hypotheses are expressed, they may be very idiosyncratic, or may be based on different areas of the given text. Reder (1976) has shown that given a fixed number of choices subjects will make highly plausible hypotheses in a

consistent and reliable manner. Whether or not subjects will make reliably similar hypotheses when asked to make spontaneous predictions remains to be explored.

Models which have described the process of comprehension (Sanford and Garrod, 1981; Kintsch and van Dijk, 1978) have not discussed the question of irregular structure. Sanford and Garrod have described the ongoing processing of input text in a model which takes no account of the order of presentation or of macrostructural influence. Kintsch and van Dijk have taken macrostructural influence into account but have not taken account of variations in the structure.

Authors, however, do use a flashback technique and readers do process text which is presented in this way. Although the eventual recall of the content may be in a sequential and goal structured form the ongoing processing of different structures must differ from a straightforward presentation.

The experiment which is to be described investigates first of all the effect of Outcome upon predictions made about events and upon hypotheses about the appearance and personality of characters.

### Methodological problems

**Problems of measuring the comprehension process.**

The main problem of measuring the process of comprehension of text is that until the text has been read the processing of it it cannot be measured. This implies that ongoing processing cannot be measured. Once a text has been read it is recalled in retrospect and Goetz has demonstrated that there are kinds of constructions that are only made once the final sentence has been read. Although the ongoing individual processing cannot be measured, the problem can to some extent be overcome by assessing the comprehension of subjects whose reading is stopped at various points in the text.

Measures of comprehension are commonly made in four ways:

1. The time taken to read the text is measured (Kintsch and Keenan, 1973; Keiras, 1978);
2. The subject rates the comprehensibility of the text (Bower, 1976);
3. The subject recalls the passage (Kintsch, 1974);
4. The subject is tested either with (a) a recognition task (Bransford and Franks, 1971) or (b) with comprehension questions (Stebbing and Raban, 1982; Barrett's taxonomy, 1972).

Collectively all these methods present problems for the measurement of the comprehension process and individually they present problems for the collection and quantification of data.

#### The Problems of methods of measurement.

##### 1. Reading Time.

The advantage of reading time measurements is that quantifiable data can be obtained by which to make comparisons between conditions. The main problem of measurements of reading time is that the experimenter has to assume that what readers are doing during the reading time is processing the text in the way in which the experimenter has predicted. In reality they may not be carrying out the particular operation or if they are, this may not be the only thing that they are doing.

In the present experiment reading time comparisons would not be a useful method of data collection as readers will be reading differing amounts of text. A separate consideration is that the experiment aims to establish what readers are doing at a certain stage at the text by collecting the kinds of predictions they are making and not by measuring the amount of time which readers spend processing the text.

## 2. Rating.

(i) When subjects rate the comprehensibility of text, they are rating textual coherence and cohesion regardless of the size or composition of the text. A judgement of this nature is useful when comparisons are being made about the relative comprehensibility of one passage against another when the content and the structure differ.

In the present experiment what is being investigated is the ongoing hypotheses which the reader makes at each point in the text and not the relative comprehensibility of each section. Therefore this is not a useful method to employ.

(ii) However rating scales can be used to investigate the importance of certain characteristics of protagonists. The problem with such investigations is that the choice of characteristics has been made by the experimenter and such characteristics may not have been selected by the subjects if they were to spontaneously describe the character. The advantages of using this method are that the range of descriptions is controlled by the experimenter. The problem of limits to the range of descriptions and of the equivalence of descriptive terms used by different subjects seems not to be encountered. However the problem of equivalence is masked by the assumption that terms which are chosen are used in the same way by both the subject and the experimenter.

In the current experiment rating scales will be used in this way with the provision that (if required) initial free descriptions of the characters will be obtained before the rating measurements are made.

## 3. Recall.

Recall has been used to compare memory for the text with the original input text. For example, Bower (1976) and Mandler and Johnson (1977) compared the order in which text was recalled with the original order; Kintsch (1974) compared the propositional content of recall with that of the original input.

The advantage of recall is that it is spontaneously produced by the subject and is not constrained by any labelling of concepts preselected by the experimenter.

The current investigation is concerned with the constructions that the subject makes in response to the given text and in particular the predictions that are made in response to a particular section of that text. In this situation recall can only be used to examine what aspects of the text are recalled or what is added to the input text. An inherent problem is that when the recall contains paraphrase it is difficult for the experimenter to discriminate between what is added to the text and what is given.

#### 4. Testing.

##### a. Recognition tasks.

Recognition tasks are usually used in experiments in which the experimenter wishes to probe whether or not certain inferences have been made. In these situations subjects are presented with a statement and asked if they recognise it. If they 'falsely' recognise a statement which involves an inference predicted by the experimenter, such recognition is taken as support for the inference having been made. The use of this paradigm has two main problems. One is that the subject may never have made the inference spontaneously but having seen it written down is prompted to make it. The other problem is that the subject may be aware that the statement was not a part of the original input but because it is plausible and may even appear to be a reasonable paraphrase of the input does not reject it.

That subjects do make plausibility judgements has been demonstrated by Reder (1976, 1979). These studies however have clearly stated that plausibility judgements differ from inference making and use different processes.

In the present investigation the central problem is one of investigating the kind of inferences which subjects make, not confirming whether or not they



have made specific inferences. Therefore the false recognition paradigm would not be a useful method of investigation.

b. Asking questions.

Traditionally educational researchers have tested comprehension by asking questions. The research reviewed in Chapter One described the skill related questions which were devised to investigate whether comprehension was a single skill or a multiplicity of subskills. One conclusion which was reached was that the most commonly used data were those which could be reliably quantified and as a consequence be collected from closed questions (which could be scored right or wrong). Open ended questions which required content analysis and subjective interpretation were generally considered to be ill-suited to a investigation of skills.

However the current investigation is concerned with the comprehension process itself and not with the skills which are related to it. In this context open ended questions which allow the reader to hypothesise about the next event and the goals of the protagonists, and to elaborate about the physical and personal qualities of the protagonists are a useful way of eliciting these speculations. However open-ended questions whilst providing the subject with greater freedom of response pose greater problems in terms of the criteria for content analysis.

Each of the methods which has been discussed poses certain problems for the collection of data about the ongoing hypotheses which are made in the process of comprehension. Some methods (for example false recognition paradigms or asking questions which can only be answered by giving explicit information from the text) are clearly unsuitable. Other methods are more suitable. For example, rating scales could be used to assess the relative importance of certain aspects of the characters and open ended questions to allow subjects to speculate about events and outcomes.

An additional method that of selection from multi-choice responses would

seem to offer potential for the collection of quantifiable data about characters. This may have some of the disadvantages of false recognition paradigms (subjects may respond to the experimenter's ideas) but the plausibility problem will be avoided by asking subjects to select concepts that are 'likely' rather than asking them to make true/false judgements. The addition of a 'not applicable' choice will allow subjects to avoid a forced choice response.

### The experiment.

The experiment which will be described was designed to investigate the process of comprehension in a text which was well structured but not written in canonical order and which involved interaction between two protagonists. The process will be investigated by the analysis of predictive inferences made by subjects about events and by analysis of their more general inferences about the characters. Three main methods will be used:

1. Open-ended questions which allow the subjects to make hypotheses about events and to speculate about qualities of the characters.
2. Multi-choice questions which allow the subjects to select adjectives and nouns which describe the appearance and personality of the characters.
3. Rating scales which allow subjects to rate the importance of certain information for an understanding of the events.

The problems which surround the methodologies have been discussed.

Because of these problems a study was carried out to assess the suitability of the content of the individual questions and the utility and reliability of each method as a tool for collecting data .

### Method.

#### Materials.

The text used was the Garp extract described in Chapter Nine. It was divided into three sections (see Fig 10:1) Outcome, Background and Events.

#### Subjects.

Eighteen subjects participated in the pilot experiment. There were equal numbers of males and females. The subjects were mature students from the Open University who were in their first year of study reading a Social Sciences Foundation Course. The subjects were drawn from two successive weekly intakes of students at Summer School. All were English native speakers and were or had been in occupations which demanded a good level of reading skill. Their ages ranged from 23 to 43 years.

#### Design.

Six subjects (three male, three female) were assigned to each of three groups: one group for each condition. In each condition subjects were given different amounts of the same text.

Condition E: Text given to subjects in this condition contained the Events section only. Data collected in this condition will provide a base line measurement of predictions about the events.

Condition BE: The text which was given to subjects in this condition contained the Background and the Events sections (in that order). Data collected in this condition will allow the effect of the background information upon the events to be measured.

Condition OBE: The text which was given to subjects in this condition contained the Outcome, Background and Events sections (in that order). The data obtained in this condition will allow the effect of placing the outcome at the outset to be compared with the two other conditions, in both of which the outcome would have followed rather than preceded the

given text if canonical order had been used.

Subjects in all three conditions were given the same set of questions (see Appendix One).

### 1. Open ended questions.

(i) Four open-ended questions were asked about the event (Questions 1-4).

(ii) Two open-ended questions were asked about the characters who participated in the events (Questions 5 and 6).

(iii) One open-ended question was asked which might elicit responses about the characters or the events (Question 7).

### 2. Multi and fixed choice questions.

(i) A multichoice question about the characters. This allowed freedom of choice between two opposing characteristics and a 'not applicable' category (Question 8).

(ii) A fixed choice question about the setting and the experience of the characters in the situation (Question 9).

### 3. Rating scales.

(i) A set of five-point rating scales which allowed subjects to rate how important it was to an understanding of the events to have information about physical, social and personality characteristics of the characters, and about the time and location of the events.

ii) A further five-point rating scale was given which allowed subjects to rate their interest in reading the rest of the book (Questions 10 and 11).

In this pilot study simple comparisons of total and percentage of total scores will be made. If more than 50% of the subjects give a similar (in the case of open ended questions) or the same response then it will be considered that the question merits further investigation using larger numbers of subjects.

(The basis on which comparisons can be made in open ended questions will be discussed under the heading of Content Analysis in the Methods section).

Fig.10:1. Text used in Experiments 1 and 2.

**Outcome.**

Jenny Fields was arrested in Boston in 1942 for wounding a man in a movie theater.

**Background.**

It was shortly after the Japanese had bombed Pearl Harbour and people were being tolerant of soldiers because suddenly everyone was a soldier, but Jenny Fields was quite firm in her intolerance of the behavior of men in general and soldiers in particular.

**Events.**

In the movie theater she had to move three times, but each time the soldier moved closer to her until she was sitting against a musty wall, her view of the newsreel almost blocked by some silly colonnade and she resolved that she would not get up and move again. The soldier moved once more and sat beside her.

**Procedure.**

The subjects were tested in two sessions. There was a week's interval between the two tests. The procedure was the same for both groups. They were seated in a section of an otherwise empty lecture theatre. There was at least one empty seat between each pair of subjects.

The texts were attached to the question sheets and divided into two equal piles, each pile containing three texts from each condition. The contents of the pile were shuffled. The texts were distributed in the shuffled order and were placed face down in front of the subjects. The male subjects received texts from one pile, the females from the other.

Subjects were asked to write down on the blank page that faced them their sex, age, occupation (previous occupation if unemployed or housewife ) and any other Open University courses that they had taken. Following this they were told :

"The task that you are about to carry out involves reading part of a story. When you have read the story you will be asked to answer questions based on it. As you answer the questions you may find that you are giving the same information in a different way. Please give the information even if you find that this is so. This is not a timed task, so when you have finished other people may still be writing. If this is so will you please sit quietly until everyone else has finished.

When you turn over the papers in front of you the story will be on the first page. Read this at your own pace and when you have finished, turn over and read the instructions about answering the questions. If you wish to turn back and reread the text whilst you are answering the questions please do. Please do not discuss this task with the other subjects until after you have left the room.

Are there any questions? Please begin."

The subjects then carried out the task. The response sheets were collected when everyone had finished.

### Content Analysis.

The main aim of the analysis was to look at the levels of agreement between subjects in their selection of responses. Responses were examined to find those for which there was at least 50% agreement between subjects within a group. The following procedure was used to analyse the open ended questions (Q.1-7).

Questions 1-3 asked the subjects to summarise the events or to make hypotheses about events which preceded or followed the given events. The responses to these questions were analysed in terms of their wider context. That is where or when the subjects set the events.

Questions 4 and 7 asked about the wider context: the kind of background knowledge that the subjects would like or the wider issues which the text might cover. The responses to these questions were analysed in terms of the text content which they were based upon.

Questions 5 and 6 asked about the characters. Question 5 asked the subjects to spontaneously generate words to describe the characters: question 6 asked them to decide who would be the main character in the narrative. Those responses to question 5 which used the same noun or adjective were grouped together and counted. Synonymous terms such as "afraid" and "frightened" were counted in the same way as those which used the same words. In question six synonymous terms which indicated the same character were counted together and the most often used term (for example, Jenny, her, she) was used as the collective label.

## Results.

### 1. Open ended questions.

The responses are presented in three sections: (i) those which answer questions about events (Q1-4); (ii) those which answer questions about the characters (Q5,6) and (iii) responses to Q7. The tables show the categories of responses for which there is 50% or more agreement.

#### (i) Questions about the events: Q.1-4, see Table 10:1(i).

The analysis of responses to Q1 ("What is happening?") shows that within the groups more than 50% of the subjects in Conditions BE and OBE give responses which have similar content. However the responses agreed upon do not differ between groups. So, reading the outcome does not change substantially the content of the summaries produced.

Analysis of the responses to Q2 shows that within all the groups more than 50% of the subjects give responses which are similar in content. The content of the responses of subjects in Condition OBE differs from that of subjects in the other two conditions. Subjects who are given the Outcome attach significance to different parts of the text and relate it to a wider context when describing prior events.

The responses to Q3 show a high level of agreement (66.6% or more) within each group. However, there is a difference in content of the responses between groups. Subjects who are given the Outcome tend to make predictions which relate to events further on in time from those made by subjects in either of the other two conditions. In addition, they make predictions which are intrinsically more 'Interesting' (Schank, 1978). Subjects who are given background information (BE and OBE) focus on events which surround Jenny.

The responses made to Q4 show that only subjects in Condition (BE) had more than 50% of agreement about the content of their predictions. The topic of their interest 'was more information about her attitudes.' Subjects in Condition OBE, who had the same background information, focussed their



Table 10:1 (i). Events.

Condition.	% Agreement	Content.
Q1. "What is happening?"		
E	50%	Sequence of events in the cinema.
BE	66.6%	Sequence of events in the cinema.
OBE	66.6%	Sequence of events in the cinema.
Q2. "What might have happened prior to the events described?"		
E	100%	Events in cinema or queue.
BE	83.3%	Events in cinema or queue.
OBE	66.6%	Experiences related to her intolerance.
Q3. "What might happen next?"		
E	66.6%	Development of relationship.
BE	83.3%	Assertive action (her).
OBE	66.6%	Legal or psychiatric events (after her arrest).
Q4. "What in particular would you like to know more about?"		
E	33.3%	Reason why she moved seats.
BE	66.6%	Background to her attitudes.
OBE	50%	Background to her actions.

interest on other and more diverse topics.

(ii) Questions about the characters: Q5-6, see Table 10:1(ii).

The words given (Q5) to describe the soldier show more than 50% agreement on at least one of the words in each condition. Subjects who were given the Outcome show some difference in response content from both other groups in that they tend to see the soldier as being 'persistent.' This attribution may be made by these subjects as a means of linking the Events to the Outcome (those who did not know the outcome made no such attributions of responsibility to him).

The responses concerning Jenny (Q5) show no agreement in Condition BE and very little agreement in Conditions E and OBE. In Condition OBE if the two descriptions "frightened" and "afraid" are added together then there is noticeable agreement between subjects in this group. However, these descriptions may have been synonyms generated by the same subjects and in that case would not represent an increased level of agreement.

The results of Q6 are interesting in that when only the Events are read, subjects are equally divided in their choice of main character. The addition of Background or both Background and Outcome establishes Jenny as the main character. Subjects performing under the latter two conditions also differ in the label they predominantly use to identify her, even though her name, Jenny, is given in both the Background and the Outcome.

(iii) The wider issues: Q7. See Table 10:1(iii).

In relation to Q7, the level of agreement about what would be the more general issues discussed in the book did not exceed 50% for any group. However 50% can be regarded as a high level of agreement for responses to a question which allowed a very wide range of options. It is particularly interesting that subjects who were given the outcome made predictions based on an inference that she was psychiatrically disturbed. Such an inference cannot be

Table 10:1(ii). Characters.

Condition.	%Agreement	Words agreed.
Q5. Words generated to describe each character.		
a: (him)		
E	66.6%	Soldier.
BE	66.6%	Lonely.
OBE	66.6%	Soldier: Persistent.
b: (her)		
E	33.3%	Annoyed: Determined.
BE	0	-----
OBE	33.3%	Frightened :Afraid.
Q6. "Who do you think will be the central character in the narrative which follows the extract that you have read?"		
E	50%	Lady.
BE	83.3%	Her.
OBE	66.6%	Jenny.

Table 10:1(iii). Predicted wider issues.

Q7. "Based on the ideas in the opening passage, what wider more general issues do you think the rest of the book might encompass?"		
E	50%	Wartime relationships.
BE	50%	War and Society.
OBE	50%	Psychiatric disturbance.

directly related to text, but is consistent with this group's responses to Q3. While the addition of the Background has a modifying effect on the responses made to both this question and Q3, the further addition of the Outcome radically alters the interpretation of the text.

Multi and fixed choice questions.

(i) Responses to the multi-choice question: Q8, see Table 10:2(i).

The responses show that there was a high level of agreement within groups for all four sections of the question (8a him and her, and 8b him and her). However, there were some differences between the words each group agreed upon. When asked to select words which described the general aspects of the characters (Q8b) subjects who were given the Outcome (Condition OBE) selected some words which were the same as those selected under the other two Conditions (E and BE). However, when asked to select words which described the characters in the situation (Q8a) there was less overlap between words chosen by subjects in Condition OBE and subjects in the other two Conditions.

(ii) Responses to the fixed choice question: Q9, see Table 10:2(ii).

Generally, the responses show a high level of agreement within each group. However, while subjects in Conditions BE and OBE were certain (100%) that the movie theater was empty, subjects in Condition E were uncertain. Furthermore, while all groups agreed that the characters did not know each other and that this was not a new experience for him, subjects in Condition OBE were uncertain whether what happened was a new experience for her. Subjects in the other two conditions were certain but made opposing decisions: this with only the Events said 'Yes', those with the Background said 'No'.

Table 10:2(i). Multi-choice questions.

Condition	%Agreement (over 50%)	Words agreed upon.
8a. Ratings of the characters reactions to the situation described in the text.		
Him.		
E	100%	Assured.
	66.6%	Controlled.
BE	66.6%	Relaxed.
OBE	83.3%	Assured: Unafraid.
Her.		
E	100%	Tense.
	88.3%	Anxious.
	66.6%	Controlled.
BE	100%	Tense.
	88.3%	Anxious: Rational.
OBE	100%	Tense: Anxious.
	88.3%	Afraid: Excitable.
8b. Ratings of the characters more general personality traits.		
Him.		
E	100%	Friendly: Sociable.
	83.3%	Self confident: Self reliant.
BE	66.6%	Friendly.
OBE	66.6%	Self confident: Sociable.
Her.		
E	88.3%	Sensitive: Hostile.
BE	100%	Aloof.
	88.3%	Sensitive.
	66.6%	Insecure: Hostile
OBE	100%	Sensitive.
	88.3%	Insecure.

Table 10:2(ii). Fixed choice questions.

Condition	%Agreement.	Words agreed upon.
Q9: "Do you think that:"		
A. "The movie theater was fairly crowded or empty?"		
E	50%	Empty/ Full.
BE	100%	Empty.
OBE	100%	Empty.
B. "The characters already knew each other?"		
E	100%	No.
BE	100%	No.
OBE	100%	No.
C. "What happened was a new experience for him?"		
E	100%	No.
BE	66.6%	No.
OBE	66.6%	No.
D. "What happened was a new experience for her?"		
E	88.3%	Yes.
BE	66.6%	No.
OBE	50%	Yes/No.

Table 10:3

Items rated as being very important additional information for a better understanding of the text. Items included in the table are those having average ratings of less than 2. The rating scale ranged from 1 (Very Important) to 5 (Not Important).

Condition.	Additional Information (Mean rating).
Q10(i): Information about her.	
E	Lifestyle (1) Her reasons for going to the cinema alone.(1) Attractiveness.(1.1) Attitude: social expectations of herself.(1.2) Disposition. (1.5) Attitude: to men. (1.5) Attitude: to her own appearance.(1.6) Age. (1.6) Status: occupation.(1.9)
BE	Disposition.(1.2) Reasons for going to the cinema alone.(1.6) Attitude: to men.(1.6) Attractiveness.(1.6) Age.(1.6) Lifestyle.(1.8) Attitude: to women.(2)
OBE	Disposition.(1) Attitude: to men.(1) Reasons for going to the cinema alone.(1.1) Attitude:social expectations of herself.(1.2) Lifestyle.(1.5) Attitude: to own appearance.(2) Age.(2)

Table 10.3 (continued).

Condition.	Additional Information (Mean rating).
Q10(ii): Information about him.	
E	Attitude: to women (1.1) Attitude: to social expectations of himself. (1.2) Disposition.(1.5) Reasons for going to cinema alone.(1.6) Age.(1.8)
BE	Attitude: to women. (1) Reasons for going to cinema alone.(1) Disposition.(1.6) Features.(1.6) Lifestyle.(1.8)
OBE	Disposition (1.5) Attitude: to women. (1.5) Reasons for going to the cinema alone.(1.5) Age.(2) Attitude: to social expectations of himself. (2) Social status.(2)
Q10(iii): Information about the setting.	
E	----- (no high ratings) -----.
BE	The country.
OBE	The neighbourhood.

---

Q11. "How interested are you in reading the rest of the book?"

(1= very interested, 5=not at all).

Condition	Mean rating
E	3.83.
BE	2.83.
OBE	2.16.



### 3. Rating Scales: Q10-11, see Table 10:3.

The responses to Q10(i) and (ii) show a high level of agreement between the groups in the rating of items (about 'her' and about 'him') considered to be 'very important' in understanding the text. Responses to 10(iii) reveal agreement between subjects in Conditions BE and OBE about the kind of background information which would be useful. The greater the amount of information given in the text, the more specifically they identify the additional information they require.

The results for Q11 show an increase in interest as text is added (from Condition E to BE, and from BE to OBE). It is not possible at present to say whether this is due to the increase in information, the content of the additional information or the structural form of presentation. However, it was noted earlier that the results for Q3 ("What happened next?") suggested that the predicted content of the rest of the story is more 'interesting' (Schank, 1978) for Condition OBE than for either Condition BE or Condition E.

### Discussion.

The main aim of the experiment was to determine whether the design and the type of questions used were suitable for investigating the effects on comprehension of varying story order and content. Specifically, two main questions were asked:

1. Are the hypotheses stated by subjects who did not know the outcome of the story similar to each other or idiosyncratic?
2. Does placing the outcome at the beginning of the text change the kind of predictions made about the events in the story and the hypotheses made about the characters?

In general, the open-ended questions produced reliable data which related to the two questions above. This will be discussed in detail below. The open descriptions of characters (Q5), however, provided less useful measures. The main problem in the analysis of Q5 was the subjective problem of deciding

which responses were synonyms and which were not. Most decisions required speculation about what the subject might have meant by the use of particular words. Such decisions made on the basis of single words were too subjective and were also complicated by the fact that the same subject may have generated synonymous words to describe the character. In the end only use of the same terms was scored.

All but two of the descriptive words given to describe the soldier were drawn directly from the text. These two exceptions ("lonely" from Condition BE and "persistent" from Condition OBE) provide some support for the proposal that giving the Outcome changes the interpretation of the information which is given about the background and events. No further information is given about the soldier in the Outcome, but the information about his behaviour in the cinema which is described in the Events is interpreted quite differently. The words given to describe Jenny tend to be idiosyncratic in all conditions.

The multi-choice questions (Q8a and 8b) provided useful data. The responses indicated that there were differences within and between groups when subjects were asked to describe the characters' reactions to the situation, but the responses were remarkably similar when subjects were asked to indicate their more permanent characteristics. Further investigation with a larger sample is needed to confirm these differences. The rating scales provided reliable measures. Subjects in all three conditions mainly agreed about which aspects of the characters it was important to know more about. The information obtained from this question will be used to produce more precise open-ended questions about the characters for the next experiment. Rating scales will also be used as a more direct measure of the characters or the events.

1. Are the hypotheses stated by subjects who did not know the outcome of the story similar to each other or idiosyncratic?

Subjects in Condition BE had high levels of agreement on all but two

questions: Questions 5b and 7. Indeed, subjects in all three conditions tended to give idiosyncratic responses to these questions. The suitability of these questions for further investigation is raised again below. Subjects in Condition E also showed generally high levels of agreement. However, they gave idiosyncratic responses to Questions 1 and 6. These idiosyncratic responses only occurred in this Condition. Once Background information was added to both these questions subjects (Condition BE) produced responses which could be placed in a single category.

The results presented in all three tables show no general differences in agreement between subjects within these two conditions (E and BE). On the whole, they indicate that subjects do make ongoing inferences which are plausible given the content of the text, that they make these on the basis of small amounts of information and that the content of these inferences is similar amongst those who have read the same portion of text. Neither the responses to the open ended nor the multi-choice questions provide support for the idea that inferences are only made when the outcome indicates the relative importance of the ongoing text.

These results need to be confirmed by an experiment which uses a larger number of subjects. It will also be necessary to obtain some independent measure of reliability of the subjective categorisation of responses to open-ended questions.

**2. Does placing the outcome at the beginning of the text change the kind of predictions made about the events in the story and the hypotheses made about the characters?**

The results section shows that there is evidence to support the idea that ongoing inferences are made about events and about characters and that these are changed by the outcome. Placing the outcome at the beginning of the text does change the kind of inferences which are made and provides the reader with the basis for making more specific hypotheses.

When subjects were asked "what might have happened prior to the events described?" (Q2) the responses given by subjects in Condition OBE refer to a wider context ("experiences which are related to her intolerance ") than those responses given by subjects who did not have the Outcome. As the speculations made by subjects in Condition OBE are based on information given in the Background section, it would seem that background information has gained a significance when preceded by the Outcome which was not apparent to subjects given the Background without the Outcome.

When subjects were asked "what might happen next?" (Q3.) those who had read the Outcome first referred to events which would take place after her arrest. Subjects in the other two conditions referred to events which would immediately follow the Events section. Whilst it might be interesting to speculate about the wider context which was evoked by the Outcome text, it must be noted that for subjects who are given it, the word "next" in the question could refer either to events that happened after "the soldier moved once more" or after "Jenny Fields was arrested". Subjects in the other two conditions did not have this choice. Nonetheless, it is interesting to note that the subjects in Condition OBE made their predictions on the basis of the temporal order of the events rather than the linear order of the text. This may be attributed to the dramatic "Interestingness" of the Outcome or the effect of placing the Outcome at the outset (or both). This requires further investigation and will be explored in Experiment Five. In the context of the present experiment, the results point out the ambiguity of Question 3 and indicate the need for tighter control.

Results from all three types of questions (open-ended, restricted choice and rating scales) support the proposal that subjects who read the Outcome in addition to the Background and Events do make speculations about events and characters in the story which differ from those made by subjects who do not have the Outcome. However it is not clear whether this is because of the

content of the Outcome or its position in the text.

There were some similarities in the responses given by subjects in Conditions BE and OBE. For example, both groups having Background information focussed on events in which Jenny (rather than the soldier) was the subject (Q3), and both sets of subjects were certain that Jenny was the main character (Q6). This shows that subjects in Condition OBE answered some questions on the basis of background information and that the Outcome may have had no effect on these. Further investigations need to separate out the effects of the two pieces of information by the inclusion of a fourth condition: Outcome and Events (OE). This will allow the effects of the Outcome to be measured against two baseline conditions: that is (OE and OBE) vs (E and BE).

#### Conclusions.

The experiment has indicated that giving the Outcome changes the nature of predictions made about events which might have taken place prior to and beyond those given in the text. Open ended questions are a useful way of eliciting these predictions, but these need to be written in a clear and unambiguous manner (so that all group are responding to the same question). Some independent measure of the reliability of the categorisation of the responses needs to be made. Open ended questions about the characters need to be formulated more precisely so that different facets are tested independently. The uncertainty about the relative effects of Background and Outcome texts needs to be resolved by the inclusion of a further condition. The next experiment will incorporate these changes into the design.

CHAPTER ELEVEN.Experiment 2. The effect of Background and Outcome on Predictions about Events and Hypotheses about characters.Introduction.

The pilot experiment which was described in Chapter Ten investigated the effect of the outcome of a story on the kinds of predictions made. The results indicated that the effect of the outcome could be measured from responses which had been elicited by the use of open-ended questions. Open-ended questions which asked the subject to predict events produced some results which suggested that the outcome caused information in the text to be interpreted differently. However, other responses (particularly those which referred to characters) were similar to those given by subjects who had background but no outcome information.

Two particular issues are raised by these results. One concerns the design of the experiment and indicates the need to incorporate a condition which will enable the effects of the Background information to be separated out from those of the Outcome. The other issue concerns the collection of data about the characters. In the discussion which followed the pilot experiment it was pointed out that there was a need to ask more specific questions about the characters (and the events) so that it would be possible to quantify the data. Data about the characters which were collected by using the multi-choice question indicated that knowledge of the outcome changed some of the hypotheses which were made about the characters but left others unchanged. This experiment aims to explore which hypotheses about the characters are changed by knowledge about the Outcome and which are unaffected. Among those that are changed (if any) a distinction will be made between those which reflect the informational content of the outcome and those which are based on

information given in the Background and Events sections.

### Method.

#### Materials.

The material used was the same Garp extract as was used in the pilot experiment (see Fig.10:1)

#### Subjects.

48 subjects took part in the experiment. There were equal numbers of males and females. All were first or second year students from the Department of Psychology at the University of Hull. Their ages ranged 18 and 49 years (median = 22yrs). All were native speakers of English. None had previously read the text.

#### Design.

12 subjects (6 males and 6 females) were assigned to each of 4 groups: one group for each Condition. In each Condition subjects were given different amounts of the same text:

Condition E: Subjects were given the Events only (this was to produce a baseline measurement for the analysis);

Condition BE: Subjects were given the Background and the Events (in that order);

Condition OE: Subjects were given the Outcome and the Events (in that order);

Condition OBE: Subjects were given the Outcome, the Background and the Events (in that order).

All subjects were given the same set of questions (see Appendix 2.)

The four conditions constitute a 2x2 design (background/no background; outcome/no outcome). This will allow the effects of background and outcome information to be separated out in the analysis of the data.

The same three kinds of questions (see Appendix 2) were asked as in the pilot study:

1.Open-ended questions.

- a. Questions about the events (Q3).
- b. Questions about her (Q1, 4b, 6b, 7, 9, 8a, 8b).
- c. Questions about him (Q2, 4a, 6a, 8a, 8b, 9).

2.Restricted choice questions.

- a. Questions about him (Q8c, 8d, 10a).
- b. Questions about her (Q8c, 8d, 10b).
- c. Questions about the setting (Q12).
- d. Questions about them (Q11).

3.Rating Scales.(Q5.)

**Procedure.**

The subjects were divided into four groups prior to the experiment. They were seated in these groups in a large lecture theatre. Subjects who were given the least amount of information (Condition E) were seated at the front of the theatre. Those with most information (Condition OBE) were seated at the back. Those in Conditions BE and OE were seated in the separate groups between groups E and OBE. This grouping was arranged to prevent subjects from inadvertently obtaining additional information. Subjects were spaced to avoid collaboration.

Subjects were each given a questionnaire which was placed face upwards,so that they could read the instructions and see what information was required of them. The pieces of text were given separately and placed face downwards. The experimenter then read aloud the following instructions which were printed on the front of the questionnaire:

"Please read the following text. It is an extract from a published book. When you have read the passage,please answer the questions on the attached sheets. This is not a test of your memory, so feel free to refer to the



text while answering the questions. Some of the questions are fairly open-ended, others are more specific; do not worry if you have to give the same information twice. There are no right and wrong answers, you are being asked for your ideas and opinions. Make your answers as detailed as you can within the space provided. Please answer the questions as spontaneously as possible."

They were then told that there was no time limit for answering the questions and were asked to sit quietly when they had finished answering the questions until everyone else had finished and the questionnaires had been collected. They were then asked if they understood the instructions. Following this they were asked to fill in the requested information on the front of the form and then to read the text and answer the questions.

#### Open-ended questions: Content Analysis.

The open-ended questions (Q1-4, Q6 and Q7) were analysed in the following way. The content of each subject's answers was listed within each group. If a subject made more than one prediction then each one was listed separately. At this stage of the analysis it was possible to note the number of predictions made by:

- a. each subject;
- b. each group;
- c. the total sample.

Once the responses were listed it became possible to categorise them. Within each group, if subjects made the same response then these were grouped together. If the responses were similar in content (for example, "she hit him": "she slapped him") they were also grouped together. The content of the groups of responses allowed them to be merged into categories of behaviour or intention.

**The Categories.**

When the method of content analysis (described above) was applied to each of the Open ended questions, some of the categories derived were applicable to more than one question. These are described below with examples. For some sets of responses, these categories are too general and other more specific categories were used. These will be described with the particular questions to which they apply.

**(i).General Categories:**

**Cinema:** Reasons which were related to the film or the cinema in a general way. For example, "someone complained about them moving about", "so that he could see the screen".

**Evasive:** One or other of the characters react in such a way that they avoid the situation. For example, "he ran away", "she left the cinema".

**Positive:** He or she react in a positive way. For example, "she called the manager", "she chatted to him", "he struck back".

**Aggressive/Threatening:** He or she takes some action which is aggressive or threatening: "she stabbed him", "to menace her".

**Mood/Emotion/Feeling:** The reactions are internal; change of mood, emotion, feeling: "she felt afraid", "he felt remorse".

**Passive:** Things happen to the character rather than the character taking action (usually described by a passive construction). For example, "she was arrested", "he was injured."

**(ii). Categories used for individual questions.**

Q1. "She first moved to another seat before "the soldier moved closer to her".  
Why did she move?

The data collected from all subjects was listed as described in the Method of Content Analysis section. The responses to this question were placed in seven categories. These categories are described more fully in Appendix 3a. (Apart from the general cinema category, the other categories are peculiar to

Question 1).

Her attitude to this particular soldier.

Her attitude to soldiers in general.

Her feelings towards men.

Actions which were made in response to the soldier.

Because of the soldiers behaviour.

"Normal" cinema behaviour.

Other.

The number of responses made by each group in each category was scored.

In addition the percentage of total group responses which fell into that category was calculated (see Table 11:1).

Reliability of categorisation.

Ten judges independently assigned the responses given to Q1 to the categories described above. There was a high level of agreement between the author and the judges:  $\text{Lambda} = 0.81$  ( $r = 0.98$ ). This experiment is fully reported in Appendix 3.

Q2. Each time she moved "the soldier moved closer to her". Why did he?

The responses to this question were placed in four categories: two of the categories have been described above:

Cinema.

Emotion.

Two other categories discriminate between the positive actions in more detail as this general category was too broad to be useful.

**Sexual:** these were reasons which were concerned with women in general as well as Jenny in particular and included very mild statements such as "to chat her up" and much stronger expressions such as "he had sexual reasons". The criterion for inclusion was that there was some intention of further advances.

Casual: these were very low key reasons which expressed no further intention. For example, "to sit beside her", "to talk to her".

Q3. Describe in detail what you think happened after: "the soldier moved once more and sat beside her".

Two general categories were used: cinema and interactive . Only 22 responses (cinema 8, interactive 14) fell into these categories. The rest referred to actions by Jenny (83) and actions by the soldier (31). Further separate analyses were carried on those categories which referred to Jenny and those which referred to the soldier.

3a. The soldier.

The responses which described the soldier's actions were placed in three categories:

Sexual and

Casual, both of which are described in the content analysis of Question 2.

Threatening: which is described under the heading of general categories.

3b. Jenny.

The responses which described Jenny's actions were placed into three of the general categories:

Feelings.

Evasive.

Aggressive.

Q4. What were the consequences of your answer to (3) for:

a) the soldier and b) her?

The responses which described the consequences both for her and for the soldier fell into the same four general categories:

Mood.

Evasion.

Positive.

Passive.

Q6. What do you think were their reasons for going to the movie theater alone?

a) His and b) hers.

The responses to both parts of this question fell into the same four categories.

Cinema: This was described as a general category.

Others: others would not or could not go with him/her for various reasons.

Company: To seek company or to meet someone (same or opposite sex).

Other: Miscellaneous reasons which did not fit any other category.

Q7. Do you think that she had an occupation? If 'yes' what do you think it might be?

It was generally agreed by a majority of subjects in all groups that she had an occupation, but the range of occupations was so large (25 occupations were listed) that it was impossible to make meaningful categorisations from them.

Q8a) Estimates of the ages of the characters.

Although this was technically an open-ended question because there were no restrictions on the answers that the subjects might give, all answers were given within a single category (years) and there was no problem of content analysis.

Q8b) Ideas of the characters marital status.

Six categories were used in the responses to this question: single, married, divorced, widowed, engaged and separated.

Q9. Please describe what their attitudes might be to the following: a) the opposite sex; b) their own appearance:

These questions were only answered by some of the subjects and the variety of answers from those who did was so great that they could not be categorised.

Q12. Specify as precisely as possible when you think the events in the movie theater took place and where the movie theater was.

The specificity of responses was measured in isolation from the accuracy. Specificity of responses was measured on a 7 point scale.

a. Location.

- 1= "Anywhere".
- 2= Non-specific (a large city, near barracks).
- 3= Country.
- 4= State or county.
- 5= City.
- 6= Specific area of the city.
- 7= Specific place in the city.

b. Time.

- 1= Longspan (1900's).
- 2= Decade.
- 3= Part decade.
- 4= Year.
- 5= Year + month or season.
- 6= Date (date/month /year).
- 7= Time of day.

## Results.

### Open ended questions.

Factorial chi-square was used to analyse the differences between responses in each of the conditions.

Q1. She first moved to another seat before "the soldier moved closer to her".

Why did she move?

The results are summarised in Table 11:1. The expected frequency requirements for Chi-square meant that two of the categories had to be removed because of their low frequencies. These were "Because of the soldiers behaviour" and "Other". There is a significant interaction between the categories and the background ( Chi-square= 24.10, df 4,  $p < 0.01$ ). Without background information (B1) predictions are predominantly that she moved for normal cinema reasons and that her actions somehow indicate that she responded to the soldier. However with background information (B2) the predictions are predominantly concerned with reasons which are related to her attitudes to soldiers and men. The presence or absence of Outcome have no statistically significant effect on the distribution of responses in each category.

Placing the outcome at the beginning of the text does not change the kinds of predictions that are made about her behaviour in the cinema prior to the given events. The kind of predictions that are made are changed significantly by the addition of background material.

Examination of the content of the responses showed that those subjects who had read the Background text based their responses on that information. Subjects who did not have this information based their responses mainly on general world knowledge about cinemas. This is the major difference between the predictions. What is interesting is that the outcome information is not used in this way. One tentative conclusion which may be drawn from these results is that the Outcome information does not affect the subjects

## QUESTION 1

CATEGORY A:	C1		C2		C1		C2		TOTAL
	B1	%	B2	%	B1	%	B2	%	
	Condition (E)		(BE)		(OE)		(OBE)		
SOLDIER	1	5.88	1	5.88	4	33.33	2	12.5	8
SOLDIERS	3	17.64	6	35.29	1	8.33	8	50	18
MEN	0	0	7	41.17	0	0	3	18.75	10
ACTION	7	41.17	3	17.64	4	33.33	1	6.25	15
CINEMA	6	35.29	0	0	3	25	2	12.5	11
TOTALS	17		17		12		16		62

## AxB and AxC SUMMARY TABLES:

CATEGORY A:	B1	B2	C1	C2
	(No background)	(Background)	(No outcome)	(Outcome)
SOLDIER	5	3	2	6
SOLDIERS	4	14	9	9
MEN	0	10	7	3
ACTION	11	4	10	5
CINEMA	9	2	6	5
TOTALS	29	33	34	28

CHI-SQUARE:	Total =	39.2903	degrees of freedom =	19	<.01
A =	5.25807		degrees of freedom =	4	n.s.
B =	.258065		degrees of freedom =	1	n.s.
C =	.580645		degrees of freedom =	1	n.s.
AB =	24.0968		degrees of freedom =	4	<.01
AC =	4.09677		degrees of freedom =	4	n.s.
BC =	.258065		degrees of freedom =	1	n.s.
ABC =	4.74194		degrees of freedom =	4	n.s.



speculations about the motivations of characters. Background information on the other hand becomes highly relevant in such speculations. The validity of this conclusion can be examined by analysis of the responses to other questions, in particular Q2.

Q2. Each time she moved "the soldier moved closer to her". Why did he?

The results shown in (Table 11:2) again show that the distribution of responses to the question is significantly affected by the presence of background information (Chi-square 8.91, df 3,  $p < 0.05$ ). The main difference is that when there is background information (B2) the responses given are allocated mainly to the sexual motives category: without background information most responses are more evenly assigned between the casual and the sexual motives categories. As the Background text gives no extra information about the soldier, none of the responses are paraphrases of information given in the text. The predictions are constructions but there is no single piece of information in the Background text to which they are directly related.

The Outcome text makes no significant change to the distribution. This supports the tentative conclusion (made in the discussion of Q1) that Outcome information does not affect the speculations about the motivations of the characters.

Q.3. Describe in detail what you think happened after "the soldier moved once more and sat beside her."

The analysis of the content of Q3 was discussed in the Content Analysis section when it was stated that the responses were sub divided into two sections: those responses which referred to the soldier (these results are tabulated in Table 11:3 ) and those which referred to Jenny (tabulated in Table 11:4).

TABLE 11.2

## QUESTION 2

CATEGORY A:	C1		C2		B1	%	B2	%	TOTAL
	Condition (E)	(BE)	(OE)	(OBE)					
CINEMA	4	17.39	1	7.14	3	25	1	6.25	9
SEXUAL	8	34.78	7	50	3	25	7	43.75	25
CASUAL	10	43.47	3	21.42	4	33.33	2	12.5	19
EMOTION	1	4.34	3	21.42	2	16.66	6	37.5	12
TOTALS	23		14		12		16		65

## AxB and AxC SUMMARY TABLES:

CATEGORY A:	B1 (No background)	B2 (Background)	C1 (No outcome)	C2 (Outcome)
CINEMA	7	2	5	4
SEXUAL	11	14	15	10
CASUAL	14	5	13	6
EMOTION	3	9	4	8
TOTALS:	35	30	37	28

CHI-SQUARE:	Total = 27.8	degrees of freedom = 15	<.05
A =	9.52307	degrees of freedom = 3	<.05
B =	.384615	degrees of freedom = 1	n.s.
C =	1.24615	degrees of freedom = 1	n.s.
AB =	8.9077	degrees of freedom = 3	<.05
AC =	4.35385	degrees of freedom = 3	n.s.
BC =	2.6	degrees of freedom = 1	n.s.
ABC =	.784615	degrees of freedom = 3	n.s.

TABLE 11.3

## QUESTION 3

CATEGORY A:	C1		C2		B1	%	B2	%	TOTAL
	B1	%	B2	%					
Condition (E)	(BE)	(OE)	(OBE)						
SEXUAL	0	0	4	40	4	44.44	1	25	9
THREAT	1	12.5	0	0	3	33.33	3	75	7
CASUAL	7	87.5	6	60	2	22.22	0	0	15
TOTALS	8		10		9		4		31

## AxB and AxC SUMMARY TABLES:

CATEGORY A:	B1 (No background)	B2 (Background)	C1 (No outcome)	C2 (Outcome)
SEXUAL	4	5	4	5
THREAT	4	3	1	6
CASUAL	9	6	13	2
TOTALS:	17	14	18	13

CHI-SQUARE:	Total = 23.5807	degrees of freedom = 11	<.05
	A = 3.35484	degrees of freedom = 2	n.s.
	B = .290323	degrees of freedom = 1	n.s.
	C = .806452	degrees of freedom = 1	n.s.
	AB = .774193	degrees of freedom = 2	n.s.
	AC = 13.4194	degrees of freedom = 2	<.01
	BC = 1.58065	degrees of freedom = 1	n.s.
	ABC = 3.35484	degrees of freedom = 2	n.s.

TABLE 11.4

## QUESTION 3

CATEGORY A:	B1		C1		B1		C2		TOTAL
	Condition (E)	%	(BE)	%	(OE)	%	(OBE)		
FEELINGS	6	50	4	16.66	2	11.11	3	10.34	15
EVASIVE	6	50	20	83.33	5	27.77	5	17.24	36
AGGRESS	0	0	0	0	11	61.11	21	72.41	32
TOTALS	12		24		18		29		83

## AxB and AxC SUMMARY TABLES:

CATEGORY A:	B1 (No background)	B2 (Background)	C1 (No outcome)	C2 (Outcome)
FEELINGS	8	7	10	5
EVASIVE	11	25	26	10
AGGRESS	11	21	0	32
TOTALS:	30	53	36	47

CHI-SQUARE: Total = 77.9157	degrees of freedom = 11	<.01
A = 8.98795	degrees of freedom = 2	<.05
B = 6.37349	degrees of freedom = 1	<.05
C = 1.45783	degrees of freedom = 1	n.s.
AB = 4.36145	degrees of freedom = 2	n.s.
AC = 45.7108	degrees of freedom = 2	<.01
BC = .0120449	degrees of freedom = 1	n.s.
ABC = 11.0121	degrees of freedom = 2	<.01

Table 11:3 shows that there is a significant interaction between the Outcome and the categories (Chi-square = 13.42, df 2,  $p < 0.01$ ). Without Outcome the predictions were that the soldier would behave in a casual manner, chat to her or discuss the film. With the Outcome information, the content of the predictions changes and the main hypotheses are that he will make some sexual or threatening move.

The results shown in Table 11:4 refer to predictions about Jenny. Again there is a significant interaction between the Outcome and the categories (Chi-square = 45.71, df 2,  $p < 0.01$ ). Without the Outcome it is predicted that she would have behaved in an evasive manner, leaving the cinema or moving seats again. When the Outcome is given then the predictions mainly hypothesise that she would have behaved in an aggressive manner; for example, stabbing him or hitting him with an umbrella. These results also show a significant interaction between the Background, Outcome and Categories (Chi-square = 11.01, df 2,  $p < 0.01$ ). In the condition where the subjects have Background and no Outcome (BE), the majority of responses predict that she would evade the situation, when Outcome is added to the Background (OBE) then the predictions are that she will behave aggressively.

When they are given the same information (that is text which includes the Outcome) and asked the same question subjects make two main types of prediction (this happens in all conditions). One is that she will behave in an aggressive manner, the other is that he will act in a way which is threatening or which anticipates a sexual encounter. The prediction about her behaviour can be associated with the information in the outcome IF the subjects have already made the inference that she did wound the man and that the given events describe the incident which lead up to the arrest. The prediction about the soldier is a different interpretation of his behaviour in the given Events and this is in some way related to the content of the Outcome. If it is to be assumed that the subjects were trying to make sense of the text, the

TABLE 11.5

## QUESTION 4a

CATEGORY A:	C1		C2		B1		B2		TOTAL
	B1	%	B2	%	B1	%	B2	%	
Condition (E)	(BE)		(OE)		(OBE)				
MOOD	10	55.55	8	40	2	11.76	2	9.09	22
EVASION	0	0	6	30	1	5.88	1	4.54	8
POSITIVE	4	22.22	3	15	4	23.52	1	4.54	12
PASSIVE	4	22.22	3	15	10	58.82	18	81.81	35
TOTALS	18		20		17		22		77

## AxB and AxC SUMMARY TABLES:

CATEGORY A:	B1 (No background)	B2 (Background)	C1 (No outcome)	C2 (Outcome)
MOOD	12	10	18	4
EVASION	1	7	6	2
POSITIVE	8	4	7	5
PASSIVE	14	21	7	28
TOTALS:	35	42	38	39

CHI-SQUARE: Total = 68.6623	degrees of freedom = 15	<.01
A = 22.5844	degrees of freedom = 3	<.01
B = .636364	degrees of freedom = 1	n.s.
C = .012987	degrees of freedom = 1	n.s.
AB = 4.81819	degrees of freedom = 3	n.s.
AC = 34.1169	degrees of freedom = 3	<.01
BC = .116883	degrees of freedom = 1	n.s.
ABC = 6.37661	degrees of freedom = 3	n.s.

TABLE 11.6

## QUESTION 4b

CATEGORY A:	C1		C2		TOTAL				
	B1	%	B2	%					
Condition (E)			(BE)		(OE)		(OBE)		
FEELINGS	12	60	14	70	1	6.25	7	33.33	34
EVASIVE	0	0	4	20	0	0	2	9.52	6
POSITIVE	3	15	2	10	3	18.75	1	4.76	9
PASSIVE	5	25	0	0	12	75	11	52.38	28
TOTALS	20		20		16		21		77

## AxB and AxC SUMMARY TABLES:

CATEGORY A:	B1 (No background)	B2 (Background)	C1 (No outcome)	C2 (Outcome)
FEELINGS	13	21	26	8
EVASIVE	0	6	4	2
POSITIVE	6	3	5	4
PASSIVE	17	11	5	23
TOTALS:	36	41	40	37

CHI-SQUARE: Total = 73.2338	degrees of freedom = 15	<.01
A = 29.8571	degrees of freedom = 3	<.01
B = .324675	degrees of freedom = 1	n. s.
C = .116883	degrees of freedom = 1	n. s.
AB = 7.2078	degrees of freedom = 3	n. s.
AC = 33.8052	degrees of freedom = 3	<.01
BC = .324675	degrees of freedom = 1	n. s.
ABC = 1.5974	degrees of freedom = 3	n. s.

predictions about his behaviour could involve two kinds of inference being made: one which would connect him with the man who was injured, and one which would establish a causal link between the wounding and his behaviour.

Q4. What were the consequences of your answer to (3) for: a) the soldier; b) for her.

Table 11:5 shows the results of the responses which were given about the soldier and Table 11:6 the results of the responses made about her. Both tables show a significant interaction between the Outcome and the categories (for Table 11:5: Chi-square= 34.1169, df 3,  $p < 0.01$ ; and for Table 11:6: Chi-square= 33.8052, df 3,  $p < 0.01$ ).

The differences between the categories is the same in the hypotheses about him and those about her. Without the Outcome it is mainly predicted that any consequences for him and for her will be a change which is made in their internal state (that is a change in their emotional state or in the way they think or feel). When the outcome is given the kinds of predictions made change and the consequences which are predicted for both characters are that things will happen to them beyond this episode. Both are depicted as passive beings: he will be taken to hospital, she will be arrested, charged or jailed. The predictions made when the outcome is given rest on the inferences that he is "the man" and that this is the incident in which she wounded him.

Q6. What do you think were their reasons for going to the movie theater alone? a) His and b) hers.

The results shown in Table 11:7 summarise the responses which were made about the soldier's reasons for going to the movie theater alone. There is no significant interaction between the kind of text given and the categories of responses. There is a significant difference in the number of responses which are placed in each category (Chi-square = 19.40, df 3,  $p < 0.01$ ). The main difference is that more responses are placed in the category Company than in



the other categories. This is the same in all conditions and it is generally agreed that he went to the movie theater to find company (either male or female).

The results which are summarised in Table 11:8 refer to the responses which are made about her reasons for going to the movie theater alone. Again there is no significant interaction between the kinds of text given and the categories which are predominantly used. The responses fall mainly into one category. The difference in numbers of responses in each category is statistically significant (Chi-square = 20.15, df 3,  $p < 0.01$ ). The category which is most used in all conditions is cinema. It is generally hypothesised that she went to the movie theater to see the film, newsreel or news. Table 11:8 also shows that the number of responses made when there is no outcome is significantly larger (Chi-square = 20.15, df 1,  $p < 0.05$ ) than when there is an outcome. However the distribution of the responses among the categories is not significantly different.

The additional information given with background and outcome does not change the kind of hypotheses that are made about why each of the characters went to the movie theater alone. In all conditions the predictions remain in the same predominant category: he went alone to seek company, she went alone to see the film.

#### Q8a) Their Ages (Jenny and the soldier).

Subjects estimations of their ages was examined using a three factor analysis of variance: where factor A was identity (Jenny vs the soldier), factor B presence or absence of background text and factor C presence or absence of Outcome text. The results showed an interaction between the Background factor and their identities ( $F = 5.69$ , df 1,44,  $p < 0.05$ ): without background he is rated older than her - 26.6 as opposed to 24.5; with background information she is rated older - 26.9 for her versus 24 for him.

TABLE 11.7

## QUESTION 6a

CATEGORY A:	C1		C2		B1	%	B2	%	TOTAL
	B1	%	B2	%					
	Condition (E)		(BE)		(OE)		(OBE)		
CINEMA	3	13.63	3	12.5	4	20	5	25	15
OTHERS	4	18.18	2	8.33	4	20	4	20	14
COMPANY	11	50	10	41.66	9	45	9	45	39
MISC	4	18.18	9	37.5	3	15	2	10	18
TOTALS	22		24		20		20		86

## AxB and AxC SUMMARY TABLES:

CATEGORY A:	B1 (No background)	B2 (Background)	C1 (No outcome)	C2 (Outcome)
CINEMA	7	8	6	9
OTHERS	8	6	6	8
COMPANY	20	19	21	18
MISC	7	11	13	5
TOTALS:	42	44	46	40

CHI-SQUARE: Total = 26.3721	degrees of freedom = 15	<.05
A = 19.3953	degrees of freedom = 3	<.01
B = .0465116	degrees of freedom = 1	n.s.
C = .418605	degrees of freedom = 1	n.s.
AB = .976741	degrees of freedom = 3	n.s.
AC = 3.58139	degrees of freedom = 3	n.s.
BC = .0465117	degrees of freedom = 1	n.s.
ABC = 1.90698	degrees of freedom = 3	n.s.

TABLE 11.8

## QUESTION 6b

CATEGORY A:	C1				C2				TOTAL
	B1	%	B2	%	B1	%	B2	%	
Condition (E)	(BE)				(OE)				(OBE)
CINEMA	11	45.83	10	37.03	8	53.33	9	56.25	38
OTHERS	4	16.66	3	11.11	2	13.33	4	25	13
COMPANY	4	16.66	7	25.92	3	20	2	12.5	16
MISC	5	20.83	7	25.92	2	13.33	1	6.25	15
TOTALS	24		27		15		16		82

## AxB and AxC SUMMARY TABLES:

CATEGORY A:	B1 (No background)	B2 (Background)	C1 (No outcome)	C2 (Outcome)
CINEMA	19	19	21	17
OTHERS	6	7	7	6
COMPANY	7	9	11	5
MISC	7	8	12	3
TOTALS:	39	43	51	31

CHI-SQUARE: Total = 28.8293	degrees of freedom = 15	<.05
A = 20.1463	degrees of freedom = 3	<.01
B = .195122	degrees of freedom = 1	n.s.
C = 4.87805	degrees of freedom = 1	<.05
AB = .0975531	degrees of freedom = 3	n.s.
AC = 1.65854	degrees of freedom = 3	n.s.
BC = .0487814	degrees of freedom = 1	n.s.
ABC = 1.80488	degrees of freedom = 3	n.s.

There is also a statistically significant interaction between the Background and the Outcome factors ( $F=5.597$ ,  $df 1,44$ ,  $p>0.05$ ). When Outcome information is added to the Background information both characters are rated older. In condition BE his mean age is 22.7 and hers is 24.3 while in condition OBE they are 25.9 and 29.5 respectively.

Q8b) The marital status of the characters.

Six categories were used in by subjects. The results are not tabulated because one category (single) was used predominantly for both characters in all conditions. The expected frequencies for the other five categories were too small to be used in a factorial chi square analysis.

---

Table 11:9

Mean Ratings of Specificity Scores in each Condition.

	Conditions			
	E	BE	OE	OBE
a) Location	4.25	4.08	5.42	5.58.
b) Time	4.0	3.9	5.0	4.25.

  

	B1	B2	C1	C2
	(No Background)	(Background)	(No Outcome)	(Outcome)
a) Location	4.83	4.83	4.17	5.50
b) Time	4.50	4.08	3.92	4.63.

Q12. Specificity of time and location.

These results show the specificity not the "factual" accuracy of responses.

- a) Subjects are more specific about the location of the events as outcome is given. Adding background makes no difference to the specificity of responses about the location.
- b) Subjects responses about time are less specific when Background information is added. They are more specific when there is an Outcome than when there is not.

Responses which are made when the Outcome is given are more specific about the setting than those made without Outcome information. Background information does not produce the same effect. The specificity of responses is either unchanged (location) or less specific (time). The responses may or may not be accurate. Measurements concerning the accuracy of responses were not made because subjects in the different conditions had differing amounts of specific information about the setting. In addition to this subjects who had outcome information had the two levels of events (those in the Events section and those in the Outcome ) to discuss.

It may be concluded that an increase in the specificity of responses about the setting is related to the Outcome but not to the Background information. It would seem that subjects who have the Outcome information are more certain of the setting of the events regardless of the accuracy of their ideas. This may be because they infer that the outcome relates to the events.

Restricted Choice Questions.Q8c. Level of self-esteem: High, average or low.

Table 11:10 shows the level of self-esteem selected in relation to Jenny. There are no statistically significant differences between the categories selected either in terms of main effect or in interactions. Overall she is generally rated as having high self esteem.

TABLE 11.10

## QUESTION 8 (SELF-ESTEEM) JENNY

CATEGORY A:	C1		C2		B1 (OE)	%	B2 (OBE)	%	TOTAL
	B1 Condition (E)	%	B2 (BE)	%					
HI	6	50	6	50	4	33.33	6	50	22
AV	3	25	2	16.66	6	50	4	33.33	15
LO	3	25	4	33.33	2	16.66	2	16.66	11
TOTALS	12		12		12		12		48

## AxB and AxC SUMMARY TABLES:

CATEGORY A:	B1 (No background)	B2 (Background)	C1 (No outcome)	C2 (Outcome)
HI	10	12	12	10
AV	9	6	5	10
LO	5	6	7	4
TOTALS:	24	24	24	24

CHI-SQUARE: Total = 7.5	degrees of freedom = 8	n. s.
A = 3.875	degrees of freedom = 2	n. s.
AB = .875	degrees of freedom = 2	n. s.
AC = 2.375	degrees of freedom = 2	n. s.
ABC = .375	degrees of freedom = 2	n. s.

TABLE 11.11

## QUESTION 8 (SELF-ESTEEM) SOLDIER

CATEGORY A:	C1		C2		B1	%	B2	%	TOTAL
	B1	%	B2	%					
	Condition (E)		(BE)		(OE)		(OBE)		
HI	6	50	3	25	5	41.66	5	41.66	19
AV	4	33.33	7	58.33	5	41.66	6	50	22
LO	2	16.66	2	16.66	2	16.66	1	8.33	7
TOTALS	12		12		12		12		48

## AxB and AxC SUMMARY TABLES

CATEGORY A:	B1	B2	C1	C2
	(No background)	(Background)	(No outcome)	(Outcome)
HI	11	8	9	10
AV	9	13	11	11
LO	4	3	4	3
TOTALS:	24	24	24	24

CHI-SQUARE: Total = 10.5	degrees of freedom = 8	n. s.
A = 7.875	degrees of freedom = 2	<.05
AB = 1.625	degrees of freedom = 2	n. s.
AC = .125	degrees of freedom = 2	n. s.
ABC = .875	degrees of freedom = 2	n. s.

Table 11.11 shows the level of self esteem selected in relation to the soldier. The results show that there is a statistically significant difference between the frequencies in each category (Chi-square = 7.875, df2,  $p < 0.05$ ). In all conditions subjects tended to rate his self esteem as high or average. There is no significant interaction between the text given and the category selected. Addition of Background and Outcome information makes no statistically significant difference to the results.

Q8d. Subjects estimations of the characters' attractiveness.

Tables 11:12 (Jenny) and 11:13 (the soldier) show that in all conditions the average category is selected more frequently than the others. This pattern is the same for both subjects in all four conditions.

Q10. Words which were selected to describe him and her a) in the situation; b) more generally (personality traits).

The responses to these questions were analysed as four separate analyses. A factorial chi square was used for each analysis. The results which are summarised in Table 11:14 show that there is a statistically significant difference in the frequency with which the words were selected (Chi-square = 75.1092, df 11,  $p < 0.01$ ). Some words were frequently selected by subjects in all conditions to describe the way that Jenny reacted to the situation these were tense and anxious. The words which were infrequently chosen in all conditions were relaxed and assured. The results also show a statistically significant interaction between the Outcome and the words selected (Chi-square = 30.3361, df 11,  $p < 0.01$ ). Without the Outcome she is frequently described as controlled and rational; with Outcome she is described as afraid and excitable.

The words which were selected to describe his reaction to the situation are summarised in Table 11:15. These results show a statistically significant difference in the frequency of words selected to describe the reactions of the



TABLE 11.12

QUESTION 8 (ATTRACTIVENESS) JENNY

CATEGORY A:	B1		C1		B1		C2		TOTAL
	Condition (E)	%	(BE)	%	(OE)	%	(OBE)		
V	4	33.33	1	8.33	3	25	3	25	11
A	8	66.66	11	91.66	7	58.33	8	66.66	34
N	0	0	0	0	2	16.66	1	8.33	3
TOTALS	12		12		12		12		48

AxB and AxC SUMMARY TABLES:

CATEGORY A:	B1 (No background)	B2 (Background)	C1 (No outcome)	C2 (Outcome)
V	7	4	5	6
A	15	19	19	15
N	2	1	0	3
TOTALS:	24	24	24	24

CHI-SQUARE:	Total = 36.5	degrees of freedom = 8	<.01
	A = 32.375	degrees of freedom = 2	<.01
	AB = 1.625	degrees of freedom = 2	n. s.
	AC = 1.625	degrees of freedom = 2	n. s.
	ABC = .875	degrees of freedom = 2	n. s.

TABLE 11.13

## QUESTION 8 (ATTRACTIVENESS) SOLDIER

CATEGORY A:	C1				C2				TOTAL
	B1	%	B2	%	B1	%	B2	%	
	Condition (E)		(BE)		(OE)		(OBE)		
V	1	8.33	2	16.66	2	16.66	0	0	5
A	8	66.66	10	83.33	5	41.66	7	58.33	30
N	3	25	0	0	5	41.66	5	41.66	13
TOTALS	12		12		12		12		48

## AxB and AxC SUMMARY TABLES:

CATEGORY A:	B1 (No background)	B2 (Background)	C1 (No outcome)	C2 (Outcome)
V	3	2	3	2
A	13	17	18	12
N	8	5	3	10
TOTALS:	24	24	24	24

CHI-SQUARE: Total = 28.5	degrees of freedom = 8	<.01
A = 20.375	degrees of freedom = 2	<.01
AB = 1.625	degrees of freedom = 2	n.s.
AC = 5.375	degrees of freedom = 2	n.s.
ABC = 1.125	degrees of freedom = 2	n.s.

## QUESTION 10a: JENNY

WORDS A:	C1				C2				TOTAL
	B1	%	B2	%	B1	%	B2	%	
Condition (E)	(BE)		(OE)		(OBE)				
RELAXED	1	1.63	0	0	1	1.63	0	0	2
TENSE	11	18.03	11	19.29	10	16.39	11	19.64	43
AFRAID	5	8.19	3	5.26	9	14.75	8	13.55	25
UNAFRAID	4	6.55	5	8.77	1	1.63	2	3.38	12
COMPOSE	5	8.19	4	7.01	4	6.55	1	1.69	14
EXCITABL	3	4.91	3	5.26	6	9.83	8	13.55	20
ASSURED	3	4.91	2	3.5	3	4.91	1	1.69	9
ANXIOUS	9	14.75	10	17.54	9	14.75	9	15.25	37
RATIONAL	8	13.11	6	10.52	4	6.55	4	6.77	22
EMOTIONA	2	3.27	4	7.01	6	9.83	7	11.86	19
CONTROLL	10	16.39	9	15.78	1	1.63	4	6.77	24
UNCONTRLD	0	0	0	0	7	11.47	4	6.77	11
TOTALS	61		57		61		59		238

## AxB and AxC SUMMARY TABLES:

WORDS: A	B1 (No background)	B2 (Background)	B1 (No outcome)	B2 (Outcome)
RELAXED	2	0	1	1
TENSE	21	22	22	21
AFRAID	14	11	8	17
UNAFRAID	5	7	9	3
COMPOSE	9	5	9	5
EXCITABL	9	11	6	14
ASSURED	6	3	5	4
ANXIOUS	18	19	19	18
RATIONAL	12	10	14	8
EMOTIONA	8	11	6	13
CONTROLL	11	13	19	5
UNCONTRLD	7	4	0	11
TOTALS:	122	116	118	120

CHI-SQUARE: Total = 111.311      degrees of freedom = 47      See tables

A =	75.1092	degrees of freedom =	11	<.01
B =	.151261	degrees of freedom =	1	n. s.
C =	.0168067	degrees of freedom =	1	n. s.
AB =	3.57983	degrees of freedom =	11	n. s.
AC =	30.3361	degrees of freedom =	11	<.01
BC =	.0168067	degrees of freedom =	1	n. s.
ABC =	2.10085	degrees of freedom =	11	n. s.

## QUESTION 10a: SOLDIER

WORDS A:	C1				C2				TOTAL
	B1 Condition (E)	%	B2 (BE)	%	B1 (OE)	%	B2 (OBE)	%	
RELAXED	4	6.77	8	14.54	6	10.34	6	12.5	24
TENSE	6	10.16	3	5.45	3	5.17	3	6.25	15
AFRAID	3	5.08	4	7.27	3	5.17	1	2.08	11
UNAFRAID	7	11.86	6	10.9	8	13.79	6	12.5	27
COMPOSED	5	8.47	4	7.27	3	5.17	0	0	12
EXCITABLE	3	5.08	6	10.9	6	10.34	7	14.58	22
ASSURED	9	15.25	8	14.54	6	10.34	8	16.66	31
ANXIOUS	2	3.38	3	5.45	2	3.44	1	2.08	8
RATIONAL	6	10.16	3	5.45	2	3.44	4	8.33	15
EMOTIONAL	3	5.08	4	7.27	7	12.06	3	6.25	17
CONTROLD	8	13.55	3	5.45	5	8.62	4	8.33	20
UNCONTRLD	3	5.08	3	5.45	7	12.06	5	10.41	18
TOTALS	59		55		58		48		220

## AxB and AxC SUMMARY TABLES:

WORDS A:	B1 (No background)	B2 (Background)	C1 (No outcome)	C2 (Outcome)
RELAXED	10	14	12	12
TENSE	9	6	9	6
AFRAID	6	5	7	4
UNAFRAID	15	12	13	14
COMPOSED	8	4	9	3
EXCITABLE	9	13	9	13
ASSURED	15	16	17	14
ANXIOUS	4	4	5	3
RATIONAL	8	7	9	6
EMOTIONL	10	7	7	10
CONTROLD	13	7	11	9
UNCONTRLD	10	8	6	12
TOTALS:	117	103	114	106

CHI-SQUARE: Total = 48.8                      degrees of freedom = 47                      See tables

A = 27.7455	degrees of freedom = 11	<.01
B = .890909	degrees of freedom = 1	n.s.
C = .290909	degrees of freedom = 1	n.s.
AB = 5.54545	degrees of freedom = 11	n.s.
AC = 7.45455	degrees of freedom = 11	n.s.
BC = .163636	degrees of freedom = 1	n.s.
ABC = 6.70911	degrees of freedom = 11	n.s.

soldier to the situation (Chi-square = 27.7455, df 11,  $p < 0.01$ ). The words frequently selected in all conditions were *unafraid, assured, relaxed*. The infrequently selected words were *anxious, afraid*. There are no statistically significant interactions between the conditions and the selection of words. The kind and amount of text given did not change the words chosen to describe his reactions.

The frequency with which words were chosen by all subjects to describe her more enduring personality characteristics (see Table 11:16) differs in a statistically significant way (Chi-square = 50.39, df 11,  $p < 0.01$ ). She is most frequently described as *aloof, intolerant, self reliant*. She is least frequently described as *sociable, tolerant and dependent*. There is a significant interaction between the presence of Background text and the words selected (Chi-square = 23.875, df 11,  $p < 0.05$ ). Without the Background text the words *sensitive and friendly* are selected with the highest frequency. With the Background text *tough-minded and hostile* were most frequently selected.

The frequency with which individual words were selected to describe his more stable personality characteristics is shown in Table 11:17. There is an overall difference in the frequency with which words were selected by all subjects (Chi-square = 60.2667, df 11,  $p < 0.01$ ). He was most frequently described as *sociable, friendly and self-confident*. He was least frequently described as *aloof, sensitive and hostile*. There is no significant interaction between the texts given and the words selected. The addition of Outcome and/or Background text does not significantly change the words selected to describe his personality.

There is some consistency with which subjects in all conditions select words to describe both the reactions to the situation and the more enduring personality characteristics of both the soldier and of Jenny. These descriptions are independent of Background and Outcome information. There is no significant change in the frequency distribution of the words selected to

QUESTION 10b: JENNY

WORDS A:	C1				C2				TOTAL
	B1	%	B2	%	B1	%	B2	%	
Condition (E)			(BE)			(OE)	(OBE)		
ALOOF	7	13.2	11	18.03	7	12.06	9	17.3	34
SOCIABLE	2	3.77	0	0	3	5.17	1	1.92	6
SELF-CON	4	7.54	5	8.19	5	8.62	4	7.69	18
INSECURE	6	11.32	6	9.83	6	10.34	5	9.61	23
TUF-MIND	1	1.88	6	9.83	3	5.17	4	7.69	14
SENSITIV	8	15.09	4	6.55	6	10.34	3	5.76	21
TOLERANT	5	9.43	0	0	3	5.17	1	1.92	9
INTOL	5	9.43	11	18.03	7	12.06	9	17.3	32
FRIENDLY	5	9.43	1	1.63	4	6.89	0	0	10
HOSTILE	3	5.66	8	13.11	4	6.89	8	15.38	23
SELF-REL	5	9.43	8	13.11	5	8.62	7	13.46	25
DEPENDEN	2	3.77	1	1.63	5	8.62	1	1.92	9
TOTALS	53		61		58		52		224

AxB and AxC SUMMARY TABLES:

WORDS A:	B1 (No background)	B2 (Background)	C1 (No outcome)	C2 (Outcome)
ALOOF	14	20	18	16
SOCIABLE	5	1	2	4
SELF-CON	9	9	9	9
INSECURE	12	11	12	11
TUF-MIND	4	10	7	7
SENSITIV	14	7	12	9
TOLERANT	8	1	5	4
INTOL	12	20	16	16
FRIENDLY	9	1	6	4
HOSTILE	7	16	11	12
SELF-REL	10	15	13	12
DEPENDEN	7	2	3	6
TOTALS:	111	113	114	110

CHI-SQUARE: Total = 79.4286	degrees of freedom = 47	See tables
A = 50.3929	degrees of freedom = 11	<.01
B = .0178571	degrees of freedom = 1	n. s.
C = .0714286	degrees of freedom = 1	n. s.
AB = 23.875	degrees of freedom = 11	<.05
AC = 1.75001	degrees of freedom = 11	n. s.
BC = .875	degrees of freedom = 1	n. s.
ABC = 2.44642	degrees of freedom = 11	n. s.

QUESTION 10b: SOLDIER

WORDS A:	C1				C2				TOTAL
	B1	%	B2	%	B1	%	B2	%	
Condition (E)	(BE)				(OBE)				
ALOOF	0	0	0	0	2	3.92	1	2.43	3
SOCIABLE	9	18.75	8	20	9	17.64	5	12.19	31
SELF-CN	8	16.66	5	12.5	7	13.72	5	12.19	25
NSECURE	3	6.25	4	10	1	1.96	3	7.31	11
TUF-MIND	3	6.25	2	5	6	11.76	5	12.19	16
SENSITIV	3	6.25	1	2.5	1	1.96	2	4.87	7
TOLERANT	7	14.58	4	10	3	5.88	1	2.43	15
INTOL	1	2.08	3	7.5	3	5.88	5	12.19	12
FRIENDLY	9	18.75	7	17.5	7	13.72	5	12.19	28
HOSTILE	0	0	0	0	4	7.84	1	2.43	5
SELF-REL	3	6.25	4	10	6	11.76	5	12.19	18
DEPENDEN	2	4.16	2	5	2	3.92	3	7.31	9
<b>TOTALS</b>	<b>48</b>		<b>40</b>		<b>51</b>		<b>41</b>		<b>180</b>

AxB and AxC SUMMARY TABLES:

WORDS A:	B1 (No background)	B2 (Background)	C1 (No outcome)	C2 (Outcome)
ALOOF	2	1	0	3
SOCIABLE	18	13	17	14
SELF-ON	15	10	13	12
NSECURE	4	7	7	4
TUF-MIND	9	7	5	11
SENSITIV	4	3	4	3
TOLERANT	10	5	11	4
INTOL	4	8	4	8
FRIENDLY	16	12	16	12
HOSTILE	4	1	0	5
SELF-REL	9	9	7	11
DEPENDEN	4	5	4	5
<b>TOTALS:</b>	<b>99</b>	<b>81</b>	<b>88</b>	<b>92</b>

<b>CHI-SQUARE:</b>	<b>Total = 84</b>	<b>degrees of freedom = 47</b>	<b>See tables</b>
	A = 60.2667	degrees of freedom = 11	<.01
	B = 1.8	degrees of freedom = 1	n.s.
	C = .0888889	degrees of freedom = 1	n.s.
	AB = 7.00002	degrees of freedom = 11	n.s.
	AC = 12.4445	degrees of freedom = 11	n.s.
	BC = .0222224	degrees of freedom = 1	n.s.
	ABC = 2.37775	degrees of freedom = 11	n.s.

describe the soldier (both his reaction to the situation and his more general personality characteristics) when the additional text is given. Judgements of his reactions and of his personality are relatively unaffected by Background and Outcome information. Overall he is judged to be self confident, sociable and friendly character who reacts to the situation in an assured unafraid and relaxed manner.

Jenny is generally (and throughout the conditions) described as having a tense and anxious reaction to the situation. This does not change. However words which are selected to describe Jenny do change when there is additional text. When there is no Outcome information she is frequently described as being controlled and rational. When Outcome information is added she is most frequently described as excitable, uncontrolled and afraid. Her more general personality characteristics are described in all conditions as aloof, intolerant and self reliant. These remain unchanged. Without Background information she is described as being sensitive and friendly. When Background information is given this is changed to tough-minded and hostile.

The results indicate that the Outcome information does not change the assessment of basic characteristics of either character. However the situational expression of Jenny's character is changed by the Background information. Furthermore Outcome information changes the description of her reactions to the situation, but not the soldier's.

Q11. Did the characters already know each other?

While some of the subjects in Condition E thought that the characters might have known each other, the addition of the Background or the Outcome information led to almost total agreement amongst subjects that they did not know each other.



Table 11:18

Frequency of responses to question 11.

Category.	Conditions.			
	E	BE	OE	OBE
Yes.	4	0	0	0
No.	8	12	11	12
Don't Know.	0	0	1	0
Total	12	12	12	12

---

Rating Scales.

Q5. To what extent do you consider the following to be responsible for what you have described as happening, and why?

There is no clear difference in the mean ratings of the responsibility of what happened for him, for her or for the situation as a function of the four conditions. However there is some indication that the presence of Outcome reduces the 'responsibility' of the situation. Furthermore it seems that the addition of any information (Outcome, Background or both) increases the level of responsibility attributed to her and reduces that attributed to him.

His mean responsibility rating is higher than either hers or that attributed to the situation. The situation is rated as being more responsible for what happened than Jenny is.

Table 11:19

Mean ratings of responses to Q5. (4 = Complete Responsibility, 0 = None.)

	Conditions			
	E	BE	OE	OBE
a) Jenny	1.4	1.75	1.8	1.8
b) the soldier	3.1	2.75	2.9	2.6
c) the situation	2.4	2.1	1.4	2.4

  

	No background	Background	No Outcome	Outcome
	a) Jenny	1.6	1.78	1.6
b) the soldier	3	2.7	2.93	2.75
c) the situation	1.9	2.25	2.25	1.9

---

### Discussion.

This experiment investigated the effect of the Outcome on the text by comparing the effects of Outcome and Background text upon responses given to a set of questions. The questions asked the subjects to predict events and to make hypotheses about the characters in the text. Open ended and restricted choice questions and rating scales were used to elicit responses. The responses to the open ended questions were placed in categories on the basis of a content analysis. The reliability of this analysis was tested by asking ten independent judges to assign responses to the chosen categories. The results showed a high level of agreement among the judges and between the judges and the author.

The Outcome information mainly affected predictions concerned with the key event (that is what happened after "he moved once more and sat beside her") and the consequences of this for him and her. Changes were also made in

the restricted choice question which described her reactions to the situation.

An examination of the predictions which are made when the Outcome is added shows that some of these are structural and create a frame into which subsequent information can be slotted (these are the predictions which look to events beyond those in the movie theater). These will be discussed in greater detail. Others are predictions which are based on additional information in the Outcome, such as those which suggest that when the soldier moves and sits beside her she will act in an aggressive manner. These predictions appear to be based upon inferences (made from information in the text) that she had carried out the wounding for which she was arrested; that the events in the text are those which lead up to the wounding incident and that the soldier is "the man" who was wounded. Other predictions rely on further chaining of the inferences: that his behaviour prior to the move (sexually motivated or threatening) is a cause of the incident and that her reaction (afraid and excitable) is linked to his behaviour. This agrees with the proposals made by Goetz (1979) that events in the text will be reinterpreted and gain a different kind of importance when the outcome is known.

Those responses which were affected by the addition of Background material concerned the events which happened prior to her moving about the cinema and certain of Jenny's more enduring personality characteristics. These are similar in type to those discussed above in relation to the Outcome, that is they were based on and could be linked by inference to information in the Background text.

Other predictions which are made by subjects who have Outcome information are linked to the position of the Outcome episode in the structure of the story. The consequence (both for him and for her) of whatever happened after "the soldier moved once more and sat beside her" is that things will happen to him/her. These predictions differ from those which are made in the other two Conditions in which the characters were in control and making things happen.

Subjects in all conditions were looking for a consequence which would end the story. Those who had no Outcome information were predicting consequences which would end the story immediately after the events in the movie theater.

Subjects who were given the Outcome looked towards a set of events beyond her arrest. The wounding and the arrest are the consequences of the inferred events which link the given Events to the Outcome and the subjects go on to predict the consequences of the arrest and wounding. It would seem that the cues given by Q4 and the Outcome information indicate that the story has a structure which goes beyond this incident and it is to this that the predictions are aimed.

The experiment supports the idea that some inferences are made only after the outcome has been read. It does not support the implication that these are the only inferences that are made or that no ongoing inferences are made. Some responses were unaffected by the differences in the Conditions. These tended to be responses which were concerned with information relevant to the situation prior to the events. For example, the reasons why they went to the cinema alone, his level of self esteem, their level of responsibility for what happened, his permanent personality characteristics and reactions to the situation and some of her permanent characteristics and reactions to the situation. These relied on stored world knowledge rather than information given explicitly in the text and these seem to have been made on the basis of information in the events, although there is no clear link between the information given and the predictions and hypotheses made. These links would seem to be akin to what Reder (1979) described as 'elaborations'. In this experiment these are not generally idiosyncratic, rather those predictions that were made show there is a high degree of similarity of content for all subjects and that the elaborations are in some way situationally related. These 'elaborations' are unlike 'scripts' (Schank 1975) (which allow the reader to assume that certain events will occur) rather they allow the subjects to

predict how characters will behave given a 'usual' situation.

Such predictions and hypotheses are remarkably similar in all conditions although those characteristics which are directly contradicted by information in the text do differ. The results suggest that inferences (other than those which are made to link the events in the outcome to events which happened earlier in time) are made in an ongoing fashion. The content analysis of the predictions suggests that these ongoing inferences create a stereotype of the characters as a framework for the comprehension of their behaviour in ongoing events in the text and for the anticipation of what is to come. The characters are modified with the input of additional information. One example of this can be taken from results of the pilot experiment where it was shown that subjects who had Background information thought that what happened in the movie theater was not a new experience for Jenny whereas those who had only Events information had thought that it was. This is a clear example of the way in which the Events relating to Jenny's behaviour are reinterpreted when subjects are given more information about her character and are able to create a new stereotype on the basis of this. This frame created from stored world knowledge about characters in a situation can be directly compared to the fact structure of Kintsch (1979). Unlike Kintsch's example which depended on very specialised knowledge, this experimental data has shown that the stereotypes can be created from a very general level of situational knowledge.

The experiment has shown that some forward looking inferences or predictions can be made in response to each part of the text. The process is not in any way confined only to the Outcome. Goetz (1979) proposed that inference making was constrained by the importance for the outcome of events in the story. This would mean that inferences are made after the outcome has been read. Goetz implied that neither ongoing inferences nor inferences which did not relate to the outcome were made. Models of story grammars have been criticised in earlier chapters of this thesis for the same reason, that is

because they have implied that such processing would only take place when the text had been read completely.

Those models and theories which rely on the story having been processed in its entirety make the assumption that inferences can only be made on the basis of what has occurred and that inference making is restricted to these backward looking inferences. The current experiment has shown that inferences are made which create a stable image of the characters and that these allow the subject to predict and hypothesise about the characters future actions (that is they are able to make forward looking inferences) on the basis of this image. These inferences are then modified by the information in the text. Some inferences or predictions are the same whether the subjects were given the Events, the Background or the Outcome first. Others change according to specific information which is given in the text and the inferences which are based upon this.

The functional significance of the Flashback sequence was not investigated directly in the experiment but speculations about its function can be made on the basis of the results obtained. It would seem that placing the Outcome at the beginning allows the reader to create a structure for the incoming text and to make predictions or forward looking inferences about the story from this. This allows a framework to be created from the outset and it is used in two ways: to accept incoming information and to make predictions which go beyond the time span of the scenario described in the text. It has been demonstrated in both the pilot and the main experiment that given an appropriate probe the predictions will go beyond the time span in both a forward and backward direction. In addition to the predictions which are related to the structure of the story, certain inferences are made about the characters on the basis of the content of the outcome. This function is not solely related to the Outcome but is shown also to be a function of Background information.

It is proposed that the structural predictions are functional at outset because the reader becomes involved in making hypotheses about future events and does not make predictions which quickly become irrelevant as the story unfolds. When the outcome is placed at the end (in canonical order) the text has to be reinterpreted to accommodate this information. By this time the reader may well have made irrelevant or incorrect predictions. Placing the Outcome at the end would seem to be functional when the outcome really is the end of the story and the information either confirms the predictions made by the reader or it is so unexpected that it causes a reflection and reinterpretation of the whole plot (as, for example, in detective stories). In the initial stages of the story the strategy of placing the outcome of a set of events at the outset allows the author some measure of control over the kinds of predictions made so that those predictions which are useful and relevant are made and the reader quickly becomes involved with the story.

It may be concluded from this experiment that the relationship between the structure and the content of a story is rather more complex than was suggested by Bower (1976) and Thorndyke (1977) and that the disruption of canonical order is not always disfunctional. It has been proposed that a restructuring of the time order can assist the reader by constraining the formation of predictions. The use of this strategy would seem to be one way in which the author has control over inference making. Those inferences which are concerned with the content of a story (those which were described by Goetz 1979) are only one kind of prediction that is made. These, along with other information in the text are based on inferences which link the information in a text in a backward direction when relevant new input makes this necessary. Predictions and hypotheses about the characters are made on the basis of stored knowledge about situations and these may or may not be radically changed by the order in which the information is presented. Placing the Outcome at the beginning allows certain information which has been explicitly given to

be used. This experiment has not clearly pointed out whether the 'image' which has been created is sustained or altered by subsequent information.

It has indicated that the reader uses both the structure and the content to make predictions about the plot and the characters. Predictions about characters are based on world knowledge about characters in a situation and predictions about events are based upon stored knowledge about story conventions and about actual temporal order of events.



CHAPTER TWELVEThe effects of sequential order on  
the processing of "Interesting" content.Introduction

Experiments One and Two investigated the effect of prior knowledge of Outcome on forward looking inferences. From these results it was concluded that there are effects due to the content of the Outcome and others attributable to the structural information that it conveys. The results have also provided the basis for speculations about the possible effects of placing the Outcome at the beginning of a text, thus creating a "Flashback" sequence in a text.

Two related issues arise from the experiments. The first was raised in the discussion of Experiment One and concerned the relative importance of content and structure in the making of inferences when a dramatic Outcome is placed at the beginning of an comparatively undramatic text. The second issue concerns the effects of the position of the Outcome on the processing of the text. Specifically, is processing affected by whether the Outcome is presented after the Events (so that the text is presented in canonical order) or at the outset (so that the text is presented in a flashback order)?

A theory which was presented by Schank (1978) gives a basis for the examination of these issues. He proposed that certain concepts were intrinsically, absolutely interesting and that it was "Interestingness" which controlled the processing of text. He argued that readers selectively processed texts and made inferences about these selected parts. The basis on which items in the text were selected was that of "Interestingness". This was defined according to several criteria. The overriding criterion was that of absolute interestingness. Those concepts which are absolutely interesting are DEATH, DANGER and "other absolute interests are: POWER; SEX; MONEY (in large

quantities); DESTRUCTION; CHAOS; ROMANCE; DISEASE and many other other concepts and issues of this type" (Schank, 1978, p15).

Schank proposed that whenever these absolutely interesting concepts are present we use them to think about and to make inferences about the consequences and causes of their occurrence. Although there are other criteria for interestingness, the processing of absolutely interesting concepts overrides the processing of all others. Some concepts are so intrinsically interesting that readers always make inferences about them.

Schank's (1978) claim that "Interestingness " was the basis for selective processing of text rests on his theory that a reader cannot process all the input in the time which he/she takes to read it because the number of inferences which it would be possible to make on the basis of the content plus those inferences generated by the interactions between the inferences could not be carried out in the time. A further reason is that the interactions between the inferences involved in the process would lead to a combinatorial explosion. Schank uses computer processes as an analogy for human processes. The result of a combinatorial explosion for a computer would mean that it had not the capacity or resources to process the inferences made when the interacting procedures resulted in a geometric progression of increases in the number of inferences. The results of this for a human can only be speculative but it is supposed that the reader's processing mechanisms would be overloaded with the inferences generated so that he/she would not be able to discriminate those inferences which were pertinent to the content nor would further processing be possible. Perhaps the reader would never get beyond processing the initial input. Schank proposed that this situation was avoided by the reader selectively processing the text and that the basis on which he/she did this was primarily the intrinsic interestingness of the concepts. Those concepts which were interesting were processed and the inferences which were necessary for this processing were made.

In order to clarify whether a reader will choose to process texts on the basis of "Interestingness" of the content and to investigate whether "Interestingness" of content is the basis for the selection of concepts for processing when texts which have the same content are given a different structure the following series of investigations were devised. The investigations will explore the interestingness of concepts in isolation, in context and when they are a part of a narrative text which is presented in flashback and canonical order.

### EXPERIMENT THREE.

#### An experiment to investigate the concept of absolute interestingness.

##### Introduction.

The first of the series of experiments explores the idea that some concepts are absolutely interesting, by asking subjects to rate 16 sentences on an interestingness scale. The sentences are made up of concepts which according to Schank's criteria are either absolutely interesting or not interesting. Average ratings for each sentence and mean ratings for "Interesting" and "Not Interesting" sentences are to be calculated to discover if there is any discrimination between them.

##### Method

###### Subjects.

There were 12 subjects (6 male, six female). All students at the University of Hull. None had taken part in Experiments 1 and 2.

###### Materials.

The following 16 sentences were each written on a separate file card: according to the criterion of absolute interestingness defined by Schank (1978) eight were "interesting" and eight were "not interesting".

**"Interesting":**

1. Riot shields and water cannon were becoming commonplace.
2. Breaking the sound barrier was an achievement for science but not for humanity.
3. The noise outside was deafening.
4. The windows rattled as the sound exploded above them.
5. With superhuman strength she tore the steel jaws apart.
6. She decided that she would never carry a gun again.
7. It was as she was walking through the trees that Joan saw the vultures circling over the body.
8. It was as she was walking through the trees that Joan stumbled over the body.

**"Not Interesting"**

9. She looked around to see if her friend was sitting at a nearby table but she could not see her.
10. She tried to catch the waitresses eye but she was unsuccessful.
11. As she waited for the waitress to bring her the menu, Mary realised that her watch must have stopped half an hour earlier.
12. As she waited for the waitress to bring her the menu, Mary realised that the time she had allowed for eating her meal was getting shorter and shorter.
13. The roses were absolutely magnificent.
14. The rain and cold were suddenly forgotten.
15. The sunshine was dazzling.
16. The bees hummed as the sun shone on the garden.

**Procedure.**

A seven point scale was indicated by a series of seven marker cards placed at equal intervals along an otherwise empty desk top. The card at the extreme left was marked "interesting", the centre card was marked "neither

interesting nor uninteresting" and the card at the extreme right was labelled "uninteresting". These were in place before each subject came into the room.

Each subject carried out the task separately. The 16 cards on which the sentences were typed were shuffled and given to the subject. Each subject was asked to :-

"Rate each sentence separately in terms of how interesting you find it. The more interesting you find it, the further to the left you must place it; the less interesting you find it, the further to the right you must place it. Do not rate the sentences against each other but consider each one in terms of its own interestingness".

The instructions for placing the cards on the scale were repeated and/or clarified if necessary. One subject asked for a clarification of interestingness. He was told "interesting in the sense of you wanting to know more about it".

The subjects were then left alone to complete the task. When the task was completed the given ratings for each sentence were recorded.

### Results

There was a clear difference between the overall mean ratings of the "Interesting" and "Not Interesting" sentences. Those which were defined as "Interesting" according to Schank's criteria are rated more interesting by the subjects. Table 12:1, however, shows that "interestingness" is not a dichotomous concept but that the sentences are rated on a continuum. Within both the "Interesting" and "Not Interesting" groups of sentences there is a difference between the mean ratings of sentences. This is the case even when sentences with similar content are rated (for example, between 1 and 2, 3 and 4 and so on between pairs). The main variation is for the "Not Interesting" sentences.

Table 12:1.

Mean ratings for sentences.

"Interesting"		"Not Interesting"	
Sentence	Mean.	Sentence	Mean.
1.	5	9.	2.75
2.	4.84	10.	1.75
3.	3.67	11.	2.75
4.	5.	12.	3.58
5.	5.75	13.	2.6
6.	5.33	14.	4.34
7.	5.83	15.	3.17
8.	5.66	16.	2.84
Mean	5.1313	Mean	2.98

---

### Discussion

The results have show that some ideas are clearly rated as more interesting than others when they are read in isolation. The sentences which were chosen to discriminate between the interesting and less interesting ideas were selected according to the criteria for absolutely interesting ideas laid down by Schank (1978). However, the "Not Interesting" sentences are 'Less Interesting' rather than "Not Interesting". Schank did propose that in absence of something absolutely interesting readers would make "Interesting" whatever was available. He did however maintain that absolute "Interestingness" was preferred and that it the basis on which readers selectively processed text:

"When items of competing interest are around, we would choose the more interesting to infer (i.e. think about)." (Schank, 1978, p28).

The question of whether those ideas which are rated as interesting are those which are preferentially processed remains. Two main aspects of the question are:

1. Does reading sentences in context change their level of "interestingness"?

2. Does "interestingness" form a basis for selective processing?

The two aspects require separate investigations.

#### EXPERIMENT FOUR.

#### Ratings of interestingness of sentences by subjects who have previously read the sentences in context.

##### Introduction

This experiment aims to establish whether having read sentences in context changes the rated levels of absolute interestingness found when the sentences were read in isolation. What needs to be established is whether subjects who have read the ideas in context would still rate the ideas as interesting in the same way as those who have only read the ideas in isolation. This measurement is required as a base line for comparison with the preferred processing task if a comparison is to be made between rated interestingness and processing of ideas.

##### Method.

###### Subjects.

Thirty six subjects took part in the experiment. Twenty four subjects (12 male and 12 female) all students at the University of Hull, had read the sentences in context in two different conditions in a prior experiment (to be described as Experiment Five) and in addition the ratings obtained from the 12 subjects in Experiment Three were used as a control.

###### Design.

The sentences were rated for interest by three groups each of 12 subjects (six male, six female). Group A had previously read the sentences in texts which were presented in one order. Group B had previously read the sentences in texts presented in a different order (the reverse order of Group A). Group C were the control group.

The reliability of the ratings between the groups was tested using Spearman-Brown split-half reliabilities. Similarities between group ratings were compared using Spearman correlations.

#### Materials and Procedure.

These were the same as described in Experiment Three.

#### Results.

Table 12.2 shows the mean rated interestingness of the individual sentences for Groups A, B and Control. On average all three groups rate the "Interesting" sentences as more interesting than the "Not Interesting" sentences. There are differences in the mean ratings of less interesting sentences between the Control Group and the groups who have already read the sentences in context. Subjects who have already read the sentences in context rate the "Not Interesting" sentences more interesting than those who have only read the sentences in isolation. However, reading in context does not affect the "Interesting" sentences.

The consistency of sentence ratings was assessed by correlating the mean ratings of half the subjects with those of the other half. The coefficients of consistency were adjusted using the Spearman-Brown Prophecy formula. All groups show a high level of consistency (see Table 12.3), indicating that subjects are rating the interestingness of the sentences in very similar ways.

Correlations between the mean ratings for the sixteen sentences in each group were computed for all pairs of groups (see Table 12.4). All the correlations are high indicating that the groups are rating the sentences in the same way. The smallest correlation is between Groups A and B and suggests that the difference in the story order in some way affects the ratings. As seen in Table 12.2, the ratings of the "Not Interesting" sentences show the greatest difference between these two groups.



Table 12:2

Mean interestingness of sentences 1- 8 (Interesting)  
 9-16 (Not Interesting)  
 by Groups Control, A and B

<u>Interesting</u>	1	2	3	4	5	6	7	8	Mean
Control (n = 12)	5	4.84	3.67	5	5.75	5.3	5.83	5.66	5.13125
Gp A (n = 12)	5.16	4.416	4.083	5.5	4.5	5.333	5.5	6.42	5.113625
Gp B (n = 12)	4.91	5.91	3.75	4.91	5.25	5.33	5.33	5.58	5.12125
<u>Not Interesting</u>	9	10	11	12	13	14	15	16	Mean
Control (n = 12)	2.75	1.75	2.75	3.58	2.66	4.34	3.17	2.84	2.98
Gp A (n = 12)	4.00	2.92	4.42	4.16	3.83	4.5	4.5	4.25	4.0725
Gp B (n = 12)	3.33	2.75	3.75	3.58	2.83	4.16	3.16	3.41	3.3715

Interestingness Scale    8 = high

1 = low

Table 12.3.

Consistency of sentence ratings for each group.

Groups	Split-half correlations.	Spearman-Brown Prophecy
A	$r = 0.53$	0.81
B	$r = 0.65$	0.79
Control	$r = 0.88$	0.93

---

Table 12.4.

Intergroup correlations of ratings of Interestingness.

Groups	Correlation .
A/B	$r = 0.79$
Control/A	$r = 0.89$
Control/B	$r = 0.93$

---

### Discussion.

The results show that those sentences which are defined as absolutely interesting have a similar average rating of "Interestingness" whether the sentences have been read in isolation or have previously been read in context. The average ratings of those sentences which are "Not Interesting" differs between the groups being higher for Groups A and B. Within group reliability of ratings shows that subjects are rating the sentences in similar ways. The between group correlations show that between the groups, the mean "Interestingness" ratings for each sentence are strongly related, although the correlations between the Control Group and each of the groups who read the sentences in context are much higher than the correlation between the latter two groups.

Schank's proposal that certain ideas are "absolutely interesting" is supported when subjects are asked to rate the interestingness of each

sentence. However, the "Not Interesting " ideas acquire "Interestingness " as a function of context. The experiment has not explored whether or not the "Interestingness" of ideas is instrumental in the preferential processing of ideas in context. Experiment Five will test this idea.

#### EXPERIMENT FIVE.

#### The effects of "Interestingness" on sentence choice in text.

#### Introduction.

Schank (1978) proposed that "Interesting" ideas were processed in preference to "Not Interesting" ones. The two preceding experiments have shown that although some ideas are clearly more interesting than others all sentences are rated as having some level of "Interestingness". Experiment Four established that subjects agreed closely about the "Interestingness" level of sentences and that reading them in context raised the level of "Interestingness" of those defined as being "Not Interesting". The data collected in Experiment Four also provided an "Interestingness" rating of individual sentences by each subject. The experiment which follows uses these individual ratings by comparing the rated level of interestingness of ideas with the subject's choice of which he/she would like to read more about.

Choice behaviour will be examined at the beginning of a text and at a stage which is half way through so that the effects of context on choice can be investigated. The effects of "Flashback" will also be investigated by presenting the same texts in canonical and reverse orders.

If the "Interestingness" of the content is the most important factor in deciding what will chosen for further processing it is expected that readers will always choose to read a text which follows what they have rated as being the more interesting of a pair of sentences.

If placing the outcome at the beginning of the text leads to a difference in the way in which text is processed, then it is expected that there will be

differences in the selection of content when the text is presented in flashback order. If it is not important then there will be no difference in selection whichever order the text is presented in.

### Method

#### Subjects.

The 24 subjects used as Groups A and B in Experiment Four had previously participated in this experiment.

#### Materials.

The texts (see below) and instructions were presented using a Vector Graphic System B Microcomputer. A key board and a Visual Display Unit (VDU) was available to the subjects. A pen and a pad of paper were provided for subjects to write the answers to questions presented after each text.

#### Design.

Four texts (see Appendix 4) were designed with the same format:

- a. A pair of sentences to choose between;
- b. A single sentence;
- c. A second pair of sentences to choose between;
- d. A single sentence.

For each text, subjects had first to choose one of a pair of sentences. Following the choice a single sentence was presented followed by a second pair to choose from. After the second choice, a final sentence was presented. The presentation was so designed that which ever choice was made the same intervening sentence was displayed as the text which followed that choice. Following each text a set of questions was presented which acted as a diversion from the repetition of the selection task.

The sixteen sentences described in Experiment Three were paired in consecutive pairs (1 and 2, 3 and 4 etc.) according to the similarity of their

content. Two pairs (that is four sentences) were incorporated into each of the four texts. One pair was placed at the beginning of the text (the Outset) the other pair halfway through (in context). The sentences were combined within the texts in such a way that two of the texts contained the "Interesting" (according to Schank's, 1978 criteria) sentences and two of the texts contained the "Not Interesting" (or mundane) sentences. An additional variable was that the texts were written in such a way that two of the texts (one interesting, one mundane) had a time sequence and thus could be described as having a narrative structure and the others were descriptive and had no time structure. This produced a 2x2 design:

	Interesting	Mundane
Time sequence	Text 3 (sentences 9-12)	Text 2 (sentences 13-16)
No time sequence	Text 4 (sentences 5-8)	Text 1 (sentences 1-4)

The four texts were presented to each subject. The order in which the text was presented was reversed for half the subjects: Group A received them in one order, Group B in the other. This allowed the same sentences to be read both with and without context (at the beginning of the text and half way through). This also allowed those texts (2 and 3) which had a time sequence to be read in canonical (Condition A) and flashback (Condition B) orders.

The order of presentation of the texts was varied between subjects to control for practice effects according to a Latin Square design.

#### Procedure.

Subjects were tested singly. Each was randomly assigned to Condition A or B and to one of the orders of text presentation within that condition. The order and condition was keyed into the computer and the texts presented

accordingly.

The subjects were seated at the computer and their familiarity with the equipment discussed. Those who were not familiar with microcomputers were assured that the task only required the subject to press the A, B and RETURN keys and to read the text on the screen. All subjects were told that the task and the instructions would all be presented on the VDU but that the experimenter would remain in or near to the room during the experiment. They were then told that the experiment was not timed and that the task and the written answers to the questions should be done at their own pace. Each was given a pen and paper on which to answer the questions. The condition and presentation order were keyed in and the initial instructions appeared on the screen:

"You will be asked to read four short stories which will be displayed on the screen. With each story you will be asked at the outset to choose which version of the text (A or B) you prefer to read. Further along you will be asked to make another choice. You will make each choice by pressing the keys A or B on the keyboard. When you have finished each story you will be given a task based on the version that you have chosen to read.

Press RETURN for further instructions."

At this point the experimenter checked that the subject was confident about the use of the keyboard, the subject pressed the RETURN key and further instructions about answering the questions appeared:

"The task will be to write answers to the following:

1. Write the next sentence.
2. Outline the way in which the story might continue.
3. Give a title to the text.
4. Summarise what you have read.

The answers are to be written in the booklet provided. Write any comments

about the texts, the questions or your answers on the last page. The instructions will be repeated at appropriate points in the task. When you are ready for the first story press RETURN to start."

The texts were then presented in the preselected order and sequence. The subject's choice of text was recorded by the computer. Between each task the subject answered questions relating to the text.

When all four texts were read and the questions answered, the subject was asked to carry out the "Interestingness" rating experiment described in Experiment 4. The results were based on these ratings.

#### Data analysis.

The difference between interestingness ratings for each pair of sentences (Experiment Four) given by each individual subject was compared to his/her choice between the sentences in that pair in the text. The balance of "Interestingness" within each pair of sentences in isolation was scored first of all. Each member of the pair was assigned either the same level of "Interestingness" or one of the pair was rated as more interesting than the other. The subject's choices in the reading task were then examined to assess the frequencies with which (a) the more interesting sentences were chosen (b) the less interesting sentences were chosen. For those pairs rated of equal interest the frequency with which each were chosen was determined.

The scores were summarised as: choices made when sentences rated equally interesting (=); choice of sentence rated as most interesting (+); choice of sentence rated least interesting (-). Total numbers in each choice category (=, +, and - scores) were calculated for each text and for each position in the text. This was done for texts presented in both orders. Scores were also calculated for the Order effect. That is the choice of sentences was separately calculated for each text in each order. This included choices made both at the outset and in the context.

The scoring for sentences 3 and 4 in Text 1 is given as an example:

Group A: Canonical (at Outset).

Balance	Choice	
	Sentence 3	Sentence 4.
3=4	2	1
3>4	5	0
3<4	2	2
Total category scores: 3=; 7+; 2-.		

---

Group B: Flashback (in Context).

Balance	Choice	
	Sentence 3	Sentence 4.
3=4	2	1
3>4	5	0
3<4	1	3
Total category scores: 3=; 8+; 1-.		

---

### Results.

The results (Table 12.5) show the relationship between rated interestingness and choice of sentence when all scores are calculated. Overall (Table 12.5a) those sentences which have been assigned the highest level of interestingness are most often selected for further reading ( 80 for those rated highest, 61 for those rated lowest). When the results are separated into those choices which were made at the outset (12.5b) and those which were made in context (12.5c) it is shown that there is a strong indication that rated interestingness is related to sentence choice made at the outset: those sentences which were rated as most interesting were chosen 42 times; those rated least interesting 27. Whereas when the choice is made in context interestingness has a much reduced effect (38 and 34).



The results suggest that "Interestingness" makes an impact at the outset of a text but that if the same sentences are encountered in context the impact of "Interestingness" is much reduced and some other factors affect the selection of text for processing.

---

Table 12.5.

Balance of levels of rated interestingness and choice of sentences: Both conditions; all four texts.

	Choice Category		
	+	-	=
a) Both choices (n=192):	80	61	51
b) Choice at outset:	42	27	27
c) Choice in context:	38	34	24

---

When the texts are presented so that the narrative text is in canonical order (Group A :Table 12.6 ) there is an overall effect of rated interestingness: that is, those sentences which are rated as most interesting are selected most often (48 as opposed to 29). When the choice effects are separated out there is a clear effect of "Interestingness" on the initial choice (27 of the most interesting sentences are chosen as opposed to 9 of the less interesting). However when the choice is made in context the interestingness does not relate to the choice (21 and 20).

Table 12.6.

Rating of interestingness and choice of sentences in all four texts presented in canonical order (Group A).

	Choice Category		
	+	-	=
a) Both choice positions (n=96):	48	29	19
b) Choice at outset:	27	9	12
c) Choice in context:	21	20	7.

---

The effects of rated interestingness on the choice of sentence when the texts are presented in reverse order (Group B) are shown in Table 12.7. It is clear that "Interestingness" has no clear effect upon the overall choice of sentences (30 high rating choices as opposed to 34 low rating ones). A similar result is obtained when the scores are separated out into choice in first position (at outset) and choice in context. The choice in context (c) does show that there are more choices of the more interesting sentences than the less interesting ones, but this is not a clear difference (17:14).

---

Table 12.7.

Rating of interestingness and choice of sentence when all four texts are presented in reverse order (Group B).

	Choice Category		
	+	-	=
a) Both choice positions (n=96):	30	34	32
b) Choice at outset:	13	20	15
c) Choice in context:	17	14	17

It might be expected that reverse order would only affect the choices in the narrative text and not in the descriptive one as only the former has a time sequence. For the narrative text, the reverse order presentation is analogous to a flashback structure. Table 12.8 Shows the results for Group B who had the texts in reverse order. For those without a time sequence the rated interestingness is clearly related to choice, but when there is a time sequence (or a narrative structure) and this is in flashback order the effects of "Interestingness" are reversed (those rated most interesting were selected 9 times as opposed to those which were rated less interesting 20 times).

---

Table 12.8.

Rated interestingness and choice of sentences (with and without a time sequence ) presented in flashback order (Condition B).

	Choice Category		
	+	-	=
Texts 2 and 3 (Time Sequence):	9	20	19
Texts 1 and 4 (No Sequence ):	21	14	13

---

The texts which had a time order (Texts 2 and 3) were compared under two conditions (presented in canonical order (Group A) and presented in flashback order (Group B). The results (see Table 12.9) show that the most "Interesting" sentences are selected when the text is presented in canonical order, but that in flashback order they are clearly not the basis of selection. These results support those shown in Table 12:8 and indicate that the processing of texts presented in flashback order is influenced by factors other than "Interestingness", whereas "Interestingness" is related to the the processing of texts when they are presented in canonical order (the main effect of "Interestingness" being on choice made at the outset).

Table 12.9

Rated interestingness and choice of sentences with a time sequence (Texts 2 and 3 ) in flashback and canonical order.

	Choice Category		
	+	-	=
Condition A (canonical):	26	13	9
Condition B (flashback):	9	20	19

---

### Discussion.

Schank (1978) proposed that there were some concepts which were intrinsically absolutely interesting, and that it was on the basis of this absolute interestingness that readers selectively processed texts: the selection of the interesting concepts constraining the number of inferences which were made about the text. This series of experiments (Experiments 3, 4 and 5) have investigated the validity of the notion of absolute interestingness and the selective use of "Interestingness" in the processing of texts.

The initial experiment established that there were concepts which, when presented in isolation, were rated as being very interesting and others as less interesting. Those concepts which were rated as being most interesting were those which had been selected on the basis of Schank's criteria and those which were least interesting were those which had been included as more prosaic and not conforming to the criteria of absolute interestingness. The results showed, however, that this interestingness was not an all or none dichotomy, but that there were levels of interestingness and that both very interesting and less interesting concepts were ranged along a continuum of interestingness. Experiment Four showed that some "Not Interesting" ideas became more interesting after they had been read in context.

Experiment Five investigated the function of "Interestingness" in the processing of text and it was found that although there were indications that "Interestingness" was related to the initial processing of a piece of text, the relationship became less clear when processing was carried out in the presence of context or when there was a variation in the canonical order of a narrative text.

The results obtained when the narrative text was presented in "flashback" show that even at the beginning of a text there are factors other than "Interestingness" which are affecting the choice of sentence. The balance between choice categories is the same whether the sentences were presented at the outset or in context (that is the less interesting sentences are chosen more often than the more interesting ones). This effect does not occur when the sentences are presented in canonical order. This result is both surprising and counter-intuitive. It may be due to a transfer effect: that subjects in the reverse order condition may be expecting a flashback sequence and therefore the factors which influence choice in a flashback conditions are operating. There is not enough data to examine this proposal in detail and therefore it remains speculative.

Schank's theory of "Interestingness" assumes a process in which the input of individual concepts from a text predominates. Such a process disregards the importance of the order in which the text (or the individual concepts ) is presented and the role of the reader is reduced to that of a supplier of a disconnected string of associations. The process itself is text driven with the content stimulating associations in the reader's mind. In this model the reader's world knowledge about what is possible or probable in a situation is largely ignored. Such a theory is strangely at odds with script theory which assumes a contribution on the part of the reader about what is normal in a situation and a selective processing which places on a "weird list" those events which do not belong to the norm.

The concept of "Interestingness" would seem to have some reality, but that it is absolute is questioned. There do appear to be levels of "Interestingness" and some ideas are clearly of a higher level of "Interestingness" than others. However, "Interestingness" would seem to be only one of the factors important in the selective processing of text. Variations in the canonical order of texts and the context in which the concepts are embedded would seem to be important factors in the process. Giving a flashback order to the text would seem to be particularly influential in affecting the selection of content for further processing.

This series of experiments have shown that when ideas are presented out of context ( as for example news headlines or book titles might be) some ideas are more interesting than others and the more interesting tend to be those that are sensational. However, as Experiment Four has shown, placing ideas in context changes their level of interestingness and when the structure of the text is changed (as in Experiment Five) the interestingness is modified in a different way.

Experiment Six will investigate further the role played by the content of the text in comprehension by looking at the reader's ratings of predictions and investigating the relative contributions of the text and the reader's world knowledge to such ratings.

CHAPTER THIRTEENThe effect of text and reader related variables on predictions.Introduction

Experiments One and Two demonstrated that when, as a measure of comprehension, readers are asked to complete a narrative text by making predictions about the next event, they will contribute some stored knowledge about what is likely to the information in the text.

The responses to open-ended questions about 'What happened next' established a range of predictions. These predictions were made in response to 'stories' which established three different situations: she was being pursued by a soldier; an intolerant Jenny Fields was being pursued by a soldier (in wartime when people were being tolerant of soldiers); the intolerant Jenny Fields was consequently arrested for wounding a man after she had been pursued by a soldier. In all these situations it was stated that she was determined not to move again.

The content analysis method allowed these responses to be placed into categories such as 'behaviour related to him', 'behaviour related to her' and into sub-categories such as 'aggressive', 'casual', 'sexual' and 'normal cinema' behaviour. This analysis allowed relatively unconstrained responses to be compared in a quantitative way and established that, when readers are given different amounts of information, different predictions are made. While it was established that the predictions go beyond the information explicit in the text, and that the type of predictions made vary as a function of the amount of text presented, the relatively open-ended nature of the method made it difficult to disentangle various possible reasons for those effects.

The present experiment aims to separate out the relative contributions made by (a) the reader's prior knowledge and (b) the information in the text to predictions made about the next event. From the predictions elicited in the earlier experiments a number of events will be selected. They will be chosen from predictions made in response to all of the text conditions. Readers will be asked to rate the likelihood of each event as the 'next event'. As the predictions are concerned with the behaviour of the characters each possible event will be presented with both the pursued and the pursuing character as subject: (e.g. She slapped his face; He slapped her face). This will provide two rating scales for each event.

A minimal text will be given to two groups of subjects (one of males, one of females) to establish baseline ratings of likely and unlikely behaviour when readers have a neutral description of two characters (one pursued by the other) in a cinema setting. The characters will be described as A and B and no reference will be made to their sexual identity or occupation. The ratings will establish what behaviours are thought to be likely in this setting and what each character is likely to do when the pursuer eventually sits next to the pursued. Ratings which are both likely and non-idiosyncratic will establish default values for the setting. In addition subjects will be asked about the sexual identity of each of the characters.

Two other texts will each be given to different groups of male and female subjects. Both these texts will contain additional background information about the situation (a wartime setting and a general public consideration for people who have the same occupation as the pursuer) and about the pursued person (intolerance of the opposite sex and particularly those of the same occupational group as the pursuer; determination not to move seats again). The minimal text can be compared with the other (the Experimental) texts to see which the predictions about behaviour remain the same and which have changed as a result of the additional information. It is to be expected that



the prediction "the pursued moves seats" will be rated less likely in the Experimental conditions because the text contains the explicit information that he/she resolved not to move again as well as the additional information about the determination and intolerance of the character. The combined effects of the description of the background and the personality of the character (the pursued person's determination and his/her persistent intolerance in face of general public consideration for people of the same occupation as the pursuer) may also affect other ratings. If so, these predictions will also be rated in a different way by both the Experimental Groups although it cannot be forecast whether subjects will rate an independent minded pursued person as, for example, more or less likely to call the manager or to sit still and ignore the pursuer or to make a violent attack on him/her.

The two experimental texts will be made different by switching the sexual identity of the characters. In one text Jenny Fields will be pursued by a male soldier while in the other James Forest will be pursued by a female nurse. Thus, although the two texts will be similar in having the same amount of additional information, they will differ in the kind of information which they contain. This will allow the effects of type and quantity of additional information (confounded in the earlier experiments) to be assessed separately.

Ratings obtained in these conditions will be compared with each other. This comparison will enable the effects of the sex and the role of the character on predictions to be assessed. If there is no difference in the ratings made by subjects in each condition, then it can be assumed that readers have made predictions on the basis of their knowledge about the roles of pursued and pursuer. If there are differences between the ratings then the sex of the character as well as the role will have influenced these predictions. It is expected that predictions about face slapping and chatting up are those which will show the greatest differences if the ratings are based on world knowledge about the stereotypes of sex related behaviour and that

face slapping will be rated as more likely behaviour when the character is female and that chatting up will be rated as more likely when the character is male. If there is conflict between the behaviour expected from the role and the sex of the character this should be shown in the interactions between the ratings when the sex of the characters is reversed for the two Experimental Conditions. How closely the sex of the characters relates to the roles of pursued and pursuer will be assessed by asking subjects in the Baseline condition (where the sex of the characters is not given) to rate the likelihood of each of the characters being male or female.

When predictions were made freely in Experiments One and Two, violent events were not predicted by subjects who did not have the Outcome information that "Jenny Fields was arrested in a movie theater for wounding a man". So it is anticipated that this outcome will be unexpected and hence be given a low likelihood rating. Extra information about the outcome (that is, pursued stabs pursuer with a scalpel) will be given in the Experimental Conditions after the enriched texts have been read and the events rated. Having been given this extra information subjects will then be asked to rate the responsibility of each character for this violent outcome. Their ratings on this scale will measure the relative effects of the sex and role of the character when subjects have to extend the framework of comprehension to encompass new and unanticipated information.

Groups of male and female subjects will be used in each condition to assess the effects of individual differences in perspective upon the ratings. It seems reasonable to expect that subjects will identify more with characters of their own sex because of their greater similarities in contributed 'world view' and that this identification will be reflected in the rating scales. It might be expected, for example, that male subjects will think it more likely than female subjects that a woman will slap a man's face whatever her role in the event. There may also be a general identification by subjects with the

character of their own sex which will lead to a difference between the ratings of male and female subjects on all scales. This should be most apparent when subjects in the Experimental Conditions read about the stabbing. It is expected that they will attribute more responsibility for what happened to the character of the opposite sex, regardless of whether he or she is the pursued or the pursuer.

From the results of the experiment it will be possible to discover the degree of agreement between people in a group; the extent to which predictions are based on a common base of shared knowledge; the extent to which such predictions are modified by additional information; and whether the predictions are based on "objective" information in the text or "subjective" prior assumptions (that is aspects of the reader's own sex role, individual experience etc).

### Method.

#### Subjects.

48 subjects participated in the experiment: 24 were male, 24 female. All were students at the University of Hull and were chosen at random from people who were entering or leaving the University Library.

#### Materials.

##### a. Texts.

Three texts were prepared (see Figure 13:1), one for each of the conditions to be described. All were based upon the 'Garp' text used in Experiments One and Two but each had been modified in the ways suggested in the Introduction, so that:

Text 1 - (Baseline (Bas) condition) contained no indication of the characters' sex, occupation or personality. Nor did it contain any details about the background to the situation.

Text 2 - ('female-pursued' condition (Jenny)) named the pursued character as Jenny Fields and described her as intolerant towards men and soldiers at a time when other people were being considerate towards soldiers. The pursuer was described as a soldier and later as he, so that his sex was unambiguous. There were also details of the wartime setting and the cinema.

Text 3 - ('male-pursued' (James) condition) used text which was identical to Text 2 except that the pursued character was named James Forest and he was intolerant of women and nurses. The pursuer was described as a nurse and later as she.

Fig.13:1. The texts

**Text 1**

In a cinema, person A has moved places several times. Each time person B has moved closer. A is now sitting against a wall and B has just moved into the next seat.

**Text 2**

It was wartime and people were being considerate towards soldiers but Jenny Fields was quite firm in her intolerance of the behavior of men in general and soldiers in particular. In the movie theater she had to move three times and each time the soldier moved closer to her until she was sitting against the musty wall, her view of the newsreel almost blocked by some silly colonnade, and she resolved she would not get up and move again. The soldier moved once more and sat beside her.

**Text 3**

It was wartime and people were being considerate towards nurses but James Forest was quite firm in his intolerance of the behavior of women in general and nurses in particular. In the movie theater he had to move three times and each time the nurse moved closer to him until he was sitting against the musty wall, his view of the newsreel almost blocked by some silly colonnade, and he resolved he would not get up and move again. The nurse moved once more and sat beside him.

b. Events to be rated as likely-unlikely.

Nine 'events' were chosen from the raw data from Experiment Two. These were chosen from the responses which had been given to the question:

'Describe in detail what you think happened after

"The soldier moved once more and sat beside her" '.

The responses were selected so that there were some from each of the four conditions. Each of the nine events was presented twice - i.e. with both the pursuer and the pursued as the subject of what happened. In the Baseline condition event statements took the form of:

A called the Manager.

B called the Manager.

In the Experimental conditions 'he' and 'she' were used:

He called the Manager.

She called the Manager.

A complete list of the 18 events is presented as a part of Table 13:1.

c. The Rating scale.

Subjects were to be asked to rate the likelihood of each event. The ratings were on an eleven-point scale.

0 = there is no possibility of it happening

1 = there is hardly any chance of it happening

2 = it is extremely unlikely

3 = it is very unlikely

4 = it is unlikely

5 = it is equally likely/unlikely

6 = it is likely

7 = it is very likely

8 = it is extremely likely

9 = it is almost certain to happen

10 = it is absolutely certain to happen.

**Extra information.**

After filling in the event rating scales:

i. Subjects in the Experimental Conditions (Jenny/James) were given extra information that either

(Jenny) Jenny Fields stabbed the soldier with a scalpel.  
or  
(James) James Forest stabbed the nurse with a scalpel

They were then asked to:

"Answer the following two questions using a rating of 0 to 10, where 0 = not at all, 10 = Entirely:

How responsible was he for what happened? ( )

How responsible was she for what happened? ( )"

ii. Subjects in the Baseline condition (who were given no extra information) were asked to rate the likelihood that

A was male

A was female

B was male

B was female.

**Design.**

8 male and 8 female subjects were allocated at random to each of the three conditions: A pursued by B (Bas), Jenny pursued by soldier (Jenny) and James pursued by nurse (James).

In all conditions subjects were asked to read the appropriate version of the text and then to rate the likelihood of each of the 18 events occurring as the 'next event', using the eleven-point rating scale, described above.

The data collected in these conditions will be analysed in the following ways:

**A. Ratings.**

1. Mean ratings for each group in each condition will be compared to examine the level of agreement between the groups (using intergroup correlations).
2. The data from the ratings in the Baseline condition will be used to establish the general level of likelihood associated with each event.
3. The data from the ratings in each of the Experimental Groups will be compared with the Baseline ratings. As the ratings are not necessarily independent (i.e. she could call the Manager because he had chatted her up, etc.) stepwise regression analyses will be used. The results will be compared to see which changes are common (and can therefore be attributed to the addition of background material) and which are specific to each of the experimental conditions and therefore to the difference in the material added (i.e. sex of character).
4. All ratings will be analysed for overall variance which can be attributed to the sex of the subjects.
5. The data from the ratings in each of the Experimental Groups will be compared to establish which of the events in the two conditions show differences related to the sex of the characters (he/she) and the sex of the subjects.

**B. Extra Information**

1. The additional ratings obtained from the Baseline group will be analysed using analysis of variance. This will establish whether pursuer/pursued roles are seen as being linked to sex and whether, in general, male and female subjects perceive the sex role differently.



2. The ratings of responsibility for the violent outcome given to the Experimental Groups will be analysed for the relative variance attributable to:

- sex of subject (male/female)
- sex of character (male/female)
- role of character (pursuer/pursued)

and to determine whether there are any interactions.

#### Procedure.

The stimulus materials were prepared and divided so that half the papers for each condition were labelled 'male' and half were labelled 'female'.

Students, aged between 18 and 25, who had just entered or were preparing to leave the University Library, were approached and asked to complete the questionnaire. Only three refused. Those who agreed were taken to tables near the Issue Desk. Subjects were spaced out in both time and location to prevent collaboration. Then they were given a pen or pencil, a copy of the questionnaire and the following instruction:

"Please read the text and written instructions very carefully before you begin. Fill in the questionnaire when you have read the text and the instructions. When you have done that please return the paper and pencil to me."

The materials were prepared with the text, the instructions, the description of the intervals on the rating scale and the list of potential next events on one side of A4 paper. The additional material was presented on a second sheet of A4 paper. The two sheets were stapled together in such a way that the second sheet could not be read without unfastening the staples.

When this was done the subject was then asked:

"Would you please unfasten the staples and then answer the remaining two questions."

Some subjects asked the purpose of the questionnaire and were told that it was an investigation of differences in responses between male and female subjects.

### Results

#### 1. General Results.

The overall means for all three conditions when related to the 18 scales (Table 13:1) show that it is likely (mean ratings > 6) that:

'the pursued person will ignore the pursuer' (Scale 3)

'the pursuer will chat up the pursued person' (Scale 10).

Neutral (between 4 and 6) ratings are given to the statements that the pursued person will:

'move seats' (Scale 1),

'walk out of the cinema' (Scale 7),

'call the Manager' (Scale 13),

'sit still and watch the film' (Scale 15),

'slap the pursuer's face' (Scale 17),

It is rated as unlikely (mean ratings < 4) that the pursuer will:

'move seats' (Scale 2),

'ignore the pursued person' (Scale 4),

'walk out of the cinema' (Scale 8),

'sit still and watch the film' (Scale 16),

and very unlikely (mean ratings < 2) that:

either character will shoot or stab the other (Scales 5,6,11,12),

the pursued person will 'chat up the pursuer' (Scale 9),

the pursuer will 'call the Manager' (Scale 14), or

'slap the face of the person being pursued' (Scale 18).

Table 13:1

Mean ratings for male and female subjects under each condition for each of the 18 scales.

Scales		1	2	3	4	5	6	7	8	9
Baseline	M	6.88	4.88	5.25	2.75	1.00	2.00	5.75	2.88	1.75
	F	7.13	3.75	7.00	5.00	1.13	1.75	5.63	3.38	1.88
	Both	7.01	4.32	6.13	3.88	1.07	1.88	5.69	3.13	1.82
Jenny	M	4.63	1.88	7.75	2.63	1.38	1.13	5.00	2.25	1.00
	F	4.63	3.38	6.75	2.75	0.38	0.88	4.50	1.88	0.63
	Both	4.63	2.63	7.25	2.69	0.88	1.01	4.75	2.07	0.82
James	M	4.38	4.75	4.75	4.75	0.75	1.00	5.13	3.75	3.63
	F	5.63	3.63	6.88	5.00	1.38	1.13	6.38	3.25	2.50
	Both	5.01	4.19	5.82	4.88	1.07	1.07	5.76	3.50	3.07
Overall	Mean	5.55	3.71	6.40	3.81	1.00	1.32	5.40	2.90	1.90

  

Scale		10	11	12	13	14	15	16	17	18
Baseline	M	6.75	0.75	1.25	6.38	0.88	4.63	3.25	5.63	1.75
	F	6.75	1.38	1.75	6.25	1.50	5.50	4.25	4.13	2.38
	Both	6.75	1.07	1.50	6.32	1.19	5.07	3.75	4.88	2.07
Jenny	M	7.50	1.00	1.00	5.75	1.25	6.50	3.88	5.13	2.13
	F	7.50	1.25	1.25	4.00	0.88	5.25	3.38	5.88	1.00
	Both	7.50	1.13	1.13	4.88	1.07	5.88	3.63	5.51	1.57
James	M	6.63	1.00	1.50	3.63	2.38	5.00	4.75	1.63	1.13
	F	6.75	1.13	1.00	4.63	3.13	5.75	4.38	2.75	2.50
	Both	6.69	1.07	1.25	4.13	2.76	5.38	4.57	2.19	1.82
Overall	Mean	6.98	1.09	1.29	5.11	1.67	5.44	3.98	4.19	1.82

## Scales 1-18:

1. The pursued person moves seats.
2. The pursuer moves seats.
3. The pursued person ignores the pursuer.
4. The pursuer ignores the pursued person.
5. The pursued person stabs the pursuer.
6. The pursuer stabs pursued.
7. The pursued person walks out of the cinema.
8. The pursuer walks out of the cinema.
9. The pursued person chats up the pursuer.
10. The pursuer chats up the pursued.
11. The pursued person shoots the pursuer.
12. The pursuer shoots the pursued.
13. The pursued person calls the Manager.
14. The pursuer calls the Manager.
15. The pursued person sits still and watches the film.
16. The pursuer sits still and watches the film.
17. The pursued person slaps the pursuer's face.
18. The pursuer slaps the pursued person's face.

## 2. Differences between male and female subjects.

From Tables 13:1 it can be seen that there is very little difference in ratings as a function of sex of subjects although some scales do show a difference in means. For example:

In the Baseline condition , female subjects rate the pursuer more likely to ignore the pursued person (Scale 4);

In the Jenny condition, the pursuer (the soldier) is rated more likely to move seats (Scale 2) by females than by males

and in the James condition, face slapping behaviour by both James and the nurse (Scales 18 and 19) is rated more likely by females than males.

Table 13:2 shows very high correlations between the means for the male and female groups in all three conditions ( $r = 0.92$  for the Baseline,  $0.95$  for Jenny and  $0.89$  for James). Point biserial correlations between sex of subject and ratings were compiled for each of the 18 scales across all 3 conditions. These again showed no significant differences between the sexes on any scale ( $N = 48$ ,  $p > 0.05$  in all cases). This null result was further supported by a discriminant function analysis of the male/female differences. The scale which best discriminated between the two sex groups (Scale 4 'the pursuer will ignore the pursued person') only accounted for 3% of the total variance ( $F = 1.65$ ,  $df 1:46$ ,  $ns$ ). Because of the lack of differences on all these scales no further analysis of the apparent sex differences was carried out.

It can be concluded that sex differences which readers bring to the text do not have any consistent effect on their ratings of various events given the textual context. This is a somewhat surprising finding, especially given the apparent relevance of sex-role related behaviour in the texts themselves. Given the lack of any sex differences between subjects, the following analyses will concentrate on the differences between the three conditions treating the subjects as three groups each of 16 subjects (that is, 8 males, 8 females).

Table 13:2

Matrix of correlations between group means (n = 18 scales).

		Baseline		Jenny		James	
		M	F	M	F	M	F
Baseline	M	1.00	0.92	0.84	0.89	0.69	0.79
	F		1.00	0.90	0.89	0.79	0.92
Jenny	M			1.00	0.95	0.66	0.86
	F				1.00	0.70	0.82
James	M					1.00	0.89
	F						1.00

Table 13:3. Comparison between pairs of conditions.

Columns 2-4 show point biserial correlations, t values and significance levels for the differences between conditions for each scale. Columns 5-7 show Beta Weights, F ratios and levels of significance for the stepwise multiple regression analyses. (At the base of each set of columns 5-7, the overall multiple correlation and its level of significance is given.)

Scale *	rpbis	t (df =30)	p	Beta Weight	F (df 1, 27)	
<b>a. Baseline/Jenny.</b>						
1	0.54	2.42	p<0.05	0.49	14.21	p<0.001
10	0.28	1.09	ns	0.43	11.11	p<0.01
9	0.40	1.62	ns	0.27	4.63	p<0.05
13	0.37	1.47	ns	0.27	3.74	p<0.01
				[ R = 0.77 F = 9.97, df 4, 27, p<0.001]		
<b>b. Baseline/James.</b>						
17	0.55	2.44	p<0.05	0.37	9.32	p<0.01
14	0.46	1.93	ns	0.49	13.45	p<0.01
6	0.28	1.09	ns	0.38	8.59	p<0.01
1	0.4	1.97	ns	0.25	4.11	p<0.1
13	0.51	2.24	p<0.05	enters at step 9		
				[ R = 0.80 F =12.26, df 4, 27, p<0.001]		
<b>c. Jenny/James.</b>						
(df 1, 25)						
9	0.69	3.62	p<0.01	0.31	4.014	p<0.1
17	0.63	3.05	p<0.01	0.63	12.265	p<0.01
2	0.29	1.11	ns	0.39	11.392	p<0.01
14	0.47	1.97	ns	0.29	5.595	p<0.05
8	0.39	1.58	ns	0.50	7.642	p<0.05
7	0.26	<1	ns	0.24	3.847	p<0.01
				[ R = 0.86 F =11.35, df 6, 25, p<0.01]		

\* See Table 13.1

### 3. Baseline Conditions.

The baseline condition provides information on the subjects' expectations when the text contains only an outline of the event sequence. The effects of adding further information and switching the characters' sexual identities can be assessed by seeing how these expectations are altered in such conditions.

In the Baseline Condition, subjects generally rate it as likely that:  
the pursuer (character B) will

'chat up the pursued' (Scale 10);

and that the pursued (character A) will

'move seats' (Scale 1),

'ignore the pursuer' (Scale 3),

'call the Manager' (Scale 13).

They give Neutral ratings to the statements that:  
the pursuer (character B) will

'move seats' (Scale 2),

'sit still and watch the film' (Scale 16);

and that the pursued (character A) will

'walk out of the cinema' (Scale 7),

'sit still and watch the film' (Scale 15),

'slap the pursuer's face' (Scale 17).

They rate it as unlikely that:

either character will shoot or stab the other (Scales 5, 6,11,12),  
that the pursuer (character B) will:

'walk out of the cinema' (Scale 8),

'ignore the pursued person' (Scale 4),

'call the Manager' (Scale 14),

'slap the pursued person's face' (Scale 18);

and that the pursued person (character A) will:

'chat up the pursuer' (Scale 9).

#### 4. Differences between the Conditions.

For each pair of conditions, two complementary analyses were carried out (see Table 13.3). Point biserial correlations between conditions and ratings were computed for each of the 18 scales, and a discriminant stepwise multiple regression analysis was carried out (with 'conditions' as the dependent variable) to assess the overall difference between conditions. The latter analysis was carried out in order to be able to take account of the relationships between scales when it was considered how "different" the behaviour of subjects was under two different conditions: examination of biserial correlations could be misleading in cases where scales are highly correlated with each other or when a scale accounts for a significant proportion of the covariance between conditions and scales but not of the variance between conditions and that scale alone.

##### 4a. Baseline versus Jenny.

Scale 1 ('the pursued person moves seats') was rated significantly more likely in the Baseline condition ( $r_{pbis} = 0.54$   $t = 2.42$ ,  $df\ 30$ ,  $p < 0.05$ ). No other single scale showed a significant difference ( $t < 2.042$  in all cases).

The regression analysis revealed that by itself Scale 1 accounted for 29.52% of the total variance ( $F = 12.56$   $df\ 1,30$ ,  $p < 0.01$ ). The group of scales which best discriminated between the two conditions were 1, 10, 9 and 13 (see Table 13:3a).

In the Jenny condition the pursued person (Jenny) is less likely than A to:

- 'move seats' (Scale 1)
- 'chat up the pursuer' (Scale 9),
- 'call the Manager' (Scale 13);

the pursuer (the soldier) is more likely than B to 'chat up the pursued' (Scale 10).



Subjects in both conditions agree it is likely that:

'the pursued will ignore the pursuer' (Scale 3), and rate all the other events as equally unlikely. The comparison of the two conditions shows that there is a high level of agreement with the exception of those scales which are listed. That Jenny is rated much less likely than A to move seats is not a surprising result given that the information "she resolved she would not get up and move again" was a part of the Jenny but not the Baseline text. That she is less likely to chat up the soldier and to call the manager is also likely to be a function of the additional information but this is a more speculative conclusion. This is discussed further in section 4b.

#### 4b. Baseline versus James.

Scale 13 ( $t\ 2.24$ ,  $df\ 30$ ,  $p<0.05$ ) 'the pursued person will call the Manager' is rated significantly more likely in the Baseline Condition (A is more likely to call the manager than James). Scale 17 ( $t\ 2.44$ ,  $df\ 30$ ,  $p<0.05$ ) 'the pursued person will slap the pursuer's face' is rated as significant more likely in the Baseline Condition (see Table 13:3b) (that is, B is more likely to slap A's face than is James to slap the nurse's face). No other single scale showed a significant difference ( $t<2.042$  in all cases).

The regression analysis revealed that Scales 17, 14, 6 and 1 best discriminated between the two conditions. The variance for Scale 13 was accounted for by the other scales. In the James condition, the pursued person (James) is less likely to move seats (Scale 1) and to slap the pursuer's face (Scale 17). The pursuer (the nurse) is more likely to call the Manager (Scale 14) and less likely to stab the pursued person (James) (Scale 6).

In the comparison between the two conditions there is again a high level of agreement between the two conditions. That James is less likely than A to move seats can be accounted for by the additional given information. That James is rated less likely to slap the nurse's face is likely to be due to differences

in the assumed sex of character A and James. This will be discussed further in Section 4c.

#### 4c. Jenny versus James.

In the Jenny condition (when Jenny is pursued and the soldier is the pursuer) it is rated significantly less likely, ( $t = 3.62$ ,  $df\ 30$ ,  $p < 0.01$ ) that 'Jenny will chat up the soldier' than 'James will chat up the nurse' (Scale 9). 'Jenny will slap the soldier's face' is rated as significantly more likely ( $t = 3.05$ ,  $df\ 30$ ,  $p < 0.01$ ) than 'James will slap the nurse's face' (Scale 17). (see Table 13:3c). No other single scales showed a significant difference.

The regression analysis revealed that Scales 9, 17, 2, 14, 8 and 7 best discriminated between the two conditions.

In comparing the nurse and soldier as pursuers, the nurse is more likely to:

- 'move seats' (Scale 2),
- 'walk out of the cinema' (Scale 8),
- 'call the Manager' (Scale 14).

In comparing Jenny and James as the pursued, James is more likely to:

- 'chat up the pursuer' (the nurse) (Scale 9),
- 'walk out of the cinema' (Scale 7)

and Jenny is more likely to:

- 'slap the pursuer's (the soldier's) face' (Scale 17).

From these results it seems likely that slapping faces, moving seats and calling the Manager are seen as 'female' acts and are not related to the role she is playing, and that 'chatting up' may be a male activity whether he is pursued or pursuing.

An analysis of variance was carried out on ratings for each pair of events (for example, pursuer ignores pursued, pursued ignores pursuer and so on) in the Jenny and James conditions. This was to determine whether ratings

Table 13:4.

Results of Analyses of Variance between sex and role of characters  
for pairs of scales showing significant effects

	Means			F ratios (df 1, 28)		
	Pursuer	Pursued	Mean	Sex	Role	Interaction
<b>Scales 1 &amp; 2: Moves seats.</b>						
Male	2.63	5.01	3.82	<1	4.82 *	<1
Female	4.19	4.63	4.41			
Mean	3.41	4.82				
<b>Scales 3 &amp; 4: Ignores.</b>						
Male	2.69	5.81	4.25	<1	19.94 ***	8.66 **
Female	4.88	7.25	6.07			
Mean	3.78	6.53				
<b>Scales 7 &amp; 8: Walks out.</b>						
Male	2.06	5.75	3.91	4.68 *	50.50 ***	<1
Female	3.50	4.75	4.13			
Mean	2.78	5.25				
<b>Scales 9 &amp; 10: Chats up.</b>						
Male	7.50	3.06	5.28	5.04 *	263.734 ***	23.75 ***
Female	6.69	0.81	3.75			
Mean	7.09	1.94				

\* =  $p < 0.05$   
 \*\* =  $p < 0.01$   
 \*\*\* =  $p < 0.001$

Table 13:4. (continued)

Results of Analyses of Variance between sex and role of characters  
for pairs of scales showing significant effects

Scales 13 & 14: Calls manager.	Means			F ratios (df 1, 28)		
	Pursuer	Pursued	Mean	Sex	Role	Interaction
Male	1.06	4.16	2.61	<1	43.56 ***	9.61 **
Female	2.75	4.88	3.82			
Mean	1.91	4.50				

Scales 15 & 16: Sits still.	Means			F ratios (df 1, 28)		
	Pursuer	Pursued	Mean	Sex	Role	Interaction
Male	3.63	5.83	4.51	<1	10.82 **	2.83 ns
Female	4.56	5.88	5.22			
Mean	4.09	5.63				

Scales 17 & 18: Slaps face.	Means			F ratios (df 1, 28)		
	Pursuer	Pursued	Mean	Sex	Role	Interaction
Male	1.56	2.19	1.88	9.38 **	35.19 **	24.01 **
Female	1.81	5.51	3.66			
Mean	1.69	3.84				

\* =  $p < 0.05$   
 \*\* =  $p < 0.01$   
 \*\*\* =  $p < 0.001$

of these events were attributable to the sex or the role of the character or to some interaction between the two.

The results of these analyses (see Table 13:4) show that for each pair of events analysed there was a main effect. There were differences between pursued and pursuer for all events except for stabbing (Scales 5 and 6) and shooting (Scales 11 and 12). The pursued is rated more likely to carry out all the remaining actions except for "chatting up" (which is rated as more likely to be done by the pursuer). There were some differences which could be attributed to the sex of the character. Females are more likely to slap faces whether they are pursuing or pursued ( $F = 9.378$ ,  $df 1, 28$ ,  $p < 0.01$ ). Male characters are more likely to chat up the opposite sex whether they are pursuing or pursued ( $F = 5.036$ ,  $df 1, 28$ ,  $p < 0.05$ ).

Other pairs of events showed significant interaction effects. The pursued person is more likely to ignore the pursued but ignoring the other is more likely in either role when the pursuer is female ( $F = 8.66$ ,  $df 1, 28$ ,  $p < 0.01$ ). The pursued person is more likely to call the Manager than the pursuer but calling the Manager is more likely in either role when the character is female ( $F = 9.617$ ,  $df 1, 28$ ,  $p < 0.01$ ). The pursuer is more likely to chat up the pursued person but the chatting up is more likely when the pursuer is male ( $F = 23.25$ ,  $df 1, 28$ ,  $p < 0.001$ ). A female is more likely to slap a male person's face than vice versa but face slapping is more likely when the female is being pursued ( $F = 24.013$ ,  $df 1, 28$ ,  $p < 0.001$ ).

#### 5. Summary of results of similarities and differences in ratings of scales between the conditions.

The results of the ratings of the scales can be summarised in three sections (see Table 13:5).

Table 13:5. Summary of Results.

a. No significant differences ( $t < 2.042$  in all cases) for scales:

- 3 Pursued ignores pursuer
- 15 Pursued sits still and watches film
- 5 Pursued stabs pursuer
- 11 Pursued shoots pursuer
- 12 Pursuer shoots pursued
- 16 Pursuer sits still and watches film
- 18 Pursuer slaps pursued's face
- 4 Pursuer ignores pursued

b. Difference between both Baseline/Jenny and Baseline/James with no differences between Jenny/James.

Scale	Differences		
	A Bas/Jenny	B Bas/James	C Jenny/James
1 Pursued moves seats	rpbis=-0.54 t = -2.42*	+	
13 Pursued calls Manager	+	rpbis=-0.51 t = -2.24	

ci. Difference between either Baseline/Jenny or Baseline/James with no differences between Jenny/James.

Scale	Differences		
	A Bas/Jenny	B Bas/James	C Jenny/James
10 Pursuer chats up pursued	+		
6 Pursuer stabs pursued		+	

cii. Difference between Experimental Conditions (Jenny/James)

Scale	Differences		
	A Bas/Jenny	B Bas/James	C Jenny/James
9 Pursued chats up pursuer	+		rpbis= 0.69 t = 3.62**
17 Pursued slaps pursuer's face		rpbis= 0.55 t = 2.44*	rpbis= 0.63 t = 3.05**
14 Pursuer calls the Manager		+	+
2 Pursuer moves seats			+
8 Pursuer walks out of cinema			+
7 Pursued walks out of cinema			+

+ discriminates between groups in Regression Analysis.

\* significant at  $p < 0.05$ .

\*\* significant at  $p < 0.01$ .

5a) Those scales on which ratings do not differ significantly and which do not discriminate between groups.

These ratings can be related to the roles of pursued and pursuer. These are not changed when additional information about the sexual identity and the personality of the character is added along with background information about the situation. The likelihood of these events is linked strongly to the scenario and they are not easily modifiable by adding new information. It is likely (whether the pursued person is Jenny, James or A and whether the pursuer is the soldier, the nurse or B) that:

'the pursued person will ignore the pursuer' (Scale 3).

Neutral ratings are given to the event that:

'the pursued person will sit still and watch the film' (Scale 15),

and it is unlikely that:

'either character will shoot the other' (Scale 11, 12),

'the pursued will stab the pursuer' (Scale 5),

'the pursuer will sit still and watch the film' (Scale 16) or

'slap the pursued person's face' (Scale 18) or

'ignore the pursued person' (Scale 4).

5b) Scales which differ significantly in the same way when additional information is given regardless of the sexual identity allotted to the characters, that is, those scales which discriminate between both Baseline versus Jenny and Baseline versus James.

The differences in ratings on these scales can be attributed to the additional information other than that of the sexual identity of the character, that is, the information about the personality of the characters and background information about the situation. Such scales can be assumed to reflect events which have weak expectations associated with them, in the sense that the baseline conditions reflect default values which are readily modified

by added background information.

The event 'the pursued person will move seats' (Scale 1), is rated less likely to happen ( see Table 13:4) in both the Jenny and the James conditions than in the Baseline condition. There is a significant difference between Jenny/Bas and this is a discriminating variable between the James/Bas condition. It has been assumed that the information given in the text about the pursued character's decision to not move again will have directly contributed to this decrease in the likelihood of this happening (see Table 13:1).

The event 'the pursued person calls the Manager' (Scale 13), discriminates between the Jenny and the Baseline conditions and is rated significantly differently in the James versus Baseline conditions (see Table 13:3b). In both conditions it is less likely that either Jenny or James will call the Manager than that A, in the Baseline condition, will call the Manager.

5c) (i). Scales on which there are differences in ratings between the Baseline/Jenny conditions or the Baseline/James conditions (but not both) but which do not discriminate between the Jenny/James conditions.

The differences in ratings on these scales can be attributed to a unique combination of all the information in the text (that is, the particular sexual identities of the particular characters in each of the experimental conditions combined with the information about their personalities and the background to the situation).

Baseline/Jenny:

Ratings on 'the pursuer chats up the pursued' (Scale 10) discriminate between subjects in the Baseline and Jenny conditions (see Table 13:3b) although the difference is not significant at  $p < 0.05$  when the scale is tested in isolation. In the Jenny condition the soldier is more likely to chat up



Jenny than B is to chat up A in the Baseline condition.

Baseline/James:

Ratings on 'the pursuer will stab the pursued' (Scale 6) discriminate between the two groups of subjects although there is no significant difference between the groups when this is considered as a single variable. B (the pursuer) is more likely to stab A than the nurse is likely to stab James but both are very unlikely indeed.

5c)(ii) Those scales on which the ratings differ between the Experimental Groups (Jenny/James).

In the experimental conditions subjects were given the same amount of additional information but the sexual identities of the characters were switched. Differences between these two groups can be attributed to the stereotype male/female sexual identities in relation to the roles of pursued/pursuer, that is, they reflect events which have default values, conditional upon the character's role as pursuer or pursued, which are modified specifically by information about sexual identity.

Table 13:5c shows those single scales on which ratings are significantly different between groups (Scales 17 and 14) and those which additionally best discriminate between the groups (Scales 2, 7 and 8).

In the Jenny condition, when Jenny is pursued and the soldier is the pursuer, she is significantly more likely to slap the pursuer's face than James when he is being pursued by the nurse.

In the James condition the nurse is significantly more likely to 'ignore the pursued person' (Scale 4) and 'call the Manager' (Scale 14) than the soldier in the Jenny condition.

The discriminating scales show in addition that the pursuer in the James condition (the nurse) is more likely to 'move seats' (Scale 2) and 'walk out of the cinema' (Scale 8) than the pursuer (the soldier) in the Jenny

condition. Also the pursued person in the James condition (James) is more likely to 'walk out of the cinema' (Scale 7) and 'chat up the pursuer' (Scale 9) than the pursued person in the Jenny condition (Jenny).

Switching the sexual identities of the pursued (Jenny/James) and the pursuer (soldier/nurse) is related to changes in the ratings on certain scales which predict what a character is likely/unlikely to do after the pursuer has moved next to the pursued. The nurse is more likely than the soldier to move seats, call the Manager, ignore the pursued and walk out of the cinema. The pursued male (James) is more likely to chat up his female (nurse) pursuer than is the pursued female (Jenny) to chat up her male (soldier) pursuer. James is also more likely than Jenny to walk out of the cinema as a consequence of being pursued.

Table 13:6.

Assigned Sex of Character A by sex of subjects in Baseline  
Condition.

Table of Means and Analysis of variance.

GROUPS	MEANS		
	TREATMENTS		MEAN
	Assigned Female	Assigned Male	
Male Subjects	6.38	4.38	5.38
Female Subjects	7.13	3.50	5.31
MEAN	6.75	3.94	5.34

Subjects = 16 (8 per group)

ANALYSIS OF VARIANCE: SUMMARY TABLE

SOURCE	Sum Sqs.	d.f.	Variance	F ratio
BETWEEN Ss	21.7188	15	1.44792	
Groups	0.311279	1	0.0311279	<1
(Male/Female Subjects)				
Ss within groups	21.6877	14	1.54912	
WITHIN Ss	165.5	16	10.3438	
Treatments (Assigned Sex)	63.2816	1	63.2816	9.14 **
Groups x Treatments	5.28113	1	5.28113	<1
Ss within Groups x Treatments	96.9373	14	6.92409	
TOTAL	187.219	31		

\*\*  $p < 0.01$

Table 13:7

Assigned Sex of Character B by sex of subjects in Baseline Condition.  
Table of Means and Analysis of variance.

MEANS			
TREATMENTS			
	Assigned Female	Assigned Male	MEAN
<hr/>			
GROUPS			
Male Subjects	2.00	7.13	4.56
Female Subjects	3.13	6.63	4.88
<hr/>			
MEAN	2.56	6.88	4.72

Subjects = 16 (8 per group)

-----

ANALYSIS OF VARIANCE: SUMMARY TABLE

SOURCE	Sum Sqs.	d.f.	Variance	F ratio
<hr/>				
BETWEEN Ss	15.9692	15	1.06461	
Groups	0.781799	1	0.781799	<1
(Male/Female Subjects)				
Ss within Groups	15.1874	14	1.08481	
WITHIN Ss	188.5	16	11.7813	
Treatments (Assigned Sex)	148.782	1	148.782	60.48 ***
Groups x Treatments	5.28076	1	5.28076	2.15
Ss within Groups x Treatments	34.4376	14	2.45983	
<hr/>				
TOTAL	204.469	31		
<hr/>				

\*\*\*  $p < 0.001$

## 6. Assigned sex of characters in the Baseline Condition.

### 6a) Character A.

Ratings of subjects in the Baseline Condition showed that it is highly likely that A is female and unlikely that A is male (see Table 13:6). This difference is highly significant ( $F = 9.1393$ ,  $df 1, 14$ ,  $p < 0.01$ ). Both male and female subjects rated the likelihood of the sex of the character in a similar manner. There was no significant difference between the ratings of male and female subjects and no interaction effects.

### 6b) Character B.

Ratings of subjects in the Baseline condition showed that it is highly likely that B is male and very unlikely that B is female (see Table 13:7). This difference is highly significant ( $F = 60.4846$ ,  $df 1, 14$ ,  $p < 0.001$ ). Male and female subjects rated the likeliness of the sex of the character in a similar manner (see Table 13:7). There was no significant difference between the ratings of male and female subjects and no interaction effects.

Table 13:8.

Attribution of Responsibility by Experimental Groups  
(Jenny) (James)  
Table of Means and Analysis of variance

TABLE OF MEANS

Condition	Treatment	Sex of subjects		MEAN
		Male	Female	
Jenny	Pursuer	5.63	7.13	6.38
	Pursued	3.00	4.13	3.56
Mean Jenny		4.31	5.63	4.97
James	Pursuer	5.50	7.75	6.63
	Pursued	3.00	2.13	2.56
Mean James		4.25	4.94	4.59
Mean Pursuer		5.56	7.44	6.50
Mean Pursued		3.00	3.13	3.06
Overall Mean		4.28	5.28	4.78

Subjects = 32 (16 per condition, 8 male and 8 female).

ANALYSIS OF VARIANCE: SUMMARY TABLE

SOURCE	Sum Sqs.	d.f.	Variance	F ratio
BETWEEN Ss	398.937	31	12.8689	
Factor C (Jenny, James Conds)	2.25024	1	2.25024	<1
Factor D (Male/Female Subs.)	15.9988	1	15.9988	1.18
C x D	1.56287	1	1.56287	<1
Ss within Groups	379.125	28	13.5402	
WITHIN Ss	436	32	13.625	
Treatments (Pursuer/Pursued)	189.062	1	189.062	23.97 ***
C x Treatments	6.24951	1	6.24951	<1
D x Treatments	12.2518	1	12.2518	1.55
C x D x Treatments	7.56128	1	7.56128	<1
Ss within Groups x Treatments	220.875	28	7.88839	
TOTAL	834.937	63		

\*\*\*  $p < 0.001$

7. Rated responsibility of each character for what happened after reading about the stabbing by the pursued.

The results (Table 13:8) show quite clearly that the pursuer is rated highly responsible and the pursued is rated as having very little responsibility for what happened. The difference between the two is highly significant ( $F = 23.9671$ ,  $df 1, 28$ ,  $p < 0.001$ ). The lack of any difference between the ratings of male and female subjects, or between the conditions in which characters were given male and female sexual identities, is interesting and surprising. The main influence in deciding who is responsible are the role factors (pursuer/pursued) and not the additional information about male and female characters, nor the extra information about the stabbing or the sex of the subject. In this respect the additional information in the text does not over-ride the situational assumptions made at the outset.

Discussion.

This experiment used a technique of asking readers to rate the likelihood of alternative 'next events' in order to examine two main aspects of the interaction between reader and text: one was the extent to which such ratings were affected by the individual characteristics of the reader; the other was how they were affected by the type and amount of information in the text.

It was assumed that male and female readers would bring different viewpoints to the presented text and that these differences would in turn be reflected in the ratings given to events (especially those which are sex-role related). However, it was found that the sex of the reader had no differential effect on ratings on any of the scales (although there were some trends as was shown in Table 13.1). In addition to the high correlations between male and female subjects on most scales in all conditions, the variance between subjects within each condition was quite small. This suggests that, despite the differences in their individual 'world experience' all subjects comprehended much the same story from the texts that they were given. This implies that all the texts elicited predictions which could be made from a common base of general, shared knowledge. Idiosyncratic, individual interpretations of the events did not appear to affect the ratings in any significant way.

The second aspect of the investigation (that is, the extent to which ratings are affected by information in the text) showed some clear results from which firm conclusions can be made.

The ratings in the Baseline Condition showed that readers develop clear stereotypes of the characters being described even when they are labelled A and B. Part of this stereotype is the sex assigned to the character by virtue of his/her role in the situation. The pursuer is regarded as being male and the pursued person as female. When a male is pursuing a female in a cinema there are some things that he is regarded as being likely to do, some things



that she is regarded as likely to do and other things that neither will do. For example, she is likely to ignore him, and he is likely to chat her up; she is unlikely to chat him up and he is unlikely to slap her face. There are also well defined limits to the kinds of behaviour which are rated as unlikely. For example, both characters are seen as equally unlikely to shoot or stab the other.

It is clear from these ratings that when readers are given an outline of a situation in a text they will bring to the text a body of general and common knowledge which defines limits for the behaviour of characters in that situation and also defines the sex of characters who will fill certain roles. It is impossible, from these ratings in the Baseline Condition, to decide how much of the attributed behaviour is a function of the attributed sex and how much a function of the given role of the character. Neither is it clear how far these stereotypes will be modified by additional information about the background of the characters and the situation.

The question of the extent to which the stereotypes of the characters are modified by additional information was resolved by examining those things that changed in common in the ratings in both Experimental Conditions when they were compared with the ratings made in the Baseline Condition. These comparisons show that additional information which delimits the behaviour of the characters changes the ratings for the particular behaviour described. For example, the additional text stated that the pursued person had made up his/her mind not to move again. Subsequently the rating scales reflected this information by rating it less likely that she/he would move seats.

It is still not possible to determine from these results the extent to which the ratings were affected by this direct information and the extent to which it was affected by other background information such as the stated fact that the pursued person was intolerant of the opposite sex, particularly intolerant of the occupational group to which the pursuer belonged and that

this intolerance was contrary to public consideration for this particular occupational group. It could well be that the information that she/he was determined not to move may have had less effect if given in isolation. Nor can it clearly be related to explicit information in the text why the pursued person became less likely to call the Manager. The important point however is that the additional information modified the expectations of the characters' behaviour regardless of the sex of the character. In other words, the information was related to the role of the character and modified the stereotype of the role.

The relative attributions of behaviour to the sex and the role of the character were unconfounded by comparing the scores on the rating scales in the two Experimental Conditions (where the background information was kept the same but the sexual identities of the characters were switched). The analyses showed clear differences between the roles of pursued and pursuer in the kind of behaviour they would display and this was in accord with the stereotypes elicited. The analyses of variance also showed that in relation to some kinds of behaviour there was an effect which was related to the sex of the character and some interactions between sex and roles of the characters. For example, the pursued person is regarded as more likely to chat up the pursuer than vice versa, and men are more likely to chat up women than women to chat up men. While there is a difference in the likelihood of male and female pursuers chatting up their pursued, a pursued male is regarded as far more likely to chat up his female pursuer than is a pursued female to chat up her male pursuer.

Face-slapping behaviour is rated in the same interactive manner. A pursued woman is far more likely to slap a pursuing male's face than a pursued man is to slap a pursuing woman's face. However, while pursuers are generally unlikely to slap the faces of the pursued, a pursuing male is slightly more

likely to do so than a pursuing female.

From these results it is clear that knowledge about the role of a character influences rating behaviour but that this can be modified by knowledge about behaviour related to stereotyped sexual identities. There is a clear stereotype of likely male and female behaviour but its expression is constrained by knowledge about the role each is enacting in the particular situation.

One part of the stereotype created by the neutral baseline text was that the pursuer was male and the pursued person female. In the Jenny text these default values were confirmed and there was very little change in the characters created as a result of the stereotyped role knowledge being modified by the background information. The one change which might be attributed to the pursuer being a soldier rather than just male is that he would be more likely to chat up Jenny. This effect may be due to the additional information about both characters especially the additional background information about the soldier given in the text. But as there is no such effect when either the nurse or B are the pursuers the result must be related to the 'soldier' and not the 'pursuer' in this case.

When the sex of the character is changed (from that expected of the role) in the James condition it is easily seen how stereotyped knowledge about sex related behaviour interacts with knowledge about role behaviour (these effects were masked in the original Jenny text). For example, it was more likely that the pursuer would call the manager when the pursuer was a female nurse than when the pursuer was stated to be male in the Jenny condition or assumed to be male in the Baseline condition.

The predictive framework which the reader uses to comprehend a text would seem to be one which largely consists of stored general 'world knowledge'. This knowledge about behaviour in a setting and its relation to the sex and role of the character is elicited by information about the setting and the

roles described in the text. This gives a 'skeletal' framework within which the specific anticipatory framework is created.

In this experiment subjects rated predictions of a violent outcome (shooting or stabbing) as being most unlikely. This result lends emphasis to the conclusion to Experiments Three, Four and Five that the "Interestingness" of individual concepts in the text is modified by other factors. In this experiment (Experiment Six) the framework created by readers to anticipate the incoming text is one such factor. Within the framework they have created, the readers do not focus on the "absolutely interesting" event even when there is opportunity to do so. When subjects were told that this unlikely event did happen and that the pursued person stabbed the pursuer with a scalpel, responsibility for this outcome was attributed to the role of pursuer regardless of the sex of the person in that role or the sex of the subject. It might have been expected that, in general, people who stab are more responsible for what happens than those who are stabbed and that men are more likely to stab than women. It might also be expected that subjects would identify with characters of their own sex and hold the opposite sex responsible for what happened. However this particular result suggests that the roles which have been created by the interaction between the characters and the setting in the text cause the reader to reject as inappropriate these more general stereotypes about male and female attackers and their victims however intrinsically interesting these might be as isolated ideas and also.

Models such as those of Schank (1981) and Bransford and Johnson (1973) support the idea of readers constructing a flexible and robust framework to accommodate incoming information. The results of this experiment are most simply interpreted in these terms. A framework is created even when there is the most limited amount of information (as in the Baseline condition). Groups who were given more explicit input from the text were able to utilise more general stored knowledge about the situation described. When such a

framework is created it is unique and related to the information given in a specific text, so that what is described can be accommodated and what is possible can be anticipated. The experiment has not considered how or even whether this framework is stored in memory. It has been concerned with the function of the framework in the comprehension of text. Frameworks may be created and exist only for the function of comprehension of a specific text. Schank (1981) proposed that knowledge about roles and behaviour in situations might be stored as general and specific knowledge and that 'scripts' would be compiled as required. Such a construction would have sufficient flexibility to cope with both general and specialised information in the text. The results of this experiment support this theory.

From the experiment, it can be concluded that when a role is clearly defined readers will draw upon their general stored "world knowledge" to construct a framework which will make sense of the information in the text. This supports the work of Pichert and Anderson (1973) who demonstrated that the same ambiguous description can be selectively processed to provide a specific setting for either the role of prospective burglar or a prospective housebuyer, providing this role is clearly given.

Stored knowledge can be either general world knowledge (the common cultural knowledge referred to in the Introduction to this thesis) or idiosyncratic and related to individual experience (also described in the Introduction). This experiment has shown the importance of general world knowledge in the comprehension of narrative texts and concludes that to be most beneficial in such comprehension it must contain relevant information about the constraints imposed by roles upon settings and by settings upon roles and also the variety of behaviours likely in such role settings.

CONCLUSIONS.Comprehension: the interaction between text and reader.

Comprehension is a term which can be used very loosely. It can be applied to the comprehension of sensory or linguistic events in isolation or in context. The linguistic context can be written or spoken. This thesis has largely considered the question of linguistic comprehension and the experimental work has been confined to the processes involved in the comprehension of written, narrative text.

In the Introduction a tentative definition of comprehension of text was attempted. Further exploration of the definition revealed that it must involve the interaction of the purpose and the knowledge of the reader with the content of the text. Usually narrative texts are read for a very general purpose (for example, entertainment). Other kinds of texts can be read with a specific purpose in mind, the purpose differing between readings and between readers. For example, scientists may read an experimental report to discover the results of an experiment or they may read it to check on the apparatus used, the design or the theoretical ideas it encompasses. With each purpose one section of the text will be the most important and the other sections will be read in relation to this. In such cases it is necessary for the text to be well ordered and the content well presented in a clear and unambiguous manner so that it can be read for each of these purposes. In a similar way other informational texts (Computer manuals, knitting patterns, gardening books etc.) have the same requirements. With narrative texts, however, it is likely that people bring general knowledge to bear upon the text rather than that they extract particular types of information from the text.

Narrative texts are not reliant on a rigid structure. It is likely that the adult reader will only read a narrative text (such as the "Garp" text used

in Experiments One, Two and Six) once and that the overt purpose will be for enjoyment. Such a general purpose will involve the reader in following the plot and the development of situations and characters and in bringing general stereotypes to the text. In this case the text can be structured and ordered in such a way that the reader can speculate about these developments and become involved in what is happening.

Experiments One and Two demonstrated that giving the Outcome of a set of events at the beginning of a text produced speculations that were different in kind from those made by readers who were not given this information. Giving the Outcome first allowed the reader to make tentative hypotheses about future events which went beyond the current episode; it also allowed the reader to link the Outcome and the Events together in an inferential chain. It is suggested that a 'knowledge' framework for these inferences is created by the information in the Outcome (in a way similar to that described by Minsky, 1975 and Kintsch, 1979): the incoming information is slotted into the terminal slots of a frame. Placing the Outcome at the beginning gives the reader a wider framework and a larger set of options within that framework. The reader is able to speculate about what happens after the Outcome event or about what caused the event to happen. Thus the Outcome given at the beginning has a double range of focus. Readers do not only ask 'What will be the consequence?' as they might when the text is given in canonical order; they also ask 'What caused this to happen?'. Placing the Outcome at the beginning enables the reader to create a framework in which many likely consequences and causes can be anticipated.

Additional information which was given by subjects in Experiments One and Two (describing the general personality characteristics of the soldier and Jenny and their reasons for visiting the movie theater) suggest that the reader supplies default values to the frame. This would support the Minsky frame theory. In these experiments the values may have been supplied as a

response to the questions posed and not as a part of the comprehension process, so this is not conclusive evidence for the supply of default values as a part of comprehension. However Experiment Six showed that readers do supply default values of sex of characters to the roles given in the text. When this was contradicted by the information in the text (in the James Condition) it became clear that expectations of role behaviour had also in part been based on the default value of sex of the character and that these too were modified by the information in the text.

The responses which are given by subjects who were not given an Outcome in Experiments One, Two and Six suggest that they process the content in a different way from those who are given the Outcome first. It would seem that they create a frame from their stored knowledge about situations and about characters in a situation in a scene by scene way ( as is described in the Scenario model of Sanford and Garrod, 1981) but that the framework is confined to the episode which is being described. In this situation readers do not speculate beyond the outcome which will end the episode and the kind of outcome which they predict to be compatible with the situation. Experiment Six gave further support to this conclusion by showing that even when readers were given the option of selecting such events as shooting and stabbing as a likely outcome, they still choose from the same range of events as those spontaneously predicted by subjects who were given no outcome information in Experiments One and Two.

It seems that the order of narrative text influences the way in which the reader processes and therefore comprehends the text in a situation which is largely unstructured by a specific purpose. This in turn allows some observations to be made about the comprehension process. Readers have many strategies which they can employ in the comprehension of written text. The structure of narrative text and the informational content dictates which of these strategies is employed. Whichever strategy is employed the reader will



eventually construct an overall framework which encompasses his interpretation of the content based on general and individual knowledge of the world.

The results of Experiments Five seem to indicate that there are important differences in the processing of texts when they are presented in flashback sequence rather than in canonical order. Experiments 1 and 2 suggested that the flashback strategy allowed the reader to construct a framework from the text structure and to build a picture of the situation and the characters which differed from that constructed by subjects who did not have the flashback structure. They did not, however, have the same amount of information when it was presented in these two forms.

Experiment Five provided further evidence for the effect of story format, showing that processing is affected by different orders of text presentation. The results demonstrated that readers selected an 'Interesting' idea for further reading when it was presented at the beginning of a text but that this same text was not necessarily selected when it was presented in the body of the text. These results are consistent with the idea that processing is related to the initial input rather than to the input of additional 'interesting' ideas and that the reader will search for information which fits the predictive schema constructed by the structure and content of information already presented. When an 'interesting idea' is processed at the beginning of a text it may well have a marked effect as it will be influential in the initial creation of an interpretative framework. However, when presented in context an 'interesting idea' will have a less noticeable effect unless it in some way extends or modifies the predictive schema which has been constructed by the reader.

Ideas or concepts may be interesting in isolation because they allow easy access to a stored schema of world knowledge. In context these schemata interact with the structure and order of the text and what is processed is an interaction of the three factors and not solely a function of the intrinsic

interestingness of the conceptual ideas.

The results of Experiment Six allow the framework to be described with more precision. The results show that a unique framework is actively constructed for the interpretation of a text. Each of the three texts had a different framework. The experimental results have shown that the framework is evoked by the text but that the predictions are based on inferred information about default values given in the text. The frameworks which were created in response to the texts encompassed both the readers' world knowledge about behaviour related to both the sex and to the roles of characters in a situation and the information given in the texts which modified that knowledge. (Interestingly, the world knowledge is that shared by all readers regardless of their sex and individual knowledge.)

The framework which is created is sufficiently flexible to accommodate unanticipated information (for example, the stabbing of the pursuer by the pursued). This information is incorporated into the framework that has already been created (in the example given the information is related to the pursued/pursuer roles) and not in relation to isolated world knowledge about victims and attackers in a stabbing incidents. This again supports the proposal that a framework is created and modified as the text progresses and contradicts the idea that a concluding piece of information is the key factor in the interpretation of the text.

#### General and Educational Implications

This research has shown that the frame work within which a text is comprehended is evoked and guided by the information in the text but in the main it is created from stored general world knowledge. In the experiments it was shown that readers can and do use the framework to anticipate future

events and to make predictions about what is likely to happen. However, it is not claimed that this is the sole function of the framework. This predictive function was accessed by the methodology used and the type of questions asked. The first two experiments also investigated other functions of the framework and it was shown that a general and common world knowledge is used to supply such background details as the motives of the characters, their contribution to what happened and their more stable personality characteristics. It would seem from this that the framework has a more general function: that of giving an overall coherence to the story and thus providing a basis for overall inference making when it becomes necessary to read "Between the lines" to answer external or self imposed questions about the situation.

The research and the conclusions drawn from it show that general world knowledge plays an important part in the comprehension of text. In Chapter One of this thesis, approaches to the teaching of comprehension and reading were reviewed and attention was drawn to the similarities and differences between two approaches: one that emphasised the extraction of meaning from text via a multiplicity of decoding skills and the other which emphasised the focussing of prior knowledge and experience to bring meaning to the text. One conclusion about the two approaches was that in the teaching of reading they were complementary rather than alternative approaches and that successful teaching of reading depended on new skills being taught in relation to existing knowledge. This research has looked at the comprehension process in skilled readers and has made the assumption that decoding skills and the ability to "read the lines" have already been acquired. It has shown the importance of world knowledge in the comprehension of text once the reader can read the lines even when the information they give is as fragmentary as that given in the Baseline text.

The current debate in education rests on a distinction between comprehension and learning ; comprehension being defined in this instance (Gardner,1982) as the general ability to reflect on the text being read. This approach defines reading as absorbing the information in the text by reading "between the lines". The experimental work in this thesis has shown that comprehension is an active constructive process involving reading the lines, between the lines and going beyond the lines to incorporate the readers' general world knowledge. The reader will learn and comprehend by creating a framework which will be evoked and guided by the information in the text. What it is he/she learns will depend upon the questions asked (set either by his /her own purpose or imposed externally by a teacher) and the relevance of the content of the text in supplying the answer. However, neither comprehension nor learning will take place if the reader has no world knowledge relevant to the text or the question. To ensure that the readers learns a specific piece of information from a text then the teacher must ensure that they have the necessary prior level of specific world knowledge and that they are directed towards that information either by a specific question or instruction or ideally by their own questions and purposes.

#### Methodology.

One of the points made in Chapter Nine (the summary of the review of models and theories about text processing) was that there are no immediately obvious methods available for the exploration of the processes involved in the comprehension of complex text. The main methodological problem is that of separating comprehension of text from recall. This is not a problem which can be approached directly. To be available for measurement a just-comprehended text must be additionally processed so that it can be presented in a form which can be measured (for example, recalled or accessed for answering questions). This involves a recoding of what has been comprehended and this

(as for example was demonstrated by Pichert and Anderson, 1977) is influenced by the probe which is used to elicit the presentation. So 'natural' comprehension can never be directly measured: the more direct the measurement technique, the more 'unnatural' the task becomes. On the other hand, while indirect measures (such as unanticipated recall) do not interfere with the comprehension process itself they tend to provide data contaminated by the effects of other processes.

The analysis of the predictions that readers make on the basis of what they have comprehended provides an indirect measure of comprehension which is relatively free from these contaminating effects. Such an analysis was used in Experiments One and Two when the predictions were mainly elicited by the use of open-ended questions. Open-ended questions allowed the subjects to freely express their predictions about preceding and future events and their hypotheses about the characters. The subsequent content analysis allowed categories to be derived from the responses rather than the categories being imposed upon the readers. It is not claimed that this method will measure everything that has been comprehended; only that it will allow readers to express some of the content of what has been comprehended rather than what might subsequently be recalled.

This method was used to explore the comprehension process rather than to test theories about it. The analyses reported demonstrated that reliable results could be obtained. However, the aim was to develop a means of exploring some of the procedures which might be utilised by readers and the external validity of a method which analyses freely given responses cannot be disputed. Further investigations are envisaged using this method on other texts which are similar in construction but differ in content. These will investigate the reliability of the method when applied more widely.

The methodology used in Experiments Three, Four, Five (choosing sentences to read more about and rating them for 'Interestingness' ) and Six reflects

the more traditional approach and indicates that specific questions might be examined more rigorously once they have been identified by the more explorative method. In particular Experiment Six showed that results comparable to those of Experiments One and Two can be obtained when alternatives are made explicit and that high levels of agreement between subjects obtained. Experiments One and Two provide a methodology for obtaining readers' predictions; Experiment Six provides a methodology for experimentally investigating the effect of text and reader related variables on such predictions. The present work has been based on the premise that comprehension is a purposeful activity. The development of techniques which enable us to measure this activity (as expressed in the reader's ongoing predictions and forward-looking inferences) is crucial to the development of adequate theories of comprehension.

REFERENCES.

- AARONSON, D. and SCARBOROUGH, H. S. (1977). Performance theories for sentence coding: some quantitative models. Journal of Verbal Learning and Verbal Behavior, 16, 277-303.
- ANDERSON, A. (1980). An Empirical Investigation of Indirect Scenario Changes. Paper presented at the British Psychological Society Cognitive Section Conference on Reading. Exeter University.
- ANDERSON, J. R. (1976). Language, Memory and Thought. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- ANDERSON, J. R. and BOWER, G. H. (1973). Human Associative Memory. Washington, D.C.: Winston.
- ANDERSON, J. R. and REDER, L. M. (1979). An elaborative processing explanation of depth of processing. In Cermak, L.S. and Craik, F.I.M. (Eds.) Levels of Processing in Human Memory. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- ANDERSON, R. C. and 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., GOETZ, E. T. and SCHALLERT, D. L. (1976). Frameworks for comprehending discourse. Technical Report No.12, Urbana, Illinois: Center for the Study of Reading, University of Illinois at Urbana-Champaign.
- ANDERSON, R. C., SPIRO, R. J., and ANDERSON, M. C. (1978). Schemata as scaffolding for the representation of information in discourse. American Educational Research Journal, 15, 433-440.

- ATKINSON, R. C. and SHIFFRIN, R. M. (1968). Human Memory: a proposed system and its control processes. In Spence, K.W. and Spence, J.T. (Eds.) The Psychology of Learning and Motivation: Advances in Research and Theory, 2, New York: Academic Press.
- ARTLEY, A. S. (1942). A study of certain relationships existing between general comprehension and reading comprehension in a specific subject matter area. Unpublished doctoral dissertation, The Pennsylvania State College.
- BADDELEY, A. D. (1976). The Psychology of Memory. New York: Harper and Row .
- BARRETT, T. C. (1972). Unpublished paper quoted by Clymer, T. What is reading? Some current concepts. In Melnik, A. and Merritt, J. (Eds.) Reading. Today and Tomorrow. London: University of London Press in association with The Open University Press.
- BARTLETT, F. C. (1932). Remembering. Cambridge, England: Cambridge University Press.
- BEAUGRANDE, R. De. (1980) Text, Discourse, and Process: Towards a Multidisciplinary Science of Texts. Norwood, New Jersey: Ablex .
- BOWER, G. H. (1976). Experiments on story understanding and recall. Quarterly Journal of Experimental Psychology, 28, 511-534.
- BOWER, G. H., BLACK, J. B. and TURNER, T. J. (1979). Scripts in memory for text. Cognitive Psychology, 11, 177-220.
- BRADSHAW, G. L. and 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. and FRANKS, J. J. (1972). Sentence memory: a constructive versus interpretive approach. Cognitive Psychology, 3, 193-209.
- BRANSFORD, J. D. and FRANKS, J. J. (1971). The abstraction of linguistic ideas. Cognitive Psychology, 2, 331-350.



- BRANSFORD, J. D. and JOHNSON, M. K. (1972). Contextual prerequisites for understanding: some investigations of comprehension and recall. Journal of Verbal Learning and Verbal Behavior, 11, 717-726.
- BRANSFORD, J. D. and JOHNSON, M. K. (1973). Consideration of some problems of comprehension. In Chase, W. (Ed.) Visual Information Processing. New York: Academic Press.
- BRANSFORD, J. D. and MCCARRELL, N. S. (1977). A sketch of a cognitive approach to comprehension: some thoughts about understanding what it means to comprehend. In Johnson-Laird, P.N. and Wason, P.C. Thinking: Readings in Cognitive Science. Cambridge: Cambridge University Press.
- BRANSFORD, J. D., MCCARRELL, N. S., FRANKS, J. J. and NITSCH, K. E. (1977). Towards unexplaining memory. In Shaw, R. and Bransford, J.D. (Eds.) Perceiving, Acting and Knowing. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- BRANSFORD, J. D., STEIN, B. S., SHELTON, T. S. and OWINGS, R. A. (1981). Cognition and adaptation: the importance of learning to learn. In Harvey, J.H. (Ed.) Cognition, Social Behavior and the Environment. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- BROWN, A. L., SMILEY, S. S., DAY, J. D., TOWNSEND, M. A. R. and LAWTON, S. C. (1977). Intrusion of a thematic idea in children's comprehension and retention of stories. Child Development, 48, 1454-1466.
- CACCAMISE, D. and KINTSCH, W. (1978). Recognition of important and unimportant statements from stories. American Journal of Psychology, 91, 651-657.
- CARRILLO, L. (1967). The language-experience approach to the teaching of reading. In Downing, J. and Brown, H. (Eds.) The Second International Reading Symposium. London: Cassell.
- CHAFE, W. (1972). Discourse structure and human knowledge. In Carroll, J.B. and Freedle, R.O. (Eds.) Language Comprehension and the Acquisition of Knowledge. Washington: Winston.

- CHARNIAK, E. (1972). Towards a model of children's story comprehension. MIT AI Laboratory Memorandum AI-TR-266, Cambridge Mass.
- CHOMSKY, N. (1957). Syntactic Structures. The Hague: Mouton.
- CHOMSKY, N. (1965). Aspects of the Theory of Syntax. Cambridge, Massachusetts: M.I.T. Press.
- CIRILO, R. and FOSS, D. (1980). Text structure and reading times for sentences. Journal of Verbal Learning and Verbal Behavior, 19, 96-109.
- COFER, C. N. (1976). An historical perspective. In Cofer C.N. (Ed.) The Structure of Human Memory. San Francisco: W. H. Freeman and Company.
- COLLINS, A., BROWN, J. S. and LARKIN, K. M. (1980). Inference in text understanding. In Spiro, R.J., Bruce, B.C. and Brewer, W.F. (Eds.) Theoretical Issues in Text Understanding. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- COLLINS, A. M. and LOFTUS, E. F. (1975). A spreading activation theory of semantic processing. Psychological Review, 82, 407-428.
- COLLINS, A. M. and QUILLIAN, M. R. (1969). Retrieval time from semantic memory. Journal of Verbal Learning and Verbal Behavior, 8, 240-247.
- COLLINS, A. M. and QUILLIAN, M. R. (1970). Does category size affect categorization? Journal of Verbal Learning and Verbal Behavior, 9, 432-453.
- COLLINS, A. M. and QUILLIAN, M. R. (1972a). Experiments on semantic memory and language comprehension. In Gregg, L. W. (Ed.) Cognition in Learning and Memory. New York: Wiley.
- COLLINS, A. M. and QUILLIAN, M. R. (1972b). How to make a language user. In Tulving, E. and Donaldson, W. (Eds.) Organisation of Memory. New York: Academic Press.
- CONANT, M. M. (1942). The Construction of a Diagnostic Reading Test. New York: Teachers College, Columbia University. Contributions to Education No.861.

- CRAIK, F. I. M. and LOCKHART, R. S. (1972). Levels of processing, a framework for memory research. Journal of Verbal Learning and Verbal Behavior, 11, 123-136.
- CRAIK, F. I. M. and TULVING, E. (1975). Depth of processing and retention of words in episodic memory. Journal of Experimental Psychology: General, 104, 268-294.
- DAVIS, F. B. (1941). Fundamentaental factors of comprehension in reading. Unpublished doctoral dissertation, Graduate School of Education, Havard University.
- DAVIS, F. B. (1944). Fundamental factors of comprehension in reading. Psychometrika, 9, 185-197.
- DAVIS, F. B. (1946). A brief comment on Thurstone's note on a reanalysis of Davis' Reading Tests. Psychometrika, 11, 249-245.
- DAVIS, F. B. (1968). Research in comprehension in reading. Reading Research Quarterly, 3, 499-545.
- DAVIS, F. B. (1971). Psychometric Research on Comprehension in Reading. Graduate Schol of Education, Rutgers University, New Brunswick, New Jersey.
- D.E.S. REPORT, (1975). A Language for Life. Report of the committee of Inquiry appointed by the Secretary of State for Education and Science under the Chairmanship of Sir Alan Bullock FBA.
- DOWNING, J. (1972). Teaching reading through the medium of i.t.a. In Morris, J.M. (Ed.) The First R., London : Ward Lock Educational.
- EDINBURGH READING TEST, (1972). Moray House College of Education in association with the Scottish Education Department and the Educational Institute of Scotland.
- FILLMORE, C. J. (1968). The case for case. In Bach, E. and Harms, R.T. (Eds.), Universals in Linguistic Theory. New York: Holt, Rinehart and Winston.

- FILLMORE, C. J. (1971). Some problems for case grammar. In O'Brien, R.J. Linguistics: Developments of the Sixties-Viewpoints for the Seventies. Monograph Series on Languages and Linguistics. 24, 35-36.
- FISHER, R. P. and CRAIK, F. I. M. (1977). Interaction between encoding and retrieval operations in cued recall. Journal of Experimental Psychology: Human Learning and Memory, 3, 701-711.
- FODOR, J. A. and GARRETT, M. (1967). Some syntactic determinants of sentential complexity. Perception and Psychophysics, 2, 289-296.
- FOSS, D. J. and HARWOOD, D. A. (1975). Memory for sentences: implications for human associative memory. Journal of Verbal Learning and Verbal Behavior, 14, 1-16.
- FRIES, C. C. (1962). Linguistics and Reading. New York: Holt, Rinehart and Winston.
- GAGNE, R. M. (1965). The Conditions of Learning. New York: Holt Rinehart and Winston.
- GALAMBOS, J. A. and RIPS, L. J. (1982). Memory for routines. Journal of Verbal Learning and Verbal Behavior, 21, 260-281.
- GANS, R. A. (1940). A Study of Critical Reading Comprehension in the Intermediate Grades. New York: Teacher's College, Columbia University. Contribution to Education No.811.
- GARDNER, K. (1982). Learning from Reading. Internal Memo to Schools Council Reading for Learning in the Secondary School Project. Prepared at the School of Education, University of Nottingham.
- GARROD, S. C. and SANFORD A. J. (1977). Interpreting anaphoric relations: the integration of semantic information while reading. Journal of Verbal Learning and Verbal Behavior, 16, 77-90.
- GARROD, S. C. and SANFORD, A. J. (1978). Anaphora: a problem in text comprehension. In Campbell, R.N. and Smith, P.T. (Eds.) Recent Advances in the Psychology of Language. New York: Plenum Press.

- GATTEGNO, C. (1962). Words in Colour: Background and Principles. Reading: Educational Explorers.
- GIBBS, R. W. and TENNEY, Y. J. (1980). The concept of scripts in understanding stories. Journal of Psycholinguistic Research, 9, 275-284.
- GOETZ, E. T. (1977). Inferences in the comprehension of and memory for text. Technical Report No.49. Cambridge, Massachusetts: Bolt, Beranek and Newman Inc.
- GOODMAN, K. S. (1967). Reading: a psycholinguistic guessing game. Journal of the Reading Specialist, 4, 126-135.
- GOODMAN, K. S. (1969). Analysis of oral reading miscues: applied psycholinguistics. Reading Research Quarterly, 5, 9-30.
- GOODMAN, K. S. (1970). Psycholinguistic universals in the reading process. Journal of Typographical Research, 4, 103-110.
- GRAY, W. S. (1960). The major aspects of reading. In Robinson, H.M. (Ed.) Sequential Development of Reading Abilities. Supplementary Educational Monographs, 90. Chicago: University of Chicago Press.
- GREENE, J. (1972). Psycholinguistics: Chomsky and Psychology. Harmondsworth, Middlesex: Penguin.
- GREGG, V. (1978). Recognition and recall. In Cognitive Psychology: Memory (part 2). Milton Keynes: Open University Press.
- GRICE, H. P. (1975). Logic and Conversation. In Cole, P. and Morgan, J.L. (Eds.) Syntax and Semantics, Vol 3: Speech Acts. New York: Seminar Press.
- GUILFORD, J. P. and FRUCHTER, B. (1973). Fundamental Statistics in Psychology and Education. Tokyo: McGraw-Hill Kogakusha Ltd.
- HALL, W. E. and ROBINSON, F. P. (1945). An analytical approach to the study of reading skills. Journal of Educational Psychology, 36, 429-442.
- HALLIDAY, M. A. K. (1967). Notes on transitivity and theme in English, Part 1. Journal of Linguistics, 3, 37-81.

- HARRIS, C. W. (1948). Measurement of comprehension in literature: 11 studies of measures of comprehension. School Review, 56, 332-342.
- HAVILAND, S. E. and CLARK, H. H. (1974). What's new? Acquiring new information as a process in comprehension. Journal of Verbal Learning and Verbal Behavior, 13, 515-521.
- HAYES-ROTH, B. and THORNDYKE, P. W. (1979). Integration of knowledge from text. Journal of Verbal Learning and Verbal Behavior, 18, 91-108.
- HERRIOT, P. (1969). The comprehension of active and passive sentences as a function of pragmatic expectations. Journal of Verbal Learning and Verbal Behavior, 8, 166-169.
- HOLYOAK, K. J. and GLASS, A. L. (1975). The role of contradictions and counterexamples in the rejection of false sentences. Journal of Verbal Learning and Verbal Behavior, 14, 215-239.
- HORNBY, P. A. (1972). The psychological subject and predicate. Cognitive Psychology, 3, 632-42.
- HUNT, L. C. (1952). A further study of certain factors associated with reading comprehension. Unpublished doctoral dissertation, School of Education, Syracuse University.
- JARVELLA, R. J. (1970). Syntactic processing of connected speech. Journal of Verbal Learning and Verbal Behavior, 10, 409-416.
- JOHNSON, M. K., BRANSFORD, J. D. and SOLOMON, S. K. (1973). Short reports: memory for tacit implications of sentences. Journal of Experimental Psychology, 98, 203-205.
- JOHNSON-LAIRD, P. N. and STEVENSON, R. (1970). Memory for syntax. Nature, 227, 412.
- JORESKOG, K. G. and VAN THILLO, M. (1971). New rapid algorithms for factor analysis by unweighted least squares, generalised least squares and maximum likelihood. Research Memorandum, 71-75, Princeton, New Jersey: Educational Testing Service.

- KAISER, H. F. (1970). A second generation little jiffy. Psychometrika, 35, 401-415.
- KATZ, J. J. and FODOR, J. A. (1963). The structure of a semantic theory. Language, 39, 170-210.
- KATZ, J. J. and POSTAL, P. M. (1964). An Integrated Theory of Linguistic Descriptions. Cambridge, Massachusetts: M.I.T. Press.
- 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.
- KINTSCH, W. (1972). Notes on the structure of semantic memory. In Tulving, E. and Donaldson, W. (Eds.) Organisation of Memory. London: Academic Press.
- KINTSCH, W. (1974). The Representation of Meaning in Memory. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- KINTSCH, W. (1976). Memory for prose. In Cofer, C.N. The Structure of Human Memory. San Francisco: Freeman and Company.
- KINTSCH, W. (1977). On comprehending stories. In Just, M.A. and Carpenter, P.A. Cognitive processes in Comprehension. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- KINTSCH, W. (1979). On modeling comprehension. Educational Psychologist, 14, 3-14.
- KINTSCH, W. and GLASS, G. (1974). In Kintsch, W. The Representation of Meaning in Memory. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- KINTSCH, W. and KEENAN, J. (1973). Reading rate and retention as a function of the number of propositions in the base structure of sentences. Cognitive Psychology, 5, 257-274.
- KINTSCH, W., KOZMINSKY, E., STREBY, W. J., MCKOON, G. and 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. and MONK, D. (1972). Storage of complex information in memory: Some implications of the speed with which inferences can be made. Journal of Experimental Psychology, 94, 25-32.
- KINTSCH, W and VAN DIJK, T. A. (1978). Toward a model of text comprehension. Psychological Review, 85, 363-394.
- KINTSCH, W. and VIPOND, D. (1979). Reading comprehension and readability in educational practice and psychological theory. In Nilsson, L-G. (Ed.), Perspectives on Memory Research. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- KLARE, G. R. (1974). Assessing readability. Reading Research Quarterly, 10, 62-102.
- KLEIN, K. and SALTZ, E. (1976). Specifying the mechanisms in a levels-of-processing approach to memory. Journal of Experimental Psychology: human Learning and Memory, 2, 671-679.
- LANGSAM, R. (1941). A Factorial Analysis of Reading Ability. Unpublished doctoral dissertation, New York University. Abstract in Journal of Experimental Education, 10, 57-63.
- LENNON, R. T. (1962). What can be measured? The Reading Teacher, 15, 326-337.
- LESGOLD, A. M., ROTH, S. F., and CURTIS, M. E. (1979). Foregrounding effects in discourse comprehension. Journal of Verbal Learning and Verbal Behavior, 18, 291-308.
- LUNZER, E. A. (1979). Analysis of the text and the reader. In Pugh, A.K. and Lunzer, E.A. Language in Use: Talk and Text. Milton Keynes: Open University Press.
- LUNZER, E. A., WAITE, M. and DOLAN, T. (1979). Comprehension and comprehension tests. In Lunzer, E. and Gardener, K. (Eds.) The Effective Use of Reading. London: Heinemann Books for the Schools Council.
- LYONS, J. (1968). Introduction to Theoretical Linguistics. Cambridge: Cambridge University Press.



- LYONS, J. (1970). Generative syntax. In Lyons, J. (Ed.) New Horizons in Linguistics. Harmondsworth, Middlesex: Penguin.
- MANDLER, J. M. and JOHNSON, N. S. (1977). Remembrance of things parsed: Story structure and recall. Cognitive Psychology, 9, 111-151.
- MANDLER, J. M. (1978). A code in the node: The use of story schema in retrieval. Discourse Processes, 1, 14-35.
- MANDLER, J. M. and PARKER, R. E. (1976). Memory for descriptive and spatial information in complex pictures. Journal of Experimental Psychology: Human Learning and Memory, 2, 38-48.
- MANEY, E. S. (1952). Literal and critical reading in science. Unpublished doctoral dissertation, Temple University.
- MASSON, M. E. J. (1979). Context and inferential cuing of sentence recall. Journal of Verbal Learning and Verbal Behavior, 1, 173-185.
- MEYER, D. E. (1970). On the representation and retrieval of stored semantic information. Cognitive Psychology, 1, 242-300.
- MILLER, G. A. and MCKEAN, K. E. (1964). A chronometric study of some relationships between sentences. Quarterly Journal of Experimental Psychology, 16, 297-308.
- MINSKY, M. (1975). Frame-system theory. In Schank, R.C. and Nash-Webber, B.L. (Eds.) Theoretical Issues in Natural Language Processing. Preprints of a conference at MIT (June 1975). Reprinted in Johnson-Laird, P.N. and Wason, P.C.(Eds.) Thinking. Cambridge: Cambridge University Press, (1977).
- MOYLE, D. (1972). The Teaching of Reading. London: Ward Lock Educational.
- NEALE, M. D. (1957). Neale Analysis of Reading Ability. London: Macmillan.
- NEISSER, U. (1967). Cognitive Psychology. New York: Appleton-Century-Crofts.
- NEISSER, U. (1976). Cognition and Reality. San Francisco: Freeman.
- NEZWORSKI, T., STEIN, N. L. and TRABASSO, T. (1982). Story structure and content in children's recall. Journal of Verbal Learning and Verbal Behavior, 21, 196-206.

- NORMAN, D. A., RUMELHART, D. E. and the LNR RESEARCH GROUP (1975).  
Explorations in Cognition, San Francisco: Freeman.
- PICHERT, J. W. and ANDERSON, R. C. (1977). Taking different perspectives on a story. Journal of Educational Psychology, 69, 309-315.
- PRATT, M. W., LUSZCZ, M. A., MACKENZIE-KEATING, S. and MANNING, A. (1982). Thinking about stories: The story schema in metacognition. Journal of Verbal Learning and Verbal Behavior, 21, 493-505.
- PURKISS, E. (1978). The effect of foregrounding on pronominal reference. Unpublished undergraduate thesis, University of Glasgow.
- QUILLIAN, M. R. (1969). The teachable language comprehender: a simulation program and theory of language. Communications of the Association for Computing Machinery, 12, 459-476.
- RAUCH, S. J. (1972). Reading in the total school curriculum. In Melnik, A. and Merritt, J. (Eds.) The Reading Curriculum. London: The University of London Press in Association with the Open University Press.
- REDER, L. M. (1976). The Role of Elaboration in the Processing of Prose. Unpublished Doctoral Dissertation, University of Michigan.
- REDER, L. M. (1979). The role of elaborations in memory for prose. Cognitive Psychology, 11, 221-234.
- REDER, L. M. and ANDERSON, J. R. (1980). Partial resolution of the paradox of interference: the role of integrating knowledge. Cognitive Psychology, 12, 447-472.
- RIPS, L. J., SHOBE, E. J. and SMITH, E. E. (1973). Semantic distance and the verification of semantic relations. Journal of Verbal Learning and Verbal Behavior, 12, 1-20.
- ROBINSON, H. A. (1972). Utilising the learner's experience in comprehension. In Morris, J.M. (Ed.) The First R: Yesterday, Today and Tomorrow. London: Ward Lock Educational.

- RUMELHART, D. E. (1975). Notes on a schema for stories. In Bobrow, D.G. and Collins, A. (Eds.) Notes on a Schema for Stories. New York: Academic Press.
- RUMELHART, D. E. (1977). Understanding and summarising brief stories. In Laberge, D. and Samuels, S.J. (Eds.) Basic Processes in Reading: Perception and Comprehension. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- RUMELHART, D. E., LINDSAY, P. H. and NORMAN, D. A. (1972). A process model for long-term memory. In Tulving, E. and Donaldson, W. (Eds.), Organisation of Memory, New York: Academic Press.
- SACHS, J. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. Perception and Psychophysics, 2, 437-442.
- SANFORD, A. J. and GARROD, S. C. (1978). Memory and attention in text comprehension: the problem of reference. Paper presented at the 8th International Attention and Performance Conference, Princeton, New Jersey.
- SANFORD, A. J. and GARROD, S. C. (1981). Understanding Written Language: Explorations of Comprehension beyond the Sentence. Chichester: John Wiley and Sons.
- SAVIN, H. B. and PERCHONOCK, E. (1965). Grammatical structure and the immediate recall of English sentences Journal of Verbal Learning and Verbal Behaviour, 4, 348-353.
- SCHAEFFER, B. and WALLACE, R. (1969). Semantic similarity and the comparison of word meanings. Journal of Experimental Psychology, 82, 343-346.
- SCHANK, R. C. (1975a). The role of memory in language processing. In Cofer, C.N (Ed.) The Structure of Human Memory, San Francisco: Freeman and Co.
- SCHANK, R. C. (1975b). The structure of episodes in memory. In Bobrow, D. and Collins, A. (Eds.) Representation and Understanding, New York: Academic Press.

- SCHANK, R. C. (1978). Interestingness: Controlling Inferences. Research Report £ 145. Yale University Department of Computer Science.
- SCHANK, R. C. (1981). Language and memory. In Norman, D. (Ed.) Perspectives on Cognitive Science. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- SCHANK, R. C and ABELSON, R. P. (1975). Scripts, plans and knowledge. In Johnson-Laird, P. N. and Wason P. C. (Eds.) Thinking. Cambridge: Cambridge University Press.
- SCHANK, R. C. and ABELSON, R. P. (1977). Scripts, Plans, Goals and Understanding. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- SHOBEN, E. J. (1980). Theories of semantic memory: approaches to knowledge and sentence comprehension. In Spiro, R.J., Bruce, B.C. and Brewer, W.F. (Eds.) Theoretical Issues in Text Comprehension. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- SLACK, J. (1978). Semantic memory. In Cognitive Psychology: Memory (part 2). Milton Keynes: Open University Press.
- SMITH, E. E. (1978). Theories of semantic memory. In Estes, W.K. (Ed.) Handbook of Learning and Cognitive Processes. (Vol 6) Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- SMITH, E. E., SHOBEN, E. J. and RIPS, L. U. (1974). A featural model for semantic decision. Psychological Review, 81, 214-241.
- SOCHOR, E. E. (1952). Literal and Critical reading in Social Studies. Unpublished doctoral dissertation, Temple University.
- SPIRO, R. J., (1980). Accomodative reconstruction in prose recall. Journal of Verbal Learning and Verbal Behavior, 19, 84-95.
- SPEARRITT, D. (1972). Identification of subskills in reading comprehension by maximum likelihood factor analysis. Reading Research Quarterly, 8, 92-111.
- S.R.A. READING LABORATORY, 1A. (1961). Chicago: Science Research Associates.

- STEBBING, J. and RABAN, B (1982). Reading for meaning: an investigation of the effect of narrative in two reading books for seven year olds. Reading, 16, 153-162.
- STEIN, B. S., MORRIS, C. D., and BRANSFORD, J. D. (1978). Constraints on effective elaboration. Journal of Verbal Learning and Verbal Behavior, 17, 707-714.
- STEIN, B. S. and BRANSFORD, J. D. (1979). Constraints on effective elaboration: effects of precision and subject generation. Journal of Verbal Learning and Verbal Behavior, 18, 769-777.
- STEIN, N. L. and GLENN, C. G. (1979). An analysis of story comprehension in elementary school children. In Freedle, R. (Ed.) New Advances in Discourse Processing, VII: Advances in Discourse Processing. Norwood, New Jersey: Ablex Publishing Company.
- STOKER, H. W. and KROPP, R. P. (1960). The predictive validities and factorial content of the Florida State-wide, ninth-grade testing program battery. Florida Journal of Educational Research, 60, 105-114.
- SULIN, R. A. and DOOLING, D. J. (1974). Intrusion of a thematic idea in retention of prose. Journal of Experimental Psychology, 103, 255-262.
- THORNDIKE, E. L. and LORGE, I. (1959). The Teacher's Word book of 30,000 words. New York: Columbia Teachers' College.
- THORNDIKE, R. L. (1971). Reading as reasoning. Address delivered to the American Psychological Association, Washington, D.C.
- THORNDYKE, P. W. (1977). Cognitive structures in comprehension and memory of narrative discourse. Cognitive Psychology, 9, 77-110.
- THORNDYKE, P. W. (1978). Knowledge transfer in learning from texts. In Lesgold, A.M. and Pellegrino, J.W. (Eds.), Cognitive Psychology and Instruction. New York: Plenum Press.
- THORNDYKE, P. W. and HAYES-ROTH, B. (1979). The use of schemata in the acquisition and transfer of knowledge. Cognitive Psychology, 11, 82-106.

- THURSTONE, L. L. (1946). Note on a reanalysis of Davis' Reading Tests. Psychometrika, 11, 185-188.
- TRAXLER, A. E. (1941). Problems of measurement in reading. Proceedings of the 1941 Invitational Conference on Testing Problems. American Council on Education, 65-73.
- TULVING, E. (1972). Episodic and semantic memory. In Tulving, E. and Donaldson, W. (Eds.), Organisation of Memory, New York: Academic Press.
- TULVING, E. and THOMSON, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. Psychological Review, 80, 352-373.
- TURNER, A. and GREENE, E. (1977). The Construction of a Propositional Text Base. University of Colorado Technical Report.
- VAN DIJK, T. A. (1972). Some Aspects of Text Grammars. The Hague: Mouton.
- VAN DIJK, T. A. (1977). Semantic macro-structures and knowledge frames in discourse comprehension. In Just, M.A. and Carpenter, P.A. (Eds.), Cognitive Processes in Comprehension, Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- VAN DIJK, T. A. (1977a). Text and Context. New York: Longman.
- VIPOND, D. (1980). Micro- and macroprocesses in text comprehension. Journal of Verbal Learning and Verbal Behaviour, 19, 276-296.
- WARD LOCK EDUCATIONAL READING WORKSHOPS, (1969, 1971). London: Ward Lock Educational.
- WEXLER, K. (1978). A review of John R. Anderson's Language, Memory, and Thought. Cognition, 6, 327-351.
- WHYTE, J. (1979). Language and Reading: Some insights from a study of Adult Literacy Students. Paper presented to the British Psychological Society Conference; London, December 1979.
- WILKINS, A. J. (1971). Conjoint frequency, category size, and categorisation time. Journal of Verbal Learning and Verbal Behavior, 10, 382-385

**BLANK PAGE  
IN  
ORIGINAL**

**TEXT BOUND CLOSE TO THE SPINE IN  
THE ORIGINAL THESIS**



You have just read part of the opening paragraph of a published book. The text is as it appears in the book. It has not been specially written for this investigation. If you have read the book before, or if you are aware of the book from which this has been taken, please let me know before you answer the questions.

Please answer the questions as spontaneously as possible within the spaces given. You may write in note form if you wish.

---

1. What is happening?

---

2. What might have happened prior to the events described?

---

3. What might happen next?

---

4. What in particular would you like to know more about?

---

5. Two characters are introduced in this passage. Give three single words to describe:

	Him	Her
1.		
2.		
3.		

6. Who do you think will be the central character in the narrative which follows the extract you have read?

7. Based on the ideas in the opening passage, what wider more general issues do you think the rest of the book might encompass?

8. The following adjectives could be used to describe the characters in the book. For each pair of adjectives, circle the one you think best describes that character. If you think neither are relevant, then circle NA.

a) Rate the following in terms of how you feel the characters reacted to the situation described in the text.

HIM			HER		
Relaxed	Tense	NA	Relaxed	Tense	NA
Afraid	Unafraid	NA	Afraid	Unafraid	NA
Composed	Excitable	NA	Composed	Excitable	NA
Assured	Anxious	NA	Assured	Anxious	NA
Rational	Emotional	NA	Rational	Emotional	NA
Controlled	Uncontrolled	NA	Controlled	Uncontrolled	NA

b) Rate the following in terms of what you feel to be the character's more general personality traits.

HIM			HER		
Aloof	Sociable	NA	Aloof	Sociable	NA
Self-confident	Insecure	NA	Self-confident	Insecure	NA
Tough-minded	Sensitive	NA	Tough-minded	Sensitive	NA
Tolerant	Intolerant	NA	Tolerant	Intolerant	NA
Friendly	Hostile	NA	Friendly	Hostile	NA
Self-reliant	Dependent	NA	Self-reliant	Dependent	NA

9. Do you think that:

The movie theater was fairly crowded or empty?	CROWDED	EMPTY
The characters already knew each other?	YES	NO
What happened was a new experience for him?	YES	NO
What happened was a new experience for her?	YES	NO

10.

The extract from the book briefly outlines events which happened to two people. What additional information would you like about each of the characters in order to gain a better understanding of the events?

Please rate each of the items below from 1 to 5 for each character. Use 1 if you think information about the item is very important, through to 5 if you think it is totally unimportant. Indicate your decision by circling the appropriate number.

	HIM					HER				
Age:	1	2	3	4	5	1	2	3	4	5
Nature of Occupation:	1	2	3	4	5	1	2	3	4	5
Educational background:	1	2	3	4	5	1	2	3	4	5
Life Style:	1	2	3	4	5	1	2	3	4	5
Disposition:	1	2	3	4	5	1	2	3	4	5
Appearance:	1	2	3	4	5	1	2	3	4	5
a) Height	1	2	3	4	5	1	2	3	4	5
b) Build	1	2	3	4	5	1	2	3	4	5
c) Features	1	2	3	4	5	1	2	3	4	5
d) Attractiveness	1	2	3	4	5	1	2	3	4	5
e) Dress	1	2	3	4	5	1	2	3	4	5
Status:	1	2	3	4	5	1	2	3	4	5
a) Marital	1	2	3	4	5	1	2	3	4	5
b) Social	1	2	3	4	5	1	2	3	4	5
c) Financial	1	2	3	4	5	1	2	3	4	5
d) Occupational	1	2	3	4	5	1	2	3	4	5
Attitudes:	1	2	3	4	5	1	2	3	4	5
a) to education	1	2	3	4	5	1	2	3	4	5
b) to men	1	2	3	4	5	1	2	3	4	5
c) to women	1	2	3	4	5	1	2	3	4	5
d) to family	1	2	3	4	5	1	2	3	4	5
e) to occupation	1	2	3	4	5	1	2	3	4	5
f) to own appearance	1	2	3	4	5	1	2	3	4	5
g) to social expectations of him/herself	1	2	3	4	5	1	2	3	4	5
Her reasons for going alone to the movie theater:						1	2	3	4	5
His reasons for going alone to the movie theater:	1	2	3	4	5					

Please also rate the following items in terms of how important you think they might be to a better understanding of the events which took place:

The time of day	1	2	3	4	5
The time of year	1	2	3	4	5
The country	1	2	3	4	5
The town	1	2	3	4	5
The type of neighbourhood	1	2	3	4	5
The kind of film programme	1	2	3	4	5

If there is any other important item of information not covered by the above questions, please write it (or them) below:

11. How interested are you in reading the rest of the book:  
(Very interested) 1 2 3 4 5 (Not at all interested)

## Appendix 2

(Questions given to Subjects in Experiment 2)

INSTRUCTIONS

Please read the following. It is an extract from a published book.

When you have read the passage, please answer the questions on the attached sheets. This is not a test of your memory, so feel free to refer to the text while answering the questions. Some of the questions are fairly open-ended, others more specific; do not worry if you have to give the same information twice. There are no right or wrong answers, you are being asked for your ideas and opinions. Make your answers as detailed as you can within the space provided. Please answer the questions as spontaneously as possible.

Before answering the text-related questions, please supply the following information:

Do you recognise this text?	YES	NO		
What is your age?	.....			
Are you male or female?	MALE	FEMALE		
Group (see letter on text)?	A	B	C	D

1) She first moved to another seat before "the soldier moved closer to her". Why did she move?

---

2) Each time she moved "the soldier moved closer to her". Why did he?

---

3) Describe in detail what you think happened after: "The soldier moved once more and sat beside her".

---

4) What were the consequences of your answer to (3) for

a) the soldier;

---

b) her.

5) To what extent do you consider the following to be responsible for what you have described as happening, and why? (Please circle the appropriate number.)

Level of responsibility

Complete

None

---

a) <u>She</u> Why?	4	3	2	1	0
-----------------------	---	---	---	---	---

---

b) <u>He</u> Why?	4	3	2	1	0
----------------------	---	---	---	---	---

---

c) <u>The situation</u> Why?	4	3	2	1	0
---------------------------------	---	---	---	---	---

---

6) What do you think were their reasons for going to the movie theater alone?

a) His?

---

b) Hers?

---

7) Do you think she had an occupation? (If 'yes', what might it be?)

---

8) Please give your ideas about the following details of the characters.

	HIM	HER
a) Age:		
b) Marital Status:		
c) Self-esteem (High, Average or Low):		
d) Attractiveness (Very, Average. Not very):		

9) Please describe what their attitudes might be to the following:

	HIS	HERS
a) The opposite sex:		
b) Their own appearance:		

10) The following adjectives could be used to describe the characters. For each pair of adjectives, circle the one you think best describes that character. If you think neither are relevant, then circle NA.

a) Rate the following in terms of how you feel the characters reacted to the situation described in the text.

HIM			HER		
Relaxed	Tense	NA	Relaxed	Tense	NA
Afraid	Unafraid	NA	Afraid	Unafraid	NA
Composed	Excitable	NA	Composed	Excitable	NA
Assured	Anxious	NA	Assured	Anxious	NA
Rational	Emotional	NA	Rational	Emotional	NA
Controlled	Uncontrolled	NA	Controlled	Uncontrolled	NA

b) Rate the following in terms of what you feel to be the character's more general personality traits.

HIM			HER		
Aloof	Sociable	NA	Aloof	Sociable	NA
Self-confident	Insecure	NA	Self-confident	Insecure	NA
Tough-minded	Sensitive	NA	Tough-minded	Sensitive	NA
Tolerant	Intolerant	NA	Tolerant	Intolerant	NA
Friendly	Hostile	NA	Friendly	Hostile	NA
Self-reliant	Dependent	NA	Self-reliant	Dependent	NA

11) Do you think the characters already knew each other?

YES

NO

12) Specify as precisely as possible when you think the events in the movie theater took place and where the movie theater was.

Date:

Location:

### Appendix 3.

#### Experiment 2a to test the reliability of response categorisation.

Responses to Q1 were assigned by the author to the 7 categories listed in Chapter 11 and described in Appendix 3a. To assess the reliability of the categories used a number of judges were asked to categorise the responses on the basis of these descriptions. The level of agreement was assessed by the extent to which subjects placed responses in the same category as the author. The data were analysed to discriminate 3 possible outcomes:

1. If the categorisation by the author was subjective and if independent judges assigned the responses idiosyncratically, there would be a random distribution of frequencies for each category.
2. If the independent judges agreed with each other but differed from the author in their interpretation of the categories this would result in an uneven distribution of the frequencies (which would be high in some categories but not those predicted by the author).
3. If the judges agreed with each other and with the author, there would be high frequencies of the responses in the categories predicted by the author and low frequencies in all others.

The collected data were tested using an Assymmetric Lambda test of forecasting efficiency that is how well the author's categorisation predicts that used by the judges.

#### Method

##### Judges.

Twelve Open University Psychological Society members who were present at an Annual General Meeting (and who only met on such occasions) were asked to independently categorise the responses. All were graduates or undergraduates who had taken Psychology courses. All had knowledge of content analysis



either as students or as a part of their working life experience.

### Materials.

Each judge was given:

#### 1. The written instructions:

"Your task is to allot each of the responses given to one and only one of the seven categories explained below (see Appendix 3a). To do this you must read the instructions, the text and the complete set of categories before you begin to write. When you have read these then read each response. When you have decided into which category the response best fits write the identifying letter of that category alongside the response in the space provided. If you change an answer please delete your first answer completely and write your new answer alongside.

This is a subjective task so do not compare your responses with those of others until after you have posted your reply.

Please read the text and the descriptions of all the categories before you begin to write, and refer back to them as you work through the responses.

WHEN YOU HAVE FINISHED: Post the completed list to me in the S.A.E. provided."

2. A copy of the text used by subjects in Condition OBE (i.e. the total text: see Appendix 3b).
3. A copy of Question 1 and the responses given to that question by subjects in the main experiment (see Appendix 3c).
4. A list and description of the categories used by the author (see Appendix 3a).
5. A SAE for the return of the categorised responses.

### Procedure.

Each judge was asked to complete the task in his/her own home, at leisure, and to return the categorised responses when completed. Ten completed response sheets were returned over a period of six weeks.

### Method of Analysis and Results.

A 7 x 7 Contingency table (Table A3d ) was produced. The seven columns (A-G) represented the categories used by the judges; the seven rows represent the categories predicted by the author. Thus those 90 responses which the author predicted would fall into category A were examined. The total number of those responses which the judges assigned to category A were placed in Column A, those assigned to Category B were placed in Column B and so on. The same procedure was repeated for for rows B-G.

The contingency table shows that the main diagonal of the table (cells AA-GG) contains a high proportion of the results as would be expected if there was a high level of agreement with the author and between the judges themselves. This was confirmed with an Assymetric Lambda test (Lambda = 0.807, Chi-square = 1218.07, df =3,  $p < .0001$ ). As Lambda is equivalent to the measure called forecasting efficiency (Guilford and Fruchter, 1973, chapter 15) it may be transformed into a correlation coefficient:  $r = 0.98$ .

### Discussion.

As is indicated by the contingency table (Table A3d ) supported by the statistical test, the third outcome has occurred. The correlation co-efficient ( $r = 0.98$ ) is extremely high and demonstrates that there is a close relationship between the assignment of responses to categories by the author and the judges. This is sufficient to conclude that the method is reliable, and that the categorisation is not idiosyncratic.

## Description of Categories

CategoriesA Jenny's feelings and attitudes towards the soldier

This category includes all responses which describe any thoughts, attitudes and feelings which Jenny or 'she' has about the soldier or him.

E.g. She felt that he was annoying her  
She had an aversion towards him  
Because she is angry with the soldier.

B Jenny's feelings and attitudes towards soldiers in general

This category includes all responses which describe any thoughts, attitudes and feelings which Jenny or 'she' has about soldiers or non-specifically "a soldier" (but not the soldier, or him, he).

E.g. She felt she couldn't sit next to a soldier  
Because soldiers made her panic.

C Jenny's feelings and attitudes towards men

This category includes all responses which describe any thoughts, attitudes and feelings which Jenny or 'she' has about men or a man (but not soldiers, the or a soldier or him, he).

E.g. Because she hated all men  
She became aware that she was sitting next to a man.

D Actions by Jenny in relation to the soldier

This category includes all responses which attribute any action which Jenny or 'she' takes to the soldier or 'he'/'him'.

E.g. So that he would notice her  
To catch the soldier's eye  
To avoid him.

E Actions by the soldier

This category includes all responses which attribute Jenny's move to an action by the soldier or 'he'.

E.g. Because he sat next to her  
Because the soldier had "chatted her up" on a previous occasion.

(continued)

## Appendix 3a

Categories (continued)F Cinema Reasons

This category includes all responses given in which Jenny or 'she' acts as a cinema-goer (i.e. not in relation to anyone else) and acts in a way a cinema-goer could be expected to act.

E.g. Her seat was uncomfortable  
She couldn't see the screen

G Others

This category includes any responses which do not belong to any other category mentioned. Please use it for this purpose only.

## Appendix 3b

## THE TEXT

Jenny Fields was arrested in Boston in 1942 for wounding a man in a movie theater. It was shortly after the Japanese had bombed Pearl Harbour and people were being tolerant of soldiers because suddenly everyone was a soldier, but Jenny Fields was quite firm in her intolerance of the behaviour of men in general and soldiers in particular. In the movie theater she had to move three times, but each time the soldier moved closer to her until she was sitting against the musty wall, her view of the newsreel almost blocked by some silly colonnade and she resolved she would not get up and move again. The soldier moved once more and sat beside her.

## Appendix 3c

## The Responses

Q. 1 Why did she move the first time?

Category

- 1. To get away from the soldier.
- 2. She didn't want to sit next to him.
- 3. To get a better view of the movie.
- 4. To avoid the soldier.
- 5. To see the newsreel - her view was blocked.
- 6. She felt the soldier was trying to contact her.
- 7. To show the soldier she wasn't interested in him.
- 8. So that the soldier could move next to her.
- 9. She felt the soldier was bothering her.
- 10. She thought the soldier was "after her"!
- 11. She was afraid of being caught.
- 12. It irritated her having to sit next to a man.
- 13. She was intolerant of men.
- 14. She didn't like men.
- 15. Because the soldier had moved nearer to her.
- 16. She thought that the soldier would behave in an intolerable way.
- 17. She was intolerant of soldiers.
- 18. She felt uncomfortable sitting next to a man.
- 19. Invasion of personal space.
- 20. She felt that if she didn't move he would approach her.
- 21. To enable another couple to sit together.
- 22. Because he did behave intolerably.
- 23. She disliked the soldier.
- 24. To enable a group of people to sit together.
- 25. Perhaps he followed her into the cinema.
- 26. Perhaps she is afraid of him.
- 27. Because he was following her.
- 28. His presence disturbed her.
- 29. To increase her distance from the soldier.

## Appendix 3d

Table A3d

Table of responses assigned to each category:n = 290 (29 responses x 10 subjects)

Categories used by subjects (Dep. variable)		A	B	C	D	E	F	G	
Categories predicted by Author (Independent variable)	A	74	5	4	2	3	0	2	90
	B	0	10	0	0	0	0	0	10
	C	0	0	40	0	0	0	0	40
	D	6	0	0	42	1	0	1	50
	E	3	0	2	1	32	0	2	40
	F	0	0	0	0	0	35	5	40
	G	0	0	3	0	0	0	17	20
		83	15	49	45	36	35	27	290

## Appendix 4: Texts used in Experiment 5

## Text 1 Order A

A: The bees hummed as the sun shone on the garden.

B: The sunshine was dazzling.

Was there ever such a summer? Old people talked about the kinds of summers there were in the days when they were young and all the world was golden.

A: The rain and cold weather were suddenly forgotten.

B: The roses were absolutely magnificent.

People gathered on the grass, and from a distance looked like a collection of flower petals. It was good to be alive.

## Text 1 Order B

A: The rain and cold weather were suddenly forgotten.

B: The roses were absolutely magnificent.

People gathered on the grass, and from a distance looked like a collection of flower petals. It was good to be alive.

A: The bees hummed as the sun shone on the garden.

B: The sunshine was dazzling.

Was there ever such a summer? Old people talked about the kinds of summers there were in the days when they were young and all the world was golden.



## Text 2 Order A

A: As she waited for the waitress to bring her the menu, Mary realised that the time she had allowed for eating her meal was getting shorter and shorter.

B: As she waited for the waitress to bring her the menu, Mary realised that her watch must have stopped half an hour earlier.

She decided that when the meal did come she would not have enough time to eat it and get to her appointment on time.

A: She tried to catch the waitress's eye but was unsuccessful.

B: She looked around to see if her friend was sitting at a nearby table but she could not see her.

So she stood up, picked up her handbag and walked from the restaurant.

## Text 2 Order B

A: She tried to catch the waitress's eye but was unsuccessful.

B: She looked around to see if her friend was sitting at a nearby table but she could not see her.

So she stood up, picked up her handbag and walked from the restaurant.

A: As she waited for the waitress to bring her the menu, Mary realised that the time she had allowed for eating her meal was getting shorter and shorter.

B: As she waited for the waitress to bring her the menu, Mary realised that her watch must have stopped half an hour earlier.

She decided that when the meal did come she would not have enough time to eat it and get to her appointment on time.

## Text 3 Order A

A: It was as she was walking through the trees that Joan stumbled over the body.

B: It was as she was walking through the trees that Joan saw the vultures circling over the body.

It lay flat, its eyes staring up at the sky.

A: She decided that she would never carry a gun again.

B: With superhuman strength she tore the steel jaws apart.

Looking upon the dead animal aroused feelings within her that were a mixture of guilt and dread.

## Text 3 Order B

A: She decided that she would never carry a gun again.

B: With superhuman strength she tore the steel jaws apart.

Looking upon the dead animal aroused feelings within her that were a mixture of guilt and dread.

A: It was as she was walking through the trees that Joan stumbled over the body.

B: It was as she was walking through the trees that Joan saw the vultures circling over the body.

It lay flat, its eyes staring up at the sky.

## Text 4 Order A

A: The windows rattled as the sound exploded above them.

B: The noise outside was deafening.

Were these demonstrations absolutely necessary? The residents on the estate remembered the time, only a few years ago, when this was a calm and peaceful area.

A: Breaking the sound barrier was an achievement for science but not for humanity.

B: Riot shields and water cannons were becoming commonplace.

The people began to get together and discuss what might be done about the problem. One answer was to move from the area.

## Text 4 Order B

A: Breaking the sound barrier was an achievement for science but not for humanity.

B: Riot shields and water cannons were becoming commonplace.

The people began to get together and discuss what might be done about the problem. One answer was to move from the area.

A: The windows rattled as the sound exploded above them.

B: The noise outside was deafening.

Were these demonstrations absolutely necessary? The residents on the estate remembered the time, only a few years ago, when this was a calm and peaceful area.