RECOGNITION MEMORY FOR PICTORIAL EVENTS:

FUSION OR FEATURES

A dissertation presented to the Victoria University of Wellington in partial fulfillment of the requirements for the degree of Doctor of Philosophy

by

Julie Leibrich

1981

CONTENTS

ACKNOWLEDGEMENTS

I wish to thank Dr. K.G. White, my supervisor, for his valued encouragement and assistance during this research; John Marwick and Kim Saffron for the strength of their intellectual contributions during its construction; and Harold Bernhardt for ensuring its completion.

I am also grateful to many friends for their warmth and support, particularly during the last two years, and to the following people for their specific skills: Ross Renner and John Whitmore for their assistance with exploratory analyses and interpretation of the ranking and rating data in Experiment IV; Linda Searle for typing the manuscript; Mark Phillips and Helen White for their help in preparation of the figures; the actors - Doris Eldridge, John Eldridge, Susan Genet, Jill Humphreys, Ron Humphreys, Jan Morrison, Paul Morrison; and the subjects.

This research was supported by a University Grants Committee Postgraduate Scholarship and a Victoria University of Wellington Research Grant (46/77).

ABSTRACT

This research investigated recognition memory for picture stories. Jenkins, Wald and Pittenger (1978) had found that when subjects viewed a slide sequence which depicted an every-day event, in a later recognition memory test they correctly rejected distractors which were inconsistent with the event but falsely accepted consistent distractors. Jenkins interpreted this result as evidence that fusion - the abstraction of visual events - determined memory performance. He argued that subjects compared the test slides to the abstracted event and accepted those which were consistent with the event. A series of experiments examined the possibility that performance was due not to fusion but to confusion with respect to the featural details of the stimulus material. This alternative interpretation argued that consistent slides had more features in common with acquisition slides than did the inconsistent slides and that the variables of semantic consistency and featural similarity had been confounded. The first experiment manipulated acquisition material and found that subjects who saw a disordered acquisition sequence falsely accepted consistent slides. The second experiment manipulated acquisition conditions and found that subjects who were inhibited from fusing the event by being required to perform a non-semantic task during acquisition falsely accepted consistent slides. Neither of these results supported a fusion interpretation since acceptance of consistent slides occurred under conditions where

fusion of the event was not expected. The third experiment manipulated the test conditions and found that acceptance of both consistent and inconsistent slides was less likely with delayed tests although fusion of the event should have led to no change in the likelihood of accepting inconsistent slides. The fourth and fifth experiments re-examined the manipulation of presentation order and demonstrated that subjects were unable to reconstruct the event from a disordered sequence and yet still falsely accepted consistent slides. Each test of the fusion interpretation which had attempted to separate the variables of features and meaning indirectly had indicated that recognition performance was not due to abstraction of the visual event. A final experiment attempted to find explicit evidence for a featural interpretation of the results by directly varying featural similarity of consistent distractor slides to slides from the originally viewed sequence while keeping the degree of semantic consistency constant. Although this experiment failed to support a featural account, the converging evidence from all experiments indicated that recognition memory for picture stories is based to a large extent on the featural properties of the stimulus material. An account of performance solely in terms of visual abstraction is not adequate. Moreover, unless the variables of featural similarity and meaning can be separated directly in the test material, this recognition paradigm is unlikely to provide a means for examining the influence of schemata on recognition memory for picture stories.

LITERATURE REVIEW

Introduction

Jenkins (1974) described fusion as the abstraction of an event from a series of related items. In a set of experiments which examined the nature of visual events, Jenkins, Wald and Pittenger (1978) attempted to demonstrate fusion of pictorial action-sequences by showing that "specific memory for individual pictures (was) outweighed by the abstract or general memory for the event experienced" (p. 139). Using a recognition memory paradigm, Jenkins found that subjects who had seen a slide sequence showing a pictorial event, correctly rejected distractors which were semantically inconsistent with the event but falsely accepted consistent distractors.

The specific thesis of my research is that in Jenkins' experiments, consistent distractors had more visual features in common with acquisition slides than did the inconsistent distractors. The variables of semantic congruence and featural similarity were therefore confounded and the differential response to distractors not necessarily due to fusion. It is more likely that the response pattern reflected a difference in the featural similarity between the distractor types and the acquisition slides. The general thesis is that abstraction cannot be demonstrated in a recognition memory paradigm unless the semantic congruence and featural similarity of the stimulus material is explicity controlled.

Background

An introductory review of the literature traces some of the seminal work which has led to and influenced the development of the idea of fusion. A difficulty in putting Jenkins' research into a theoretical and historical context is that there is no *direct* developmental line of ideas on which fusion can be placed. However, there are numerous indirect relationships between fusion and other lines of research in cognition. The dominant general connection is the assumption that perception is determined by knowledge about the structure of events. This has more generally been called world knowledge (Norman & Rumelhart, 1975a) and a guiding principle in many major studies of cognition has been that perception, and in turn, memory, is determined by world knowledge.

This supposition has taken many forms, undergone various refinements, and generated several applications throughout the literature. It underlies theories of perception and memory which propose cognitive representations such as schemata (Bartlett, 1932; Bobrow & Norman, 1975; Piaget & Inhelder, 1973), propositions (Anderson & Bower, 1974; Kintsch, 1970, 1974, 1977), semantic macro-structures (van Dijk, 1977), frames (Minsky, 1975), and scripts (Shank & Abelson, 1975). It is the hall-mark of accounts of cognitive processing which emphasise dichotomies of global versus featural processing (Navon, 1977), schema versus distinctive feature processing (Pick, 1965), semantic versus structural processing (Craik & Lockhart, 1972), semantic versus episodic encoding (Tulving, 1972), and propositional versus analogue processing (Palmer, 1975a; Pylyshyn, 1975). The idea that world knowledge controls perception is at the root of studies of the influence of context (Biederman, 1972; Palmer, 1975b), organisation (Bower, 1972; Mandler & Johnson, 1976), and meaning (Norman & Rumelhart, 1975a, 1975b), and the processes of hypothesis testing (Bruner, 1957), analysis-by-synthesis (Neisser, 1967), visual abstraction (Posner, 1969; Posner & Keele, 1968, 1970) and linguistic abstraction (Bransford & Franks, 1971, 1972).

The following review presents a selected cross-section of related research. The specific connections which governed selection of these studies were that they adopted a contextualist approach to cognition, they attempted to demonstrate explicitly the influence of world knowledge, the major variable used in the demonstrations was organisation and the stimulus material was predominantly pictorial.

Problems of Definition

Many of the crucial terms and concepts used in this area of cognitive psychology are only loosely defined. They tend to have both general and specific meanings, lay usages and paradigmatic associations. Often, within the literature, the particular way in which a term or concept is used is unclear and at times the meaning of the explanatory phrases appears to be idiosyncratically associated with specific researchers. Where the term in question is uniquely associated with a particular writer (for instance, as 'analysis-by-synthesis' is with Neisser, (1967), little confusion arises because the precise meaning of the phrase has been well specified at conception and in general its definition is used consistently by subsequent writers. However, with the majority of the terms encountered (for example, the term 'schema'), serious complications arise for the reader who is attempting to draw general conclusions from apparently related research. Difficulties encountered with the term schema will be outlined here because it is a central term in this thesis.

Bartlett (1932) foresaw the problems of definition when he introduced Head's (1920) term into his study of remembering. He strongly disliked the term schema as being "at once too definite and too sketchy" (p. 201). However, he considered it was the best term to describe the underlying factor which demonstrated the constructive process in perceptual processing. Bartlett equated 'schema' with 'organised setting' and more elaborately defined the term in this way:

> 'Schema' refers to an active organisation of past reactions, or of past experiences, which must always be supposed to be operating in any well-adapted organic response. That is, whenever there is any order or regularity of behaviour, a particular response is possible only because it is related to other similar responses which have been serially organised, yet which operate, not simply as individual members coming one after another, but as a unitary mass. (Bartlett, 1932, p. 201).

Where the term schema has been used as an explanatory concept in work subsequent to Bartlett's studies, the meaning has most often been different from that proposed by Bartlett, and consequently the term schema refers to a variety of structures and processes. Winograd (1972), for example, described a schema as being, at the simplest level, "a description of a complex object, situation, process, or structure" (p. 72), the emphasis being on complexity rather than on organisation. Whereas Tulving (1972) stressed the temporal aspect of schematic organisation, Allen, Siegel and Rosinski (1978) have emphasised the spatial aspect. Perfetti and Lesgold (1977) described a schema as a "relatively small bundle of information about a concept" (p. 174), again removing the emphasis from organisation and placing it on the notion of information. Neisser's 1976 definition was far more expansive:

> a schema is that portion of the entire perceptual cycle which is internal to the perceiver, modifiable by experience, and somehow specific to what is being perceived. The schema 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).

Although there is a superficial similarity between these definitions, the meaning of the term schema varies in quite fundamental ways with each author.

Some writers consider that distinction between certain terms is unnecessary. White (1974) used the terms 'schema formation' and 'abstraction' interchangeably, and Garner (1962) equated organisation with meaning. Other writers have attempted to subsume terms under one particular concept, for example Friedman (1979) defined theories which invoke schemata and scripts as "frame" theories. It did not seem useful to the present writer either to create new generic terms for these many-faceted concepts or to treat them as though they were indistinguishable. Rather, an attempt has been made to describe terms and concepts as specifically as possible as they relate to the usage described by the individual writers referred to in this thesis.

Context and Features in Perceptual Processing

The studies described in this section have examined the effect of context on the perception of pictorial stimuli. They show that context is a crucial determiner of object perception but also that the specific features of objects are relevant in perceptual processing decisions. Data from these studies appear to be conflicting unless they are accounted for by a model of perceptual processing which incorporates an immediate and continuing interaction between perception of individual features and the total composition of those features.

The main emphasis of this section is on the work of Biederman and his associates who examined the influence of schemata - the formation of a holistic representation of a scene - by studying the effect of context on the perceptual processing of briefly presented visual images (Biederman, 1972; Biederman, Glass & Stacy, 1973; Biederman, Rabinowitz, Glass & Stacy, 1974).

Biederman's research is discussed in detail here for three main reasons. Firstly, contextual studies were a major force behind the polarization of the 'forest versus trees' debate in visual perception, and more recently a catalyst for reconciliatory models. The 'forest versus tree' distinction refers to the precedence of either global structuring of a visual scene or feature-by-feature perception (Navon, 1977). This thesis argues that a similar polarization has occurred in studies of memory where a distinction has been made between the processing of the semantic characteristics of the material and the features by which it is defined.

Secondly Biederman's stimulus material consisted of photographs of real-world scenes, as did that in the present experiments. This choice was determined by the belief that the study of cognitive processes will be advanced by the use of stimulus materials more representative of every-day perception. Biederman (1972) argued that the use of simple stimuli in perceptual research had resulted in models which account for laboratory-situation rather than real-world perception; and that these are not necessarily one and the same. We do not normally see "either a single item surrounded by homogeneous space or an array of unrelated ('random') items" (p. 74). When we look at our world, the objects we see are set in relation to other objects. Rarely do we look at a thing in isolation, for it usually has a setting - a meaningful context. It is the factor of meaningful context which Biederman believed to have been misrepresented or ignored in many models of perception. Such extensions have been valuable in studies of perception, and more recently in environmental studies of memory (Allen, Siegel & Rosinski, 1978; Pezdek & Evans, 1979; Pittenger & Jenkins, 1979).

Thirdly Biederman's examination of the influence of context on visual perception and in turn of the question of holistic and/or featural processing used disorganisation of stimulus material as a way of manipulating its meaning. The variable of organisation proved to be the crucial variable which tested fusion in the present research.

Biederman's studies of the context effect

> A number of theoretical issues present themselves when one attempts to account for the context effect, that is the advantage of coherent over jumbled scenes. One issue concerns identification of the functional units involved in the perception of scenes. Is the functional unit an individual object, or does an observer have access to more global units or schema?

> A second issue is the determination of the locus, in the sequence of processing, where context has its effect. Is it in the initial manner in which objects are physically processed - in the initial segmentation, testing, and weighing of features? Or does the context influence a stage subsequent to that involved in the physical processing, so that physically ambiguous stimuli are interpreted to be consistent with other aspects of the scene already identified? (Biederman, 1972, p. 79)

Biederman (1972) set out to demonstrate that meaningfulness is an essential feature of the very earliest stage of perception. He measured the accuracy of subjects' object recognition performance when the context which surrounded the object was disordered. A disordered context was produced by sectioning black and white photographs of every-day scenes and presenting the sections in a random arrangement. This type of presentation was compared with an intact version of the photograph where the sections had been replaced in their correct positions. Each section contained one well-defined object. In this way the meaningfulness of an object's setting was manipulated by disorganisation independently of the featural complexity of the scene. The subject's task was to decide which one of four objects presented without any context, was shown in a cued section of the photograph. The effect of disorganised presentation was assessed in three ways: presentation time was varied, the complete photograph being shown for either 300, 500, or 700 msec; the target section was indicated either before or after presentation of the slide; subjects were shown the four alternative objects either before or after presentation of the slide.

Overall, subjects recognised objects more accurately when the surrounding context was organised. Duration of presentation did not produce any consistent effect. Recognition was more accurate when the section was cued before presentation of the slide and when the subjects had seen the alternatives before the slide. However, even when subjects knew where to look (cue before) and what to look for (alternatives shown before) disorganisation impeded recognition.

Biederman, Glass and Stacy (1973) extended these findings by measuring the effect of disorganised context on reaction time. Before each scene was presented, subjects were shown a picture of one object and their task was to indicate as quickly as possible whether the object appeared in the subsequent scene.

The interest of this experiment centred on the 'No' responses, of which there were two types. When an object did not appear in the subsequent scene, either it *could have* appeared in such a scene (for instance, the object 'cup', followed by a kitchen scene) or it *could not have* appeared in such a scene (for instance, the object 'cup' being followed by a street scene). There were, therefore, two categories of 'No' responses - 'Possible-No' and 'Impossible-No'. Overall, subjects made 'Possible-No' decisions more slowly than 'Impossible-No' decisions, and disorganisation delayed responses in both categories. However, subjects were most delayed by a disorganised context when the object could have, but did not, appear in the subsequent scene.

Biederman et al. explained these findings in terms of an analysis-by-synthesis model of perception, of which "the idea that a holistic representation can precede - and facilitate - the processing of specific parts is a fundamental tenet" (1973, p. 26). They reasoned that if the search for an object involved a process where specific objects were identified in the course of building up a representation of an entire scene, then context would not affect the time or accuracy of object identification. However, since disorganisation delayed performance, the subject formed a holistic representation first, and only later attended to detail.

This interpretation was given additional support by the difference in reaction times for the two types of 'No' responses. It was proposed that in the 'Impossible-No' situation where reaction times were shorter, the subject viewed the scene and abstracted its general meaning. Since the setting was inappropriate for the target object, the search was terminated. In the 'Possible-No' situation, when the subject discovered that the setting was an appropriate one for the target object, the search continued until the individual objects in the scene were identified and rejected. "Here achieving a schema was insufficient and S would have to engage in detailed feature processing and object identification to determine if the target was in the scene" (1973, p. 26). The effect of jumbling was greatest in the 'Possible-No' condition since not only would the formation of the holistic representation be more difficult but there would be the additional hindrance of identifying an object in a jumbled context.

In the second study, the model of perceptual processing proposed by Biederman et al. seemed to depend on a serial processing in which a holistic representation is formed before detailed analysis is undertaken. In later work, however, Biederman specifically proposed a parallel model of perceptual processing, where a holistic representation and detailed analysis are apprehended simultaneously.

Before describing his later experiments it should be noted that a parallel model is not inconsistent with the findings of the second study since Biederman's interpretation of the differences in reaction times was based upon an 'additive' approach (Sternberg, 1966). Such interpretations must allow for the fact that exact time taken for the individual processes which contribute to the overall latency is not known, and are based on approximate estimates of the relative speed of each operation and on the effects of any interaction between them.

Biederman's work at this point had not explicitly demonstrated a 'forest before trees' type of perceptual processing. However, it *had* demonstrated that the 'forest' is at least present at the very beginning of perceptual processing, since the findings could not readily be interpreted in terms of a feature-by-feature analysis preceding a holistic representation. If this position were held, then one would predict that there would be no overall difference between the 'Possible-No' and 'Impossible-No' conditions, since in both cases the decision would be based on the elimination of each individual object (making a holistic representation redundant). Secondly, therefore, one could not predict an additive effect of disorganisation in the 'Possible-No' condition.

Biederman, Rabinowitz, Glass and Stacy (1974) examined the effect of disorganisation on the identification of a total scene. They were, therefore, concerned not with the effect of a disorganised context on an object but with the effect of disorganisation on the perception of an overall representation of a scene. Subjects were shown a variety of scenes, again either intact or jumbled. Exposure times varied between 20 msec and 300 msec. Before viewing a scene, the subject was shown two verbal descriptions, and was required to select the one which best fitted the subsequent scene. The pairs of labels were similar (for instance, 'shopping plaza' versus 'busy road and stores') or dissimilar (for instance 'lawn in back of house' versus 'kitchen'). Correct responses were predetermined by judges. Disorganisation affected accuracy, particularly at exposure times of 100 and 300 msec, and decisions were more accurate when labels were dissimilar.

In this third study, Biederman appeared to change the direction of his argument. In his interpretation of the two earlier studies he proposed two types of operations - the perception of a holistic representation and the identification of individual objects. He reasoned that the holistic representation is formed first and facilitates the perception of individual objects. However, the interpretation of the results of the third study emphasised the immediacy of the perception of individual objects.

Biederman argued that when labels are dissimilar, subjects may base their decision on the identification of certain 'diagnostic' objects. For instance, if a subject were given the labels 'fire-place' versus 'parking lot in front of stores' he may view the scene and once he had identified a car, decide that the scene is not of a fireplace. If the labels were similar, for instance, in the 'shopping plaza' versus 'busy road and stores' comparison, he may, having identified a shop and some people, not have enough information to arrive at a decision. It is at this point, that "perhaps S also employed a second, more holistic mode of processing which made use of the spatial relations among the objects in the scene to help him distinguish, say, 'shopping plaza' from 'busy road and stores'"(Biederman et al. 1974, p. 599).

Although it appears from this interpretation that Biederman now suggested a 'trees before forest' approach to perceptual processing, he concluded this series of experiments by postulating a parallel model which he considered the best interpretation of all the available data. "It is likely that S simultaneously handles the information from a scene with both modes. That is, individual objects would be identified along with the attainment of the overall scene characterisation" (p. 600). Later experiments by Biederman which examined specific kinds of relationship violations in incongruous scenes confirmed his view that "knowledge of physical relations is not necessarily available prior to knowledge of semantic relations - indeed, they may be one and the same" (1977, p. 88).

Extension of Biederman's Studies

Biederman's research has been extended in several directions, and the results in general have converged towards a dual-processing (parallel) model of visual perception. Antes (1977), expanded the research to include an examination of the informativeness of pictures; Hock and his colleagues (Hock, Gordon & Corcoran, 1976; Hock, Gordon & Whitehurst, 1974; Hock, Romanski, Galie & Williams, 1978) studied the effect of familiarity of pictures, and Palmer (1975b) investigated the degree of contextual meaningfulness.

Antes (1977) combined Biederman's method with a technique developed by Mackworth and Morandi (1967) which measured accuracy of perception as a function of the 'informativeness' of sections of pictures. He found that both recognition and location accuracy depended on the rated informativeness of a probed section, as well as the section's location with respect to the point of focus. Like Biederman, he interpreted his results in terms of a model of picture perception based on two kinds of processing - identification of individual objects and a holistic characterisation of the scene.

A main aim of Hock's work (Hock, Gordon & Corcoran, 1976; Hock, Gordon & Whitehurst, 1974; Hock, Romanski, Galie & Williams, 1978) was to examine the effects of 'familiarity', 'plausibility', and 'belongingness' on recognition memory for organised versus disorganised scenes. Hock attempted to refine Biederman's model of perceptual processing by incorporating these additional factors, and examined whether 'real-world' schemata were generalisable to all scenes or whether they were functional only with familiar arrangements of objects.

Hock hypothesised that it is the familiarity of inter-object relations that determines the influence of schemata. He systematically varied arrangements and selections of objects in scenes to assess subjects' performance on same-different comparison perceptual tasks, word-to-picture matching tasks, and recognition memory tasks. Although the intention was to formulate a general model of perceptual processing, the results were difficult to interpret except in terms of individual differences.

Whilst some subjects were influenced only by plausibility, others were affected only by familiarity and belongingness on the same-different perceptual task. It was inferred that some subjects used internalised rule systems to assemble arrays of objects into organised scenes, whereas others dealt with each object on an individual basis. Similarly, individual differences were found in the word-to-picture and recognition memory tasks. For example, on a task requiring subjects to indicate when there was a match between a word and a simultaneously presented picture (where the word was either a specific description or a general category description, e.g. picture of a collie dog being paired with either 'collie' or 'dog') it was found that some subjects adopted an analytical mode of processing, concentrating on the minimum number of features necessary for identification, whereas other subjects adopted a structural mode of processing, concentrating on all the information in the pictures.

Hock and his colleagues concluded that the fact that there were individual types of responses to such tasks indicated that seemingly conflicting theories of visual processing might be compatible. In particular they suggested that Bruner's (1957) theory which stressed the importance of placing stimuli into categories on the basis of featural information extracted from the stimuli and Neisser's (1967) theory which emphasised a synthetic rather than an analytic approach, might both be valid, depending upon the particular perceptual processing mode used by different individuals.

An attempt was made to resolve the apparent conflict by proposing a two-stage model of perceptual processing in which schemata are formed at the first stage and govern the processing of details at the second stage. They suggested that there are three levels of real-world schemata in scene recognition. At the 'deepest' (most abstract) level, are schemata comprising generalizable rules that specify the relational possibilities of objects. These abstract schemata are the basis for the formation of episodic schemata to represent the particular scenes. A third type - categorical schemata - occupies the level of abstractedness intermediate to deep-level rule systems and surface-level episodic schemata.

Hock's suggestion that the 'trees before forest' and 'forest before trees' models may not be incompatible has merit in that some combination of both seems essential if apparently conflicting data is to be explained. However, the model suggested invokes categories of schemata for which there appears to be little empirical support. It also has a rigidity of formulation which may limit rather then elucidate an understanding of perceptual processing. An alternative way of combining the two main types of perceptual processing models has been suggested by Palmer (1975b, 1975c, 1977).

In discussions of the nature of perceptual representation, Palmer (1975a, 1978) has argued that when different theories of perceptual processing are presented within an adversary context, few advances will be made in understanding perceptual processing since more extreme positions are then chosen to enhance defence. A compromise is needed to develop theories which more adequately account for conflicting data, and models which provide more realistic representations of the perceptual operations involved. Palmer demonstrated the need for such a synthesis in a study which developed the work begun by Biederman.

Palmer (1975b) examined the effect of context on object identification by varying the appropriateness of the context of an object whilst keeping the organisation of the scene intact, rather than by varying its spatial order. Specific objects were paired with scenes which were either appropriate (for instance, a loaf of bread on a kitchen bench) or inappropriate (for instance, a mail box on the bench); the objects were also presented without context. When an object was paired with an inappropriate context, it could be either similar in features to an object for which the context would be appropriate (as in the above example where the mailbox was similar in shape and size to the loaf of bread), or dissimilar in features (for example, a drum on the kitchen bench).

Subjects were more likely to identify a briefly presented picture of an object correctly when it was preceded by an appropriate contextual scene than when it was preceded by a blank slide. They were least likely to identify the object when it was preceded by an inappropriate context. Confidence ratings of responses were a function of the visual similarity between the target object and the object to which it was to be matched, subjects being less confident with visually similar objects. Palmer concluded that the experiment showed that both the context and the sensory characteristics of the presented object determined the response made by the subject. That is to say, that neither a 'forest before trees' nor a 'trees before forest' explanation adequately accounted for the data. Palmer (1975c, 1977) developed an interpretation of these findings in terms of a visual processing model which involved simultaneous use of both holistic and featural encoding. His model, although similar to Biederman's emphasised the interactive nature of the process and was set within the framework of the 'parsing problem':

> How can someone recognise a face until he has first recognised the eyes, nose, mouth, and ears? Then again, how can someone recognise the eyes, nose, mouth, and ears until he knows that they are part of a face? (Palmer, 1975c, p. 295).

In an attempt to resolve the parsing problem he adopted the premise that the fundamental factor in perceptual processing is neither the holistic representation nor the featural characteristics of the stimulus but the interaction which takes place between these two aspects of stimuli:

> The proposal....incorporates a constant interplay between the external sensory information and internal conceptual information. *This interaction is the heart of the perceptual system.*[1] Sensory features 'look for' possible interpretations within the available conceptual schemata, and the possible interpretations 'look for' confirming sensory information among the features being extracted. Generally speaking, the facilitating effect of this type of system is that once a member within a schema has been advanced as a candidate interpretation, the rest of the units within the schema provide 'expectations' about what else should be found and where these things should be located. (Palmer, 1975c, p. 295).

The model proposed by Palmer, therefore, was based on the belief that all types of information (from the most specific to the most general) defining a given stimulus are simultaneously available and are used in the perceptual processing of that stimulus. Any model which embodies a constant interplay between these types of information must accordingly provide for 'communication' between them.

Palmer's model defined perceptual representations as "highly

[1] Italics added

organised data structures containing many embedded levels of detail" (1977, p. 442). Many levels of structure are needed to contain all information about a stimulus which includes its global properties, its component parts and the specific perceptual relationships between them. The component parts in turn have global properties and further component parts.

Such a view of the information contained within a stimulus is conceptualised as a multileveled hierarchical structure of parts and wholes, each of which has a representation both of the holistic properties and the component structure. Each structural unit which contains information about a stimulus is viewed as a level in a hierarchy, where the highest level structural unit represents the figure as a whole with global properties of the most general nature, the next-lower level structural units define less general aspects, and so on. The entire network which defines one stimulus and which is dominated by the highest structural unit is called its schema . "The schema integrates all of the information known about the scene, object, or part into a systematic framework used during perceptual processing" (1977, p. 444). Integration occurs by a testing process between the individual structural units and between the schema and the to-be-identified stimulus, with the aim of achieving the greatest concordance in the most efficient manner.

The approach adopted by Palmer is basically functional in that rather than describing the processing of a given stimulus in terms of a predetermined perceptual system, he describes the system in terms of the stimulus - a view in many ways consistent with

Gibson's (1966) ideas of perception. Specifically he makes the point that only the amount and kind of information needed to discriminate any given stimulus is used in its processing. The system itself, therefore, provides for potential use of all levels of information, and more importantly provides both for the formation of schemata in the first instance and for their later use and modification - a provision without which the parsing problem would undermine any model of processing.

Summary

The data from the perceptual processing studies described in this literature review cannot be accounted for adequately either by 'schema-based' or by 'feature-based' theories. Biederman's studies showed that spatial disorganisation impaired accuracy of object recognition in briefly presented pictorial scenes even when the target object and its focal position were cued. They also showed that subjects took longer to decide that an object was not present in a scene when it could reasonably have been expected to appear in the scene than when its presence would not have made sense. Moreover the decision time was even longer when the scene was disorganised. In the studies which examined verbal labelling of total scenes, accuracy was impaired both by disorganisation of scene presentation and by similarity of label choices. Biederman concluded that the combination of these findings showed that both the overall composition of a scene and its component parts determine perception at the most immediate stage.

These conclusions were confirmed by studies which extended the variables of interest. Antes showed that the informativeness of individual portions of pictures was as important as their context. Hock found that when the variables of familiarity, plausability, and belongingness were introduced into contextual studies, some subjects processed material in a global way but others focused their attention on individual features. Palmer demonstrated that individual features become relevant when the appropriateness of an object's context was varied.

A perceptual processing model such as that proposed by Palmer, specifying a constant interaction between schemata and features, is needed to make sense of the context studies which showed that both the holistic scene and the specifics which define it are used in the formation of its perceptual representation.

The Role of Schemata in Memory

The attempt made by Palmer and others to reconcile conflicting theories of perception provides an important guideline for developing accounts of the processes of visual memory. In the study of memory, a dichotomous approach towards aspects of the retained material has tended to polarize accounts of memory processes into schematic-based and featural-based theories in much the same way as in the study of visual perception. In particular a distinction has been made between the semantic and featural aspects of stimuli. This in turn has led to distinctions between the semantic and featural aspects of memory itself. Such distinctions tend to generate models of memory which emphasise one aspect at the total expense of the other. It will be argued here that this has been the case particularly with the schematic-based theories of memory which have placed such emphasis on the semantic aspects of the stimuli that the role of featural aspects has been overlooked.

An examination of the data on which some major studies of memory are based shows that findings which are used to support notions of schemata and abstraction can be explained empirically in terms of the featural aspects of the stimulus material without any reference to the semantic. It is not argued here that semantic, any more than featural aspects, are unimportant but rather than interpretations of data which do not give sufficient attention to the featural aspects of stimuli may result in misleading explanatory concepts.

The main focus of this section is the work of two research groups which have attempted to demonstrate the effects of schemata on recognition memory. The first group is represented by a series of experiments by Jean Mandler and her colleagues (Mandler & Johnson, 1976; Mandler & Parker, 1976; Mandler & Ritchey, 1977). The main reason for examining Mandler's work in detail, was that like Biederman, she varied the spatial organisation of pictorial scenes as a way of manipulating meaning and attempting to demonstrate the influence of world knowledge. Moreover, Mandler's procedure of examining memory for specific information via systematic transformation of distractors was adopted in the present research as a way of standardizing distractors.

The second group is represented by the linguistic abstraction work of Bransford and Franks (1971, 1972), who claimed not only to have demonstrated that a holistic representation is formed from separate units of information presented over time but also that the holistic representation is remembered and not the individual units of information. Bransford and Franks' studies provide an investigation of verbal material which closely parallels Jenkins' studies of fusion in visual material.

Mandler's studies of spatial organisation

> If one assumes that schemata....are used during the encoding process then the extent to which pictures fit those schemata should influence how well they are remembered and perhaps what types of information are remembered. We are using the concept of schema to refer to an internal structure, developed through experience with the world, which organises incoming information relative to previous experience. (Mandler & Parker, 1976, p. 39).

Mandler argued that the extent to which a picture was organised would affect the extent to which it 'activated' a real-world schema. Therefore, manipulation of the organisation of a picture could provide a means for examining the influence of schemata on memory for pictures. Specifically, Mandler predicted that retention of certain types of information would be enhanced by activation of a real-world schema, whilst others would not be affected. A classification system was developed to describe various types of pictorial information.

The classification was not designed to give an exhaustive analysis of the kinds of information which are contained in a picture, but only to identify some main types of information. The underlying methodological rationale was that examination of memory for specific types of information, may show what aspects of pictorial stimuli determine the accuracy and durability of memory. To examine particular types of information, Mandler systematically varied the distractors used in recognition tests by making transformations of pictures shown during acquisition.

A complete list of transformations (and the types of pictorial information they varied) used during the course of Mandler's experiments, is set out below:

Type of Information	Type of Transformation
1. Inventory: specified the objects contained in a picture.	Type: an object replaced with an object of the same shape and size but of a different conceptual class.
2. Descriptive: specified the figurative detail of objects in the inventory.	Token: a change in the figurative detail of an object.
3. Spatial Location or Relation: specified where objects were in relation to one another.	Rearrangement: an interchange of the position of two objects. Size: an object made larger or smaller. Orientation: a reversal of an object with respect to left-right orientation.
4. Spatial Composition: specified areas of filled versus empty space in the overall composition of the picture.	Move: an object moved slightly. Deletion: an object removed. Addition: an object added.

Mandler and Parker (1976) suggested that real-world schemata should particularly facilitate memory for spatial location (or relation) information. Transformations of spatial location, therefore, should be more easily detected in organised (as opposed to unorganised) pictures. Other types of information should be less dependent on schemata and, therefore, be less affected by manipulations of organisation.

Subjects were shown black and white photographs of complex line drawings. Each drawing, which was presented for 20 seconds, consisted of eight well-defined objects which were either organised to represent a real-world scene or unorganised, merely representing a collection of unrelated objects. In this early experiment, two types of memory test were used and the procedure was not typical of the later experiments. To assess memory for descriptive information, subjects were given an object recognition test. They were shown pictures of each of the eight objects which had appeared in the scene. The distractors varied figurative detail, size, and orientation.[2] When the subjects had selected the eight objects which they thought were present in the original picture, they were asked to place them in a blank frame in the position in which they had originally been shown. This spatial reconstruction test measured memory for spatial location. The retention period was varied by testing subjects either immediately and then again one week later, or only after a one week delay.

[2] In this experiment, where size and orientation of an object were varied in isolation from other objects, spatial relationships were not affected and all three transformations were, therefore, described as varying descriptive information.

The object recognition test showed that memory for figurative detail was better than memory for size or orientation. Descriptive information was more accurately recognised in the immediate test than in either of the delayed tests (which were not significantly different from each other). Organisation did not affect memory for descriptive information but affected memory for spatial location since organised pictures were reconstructed more accurately than unorganised pictures.

It was concluded that spatial relationships between objects are easier to remember when the objects are arranged in an organised (real-world) pattern than when they are unorganised. Memory for descriptive information, however, is not affected by the way in which objects are arranged. On the basis of these findings, the authors suggested that "real-world schemata have less effect on recognition of descriptive detail than on spatial relationships among objects" (Mandler & Parker, 1976, p. 46).

Mandler and Johnson (1976) extended the examination of memory for different types of pictorial information. As in the earlier study, subjects were shown either organised or unorganised collections of objects. Memory for all four types of information (inventory, descriptive, spatial location (or relation), and spatial composition) was assessed by systematically transforming distractors in an immediate recognition test. Five transformations were used - type, token, rearrangement, deletion, and move.

Organisation did not affect memory for inventory or descriptive information. However, spatial location (or relation) information was more accurate for organised pictures whereas memory for spatial composition was more accurate for unorganised, unrelated collections of objects.

The interpretation given of the differential effect of organisation was that spatial location information was the key to a picture's meaningfulness and this type of information was therefore particularly important in organised pictures. Mandler and Johnson suggested that when a subject views an organised scene, attention is focused on spatial location information to the detriment of spatial composition information. However, when a subject views an unorganised scene, which has no overall meaning, spatial location information is less relevant and, therefore, spatial composition information is processed more thoroughly than in an organised scene.

Mandler and Ritchey (1977) examined the nature of memory for spatial information in greater detail. Three more transformations were added to the set of distractors. Spatial composition was now varied not only by deletions and moves but also by the addition of an object. Spatial location (or relation) information was tested by varying an object's size and orientation as well as by rearranging objects. Inventory information was still assessed by the type transformation, and descriptive information by the token transformation. Pictures were presented for ten seconds and organisation was manipulated as in previous experiments. Recognition tests were either immediate, or after delays of one day, one week, or four months.

Particular interest lay in the four month condition. Since research with verbal material had indicated that semantic content rather than structural presentation tends to be well retained over long periods, Mandler and Ritchey predicted that information more relevant to pictorial meaning might be expected to be retained for longer than other types of information. They proposed that spatial location (or relation) and inventory information are more central to a picture's meaning than either spatial composition or descriptive information.

The main prediction, therefore, was that spatial location (or relation) and inventory information would be better recognised than either spatial composition or descriptive information after a four month delay and also be more affected by the manipulation of organisation. If an unorganised picture is meaningless and does not activate a real-world schema, then spatial location (or relation) and inventory information should not be better remembered than spatial composition and descriptive information, given long retention intervals where the acquisition stimuli had been unorganised.

The major finding was a significant three-way interaction between type of transformation, retention period, and organisation. Recognition of spatial location (or relation) information declined only marginally over four months given organised pictures, whereas there was a highly significant decline over the same period for unorganised pictures. Where transformations varied addition or deletion (spatial composition information) and type (inventory information), there was a significant decline in performance over four months with organised pictures. However, there was no decline with unorganised pictures. Moreover, in the case of the spatial composition transformations, immediate testing performance was better with unorganised than organised pictures. Mandler and Ritchey accounted for this effect by arguing that addition and deletion transformations varied not only spatial composition but also inventory information:

> At immediate test, the spatial composition information is available, but as it drops out, these changes can be detected on the basis of the inventory of objects. Since inventory information is better retained in organised pictures, recognition of addition and deletion in unorganised pictures suffers in the long run. (Mandler & Ritchey, 1977, p.395).

Mandler and Ritchey concluded that since disorganisation impaired recognition of inventory and spatial location (or relation) information, these types of information are contained in the scene schema activated by real-world organisation. However "the scene schema does not include descriptive information about the objects or the overall spatial composition of the scene" (p. 395), since these types of information are processed independently of organisation.

Summary and criticism

A criticism of Mandler's experiments is that they involve an ignoratio elenchi. They first proposed that the activation of real-world schemata facilitates the retention of certain types of pictorial information. They next assumed that organised as opposed to unorganised scenes activate real-world schemata. They then predicted that certain kinds of information would be better remembered in organised pictures. On observing this to be the case, they concluded that the activation of real-world schemata was responsible for superior retention. It is obviously unnecessary to postulate real-world schemata at all, far less maintain that the data support their existence. The point demonstrated was that spatial organisation is a critical variable in memory for pictures and affects the featural aspects of pictorial material which are related to spatial relationships.

The empirical findings of Mandler's experiments were that the variable of spatial organisation did not affect memory for the figurative detail of individual objects in a scene, nor their unique movement, addition, or deletion. Although it was argued that inventory information was affected by organisation, this was not a consistent finding since neither Mandler and Johnson (1976) nor Mandler and Ritchey (1977) obtained an effect at immediate testing. The variable of spatial organisation affected precisely what it was manipulating - the spatial relationships between a group of individual objects. Moreover this did show that meaningful contextual arrangements of objects were easier to remember than meaningless arrangements under some circumstances. However, memory for the specific components of the scenes was also relevant to overall recognition performance. The experiment did not demonstrate the existence of real-world schemata.

Bransford and Franks' studies of linguistic abstraction

Bransford and Franks (1971), aimed "to lend some precision to Bartlett's (1932) notions of abstract schemas as *what is learned*" (p. 332). Specifically, they attempted to demonstrate the phenomenon of 'linguistic abstraction' - the spontaneous organisation and integration of information from a series of discrete but related sentences into a coherent holistic story.

The experimental technique designed to demonstrate abstraction involved the presentation of a story by its component parts over a series of trials. Subjects were presented with four stories each of which consisted of four basic ideas (or propositions). Trials consisted of the presentation of sentences which could contain one, two, or three of the four propositions contained in the story. For example, the celebrated 'ants and jelly' story consisted of four propositions:

The ants were in the kitchen.

The jelly was on the table.

The jelly was sweet.

The ants ate the jelly.

The four sentences above were referred to as 'ones" since they each contained only one proposition. 'Twos' were made by combining ones , for instance - the ants in the kitchen ate the jelly. Similarly, 'threes' could be formed - the ants ate the sweet jelly which was on the table. In acquisition, subjects were shown selections of ones, twos, and threes for each story in the group but were never shown any of the complete four-propositional sentences - for example, in the ants story the complete sentence was - the ants in the kitchen ate the sweet jelly which was on the table.

Learning was incidental, with subjects being asked elliptical questions about sentences as they were presented (e.g. Did what? Where?). In the subsequent recognition test, distractors consisted of any ones , twos and threes which had not been shown during acquisition, the fours which contained the entire story, and some 'non-cases' in which propositions from different stories had been combined. Subjects were asked to distinguish between old and new sentences and to rate their confidence in the correctness of their decision on a scale of 1 to 5 where 1 indicated low confidence and 5 indicated high confidence. In their data analysis of recognition memory performance, Bransford and Franks assigned a plus to any "Yes, seen before" responses and a minus to any "No, not seen before" responses, thereby creating a scale from -5 to +5, where +5 meant that the subject was highly confident that he had seen the test stimulus before.

There were two main findings. Subjects did not discriminate between old and new sentences with the exception that they detected the non-cases as new and also tended to discriminate between new and old ones.[3] The second finding which is known as 'the complexity effect', was that the more complex a sentence was (i.e. the greater the number of propositions it contained), the more certain were subjects that they had seen it before. Subjects were virtually always certain that they had seen the fours, even though these sentences had not been shown during acquisition.

Bransford and Franks argued that during processing of the acquisition sentences, subjects formed a unified representation of the meaning of each story and that during recognition they compared the test sentences against this 'holistic memory trace. They proposed a schema-based model of memory where individuals construct holistic concepts on the basis of separate units of information. Recognition of an item seen during acquisition is determined by the extent to which it deviates from the 'holistic concept. False acceptance of new items is monotonically related

[3] It should be noted that in the first two of the three experiments in this study the discrimination performance is not particularly meaningful given that the proportion of olds to news was 88 to 12 in the first experiment and 100 to 0 in the second and there was no attempt to measure response bias. In the third experiment there was an equal proportion of olds to news.

to the degree to which new items are consistent with the holistic concept. In other words, it is the abstracted idea or event which is stored in memory rather than the specifics which have defined it.

The possibility that the syntactic structure of the items play a role in remembering was not discounted by Bransford and Franks. They suggested that subjects retain "information about the general style in which the semantic information was originally expressed" (1971, p. 349). However, the dominant interpretation of their data was that a schema is formed via temporal integration when a series of related items are presented, and that this schema is the main determiner of memory performance.

These studies of verbal abstraction were interpreted by Jenkins directly in terms of fusion:

> The subjects have used the various strands repeatedly available in the texture of the experiment to construct four events that are *completely* described by the four long complex sentences. Once the *fusion* of strands into events (has) occurred....the subject cannot perform an analysis to recover the exact pattern of input that furnished support for the construction that he made. (Jenkins, 1974, p. 790).

Bransford and Franks' seminal study has been influential in the study of abstraction. In a 1972 review of studies of linguistic abstraction which has used their procedure, Bransford and Franks reported replications of the basic response pattern they had observed under conditions where acquisition instructions were varied, both recall and recognition test procedures were used, and where the types of stimuli included both abstract and concrete sentences, visual geometric patterns, and connected discourse. There seemed to be strong support for an abstraction interpretation.

Extensions and criticism of Bransford and Franks' paradigm

Although Bransford and Franks appeared to provide a direct method for examining linguistic abstraction, from which extensive study of the phenomenon could develop, alternative models have been proposed to account for these data. Singer and Rosenberg (1973) attempted to construct a model of the recognition process by proposing that with related material from which a central concept has been abstracted, individual items are assigned weights proportional to their centrality to the total concept. They suggested it is the relative weighting that determines recognition performance.

Reitman and Bower (1973) proposed what Flagg (1976) later termed the 'tally model'. This model assumed that as each item is presented during acquisition, subjects store the entire item but the strength of the memory tag given to any one item is determined by the number of components of which it is composed. Tags are stored in memory not only for the total item, but for all its component sub-items. During the recognition test, subjects tally the number of tags associated with any one item and the greater the number of tags an item has, the more likely a subject is to judge that he has seen it before. With such a system, complex sentences necessarily accumulate larger tallies and are, therefore, falsely accepted.

Flagg's (1976) test of this model argued that if all 'ones from a complex sentence were presented (and no twos , threes , or fours), then in the recognition test the abstraction and tally models would predict different results. According to Bransford and Franks' formulation, subjects would think that they had seen the four before because they would have spontaneously integrated the propositions presented in the individual ones and compared the four to the integrated schema. In contrast, the tally model would predict a decrement in recognition confidence as the items became more complex, since the acquisition items had each been assigned only one tag, and overall there would be a smaller number of tags associated with the selected sentence than with other fours . The results supported the tally model.

An 'attribute frequency model' proposed by Neumann (1974) argued that recognition performance is a function of the frequency with which each item and each combination of items has occurred during learning. He tested this explanation using geometric stimuli and found that when the frequency with which individual items and combinations of items presented during acquisition were varied systematically, those items having a high frequency of presentation tended to receive the highest positive confidence rating. He argued that a frequency rather than abstraction interpretation of the data was more plausible.

Although Bransford and Franks' studies have stimulated a development of abstraction-type theoretical formulations, consideration of their merits is not an issue if the data on which they are based were actually an artefact of the procedure used. That the results of Bransford and Franks' experiments are unlikely to be due to the underlying semantic content of the series of items is indicated by Katz and Gruenewald's (1974) and Reitman and Bower's (1973) findings that the complexity effect occurs with meaningless stimuli. More specifically, and most pertinent to the present thesis, are the arguments put forward by White (1974) and James and Hillenger (1977) that the method designed by Bransford and Franks confounds abstraction of the schema with confusion with respect to the surface structure of the individual items.

In a learning paradigm, where remembering specific features is not required and where the individual items are extremely similar, it is likely that subjects cannot distinguish between olds and news because they are confused with respect to the features of the material, not because new items are semantically consistent with the abstracted schema. The monotonic relationship between complexity and confidence in these experiments may therefore reflect the *degree* of confusion rather than an underlying process of abstraction.

White (1974) argued that Bransford and Franks' procedure of using a semantic task in the incidental learning procedure minimised retention of the individual items of which the complete story was comprised. He, therefore, compared two types of learning procedures - an intentional condition where subjects were instructed to listen to each of the sentences carefully in order to remember them for a later recognition test, and an incidental condition where subjects were asked to rate each acquisition sentence for its imagery value. White found that whilst the performance pattern of the incidental group closely resembled the results of the Bransford and Franks' studies, this was not the case with the intentional group where subjects tended to discriminate between olds and news and where the complexity effect was not present since subjects were no more confident in their false acceptance of fours than they were of twos.

A more fundamental criticism made by White was that the combined measure of recognition accuracy and confidence rating used by Bransford and Franks confounded two different types of responses. He, therefore, performed an analysis of his data where recognition performance and confidence in performance were treated as separate measures. Such a separation showed even more clearly that subjects in the intentional condition responded differently from subjects in the incidental condition.

The recognition performance data provided strong evidence that subjects in the incidental condition did not discriminate between olds and news (with the exception of ones) whereas in the intentional condition subjects were able to make this discrimination. The confidence rating data showed no difference either between the intentional and incidental condition or across the four sentence types. That is to say when the confidence measure was considered separately from the recognition performance, subjects were no more confident about their responses to four-propositional sentences than to any other - the complexity effect was not present. White suggested that when specific features of sentences are emphasised in a recognition memory task there is a weaker tendency to abstract ideas. When there is abstraction there is little discrimination. Abstracting ideas is thus confused with failure to discriminate.

James and Hillenger (1977) similarly have argued that:

> Although related sentences may be integrated, Bransford and Franks' results can be predicted on the basis of interference arising from the differing degrees of confusability among sentences of differing complexity. If that can be done, any reference to integration is unnecessary. (James & Hillenger, 1977, p. 712).

They tested Bransford and Franks' procedure by reducing the similarity and/or list structure of stimuli and found that under these conditions, the ability to remember specific sentences increased and the effect of monotonicity between confidence ratings and sentence complexity decreased.

Using an analysis similar to that suggested by White, they also found that when confidence ratings were assessed independently of recognition response, the monotonic relationship was no longer present. James and Hillenger argued convincingly that a confusion interpretation not only accounts for Bransford and Franks' results, but also accounts for them with far greater economy than does any other interpretation. They concluded that:

> Although the particular semantic integration hypothesis proposed by Bransford and Franks is false, the potential role on integration cannot be denied. However, any integration occurring in this paradigm must be confounded with the confusion factor....Our purpose is not to deny the importance of integration, but to propose that the search for a paradigm relevant to its study must continue. (James & Hillenger, 1977, p. 720).

Bransford and Franks' method of examining verbal memory has been extended to memory for pictorial material (Cortis Park & Whitten II, 1977; Pezdek, 1978). However, neither of these studies have attempted to clarify the possible confounding of features and meaning or of confidence rating and recognition performance.

Pezdek (1978) designed cartoon drawings which could contain either one, two, three or four components of a total cartoon picture. Subjects were shown cartoons during acquisition, along the lines defined by Bransford and Franks, but were given intentional learning instructions, being asked to remember each individual item. An intentional rather than incidental learning paradigm was adopted so that it might later be argued that any demonstration of abstraction could be said to be attributable to a spontaneous integration of the component parts.

Although Pezdek found that when subjects falsely accepted new items, the confidence in their decision was monotonically related to the complexity of the distractor, old items were responded to significantly more accurately than were new items. The pattern of recognition performance found in Bransford and Franks' studies, therefore, did not appear to extend to pictorial material. In a second experiment, Pezdek incorporated a modification which provided a test of Flagg's tally model. The information for a selected 'key' cartoon picture was presented solely as ones during acquisition, whereas the other three cartoon pictures were presented as combinations of ones , twos , or threes . Recognition confidence increased with complexity of the test item only with cartoon series which had been presented in the standard manner and in fact decreased with increasing complexity of the key cartoon items. Pezdek, therefore, supported the tally interpretation of Bransford and Franks' studies.

Pezdek's experiments are equally well accounted for by a 'confusion' hypothesis. Like White (1974), Pezdek obtained a high degree of recognition accuracy when subjects were given an intentional learning procedure. This may well have been enhanced by the fact that the cartoon line drawings were simple black and white drawings and also by the spatial aspects of pictorial stimuli (with the component parts of each cartoon being in separate portions of the overall frame). These aspects of the stimuli are likely to have made the items less similar to each other in terms of features than were the kind of sentences used in the standard Bransford and Franks experiments. Secondly, and more importantly, Pezdek, like Bransford and Franks, confounded confidence rating with recognition accuracy in her data analysis.

Cortis Park and Whitten II (1977) also attempted to replicate Bransford and Franks' studies. Unlike Pezdek, they used photographs of real-life events and directly compared the performance of subjects who saw pictorial stories with the performance of subjects who saw sentence equivalents and who were asked either to rehearse the sentences or to construct them into images. They found that whilst both types of sentence presentation produced results comparable with those obtained by Bransford and Franks (when the combined recognition and confidence measure was used), the group who saw pictorial material discriminated between olds and news and did not exhibit the complexity effect. These results were found with both an intentional and an incidental learning procedure.

Although there was no explicit attempt either to separate features from meaning, or confidence rating from recognition response, the authors suggested that the salience of the featural properties of pictures compared to sentences may have accounted for their results:

> Previous abstraction research has used stimuli that were either more abstract (dot patterns, sentences, digit or letter strings) or more impoverished visually (geometric forms) than real-world scenes. When one considers how markedly different real-world scenes are from these other stimuli, particularly on the dimensions of specificity and richness of detail, it seems less surprising that the scenes are not recognised in a manner analogous to these other stimuli. (Cortis Park & Whitten II, 1977, pp. 536-7).

The 'Pictorial Superiority' Effect

Early studies suggested that memory for pictorial material was virtually limitless. Shepard (1967) showed subjects over 600 pictures. In a subsequent forced-choice recognition test, 99.7% were correctly recognised. A week later recognition performance had dropped to 87% correct choice and only after 120 days did recognition memory for the pictures fall to chance level. Later studies produced even more impressive results. Standing (1973) presented subjects with pictures over a period of five days and estimated a total retention of 6,600 items from a pool of 10,000. Standing, Conezio and Haber (1970) showed subjects 2,560 photographs, each appearing for five to ten seconds. In this situation, recognition performance exceeded 90% accuracy in a paired forced-choice recognition procedure even with a mean retention interval of 1.5 days. High levels of accuracy have been demonstrated both with recognition and recall procedures (Bousefield, Esterson & Whitmarsh, 1957) and retention seems to be excellent over long periods of time (Nickerson, 1968).

Memory for pictorial material appears to be superior to memory for corresponding verbal descriptions, both with simple material (Davies, Milne & Glennie, 1973; Jenkins, Neale & Deno, 1967; Paivio & Csapo, 1969, 1973) and with complex material (Dallett & Wilcox, 1968; Nelson, Metzler & Reed, 1974). Pictures result in better performance than words in a serial position recall task involving both a temporal and spatial component (Allik & Siegel, 1974). In a 1969 study of memory for visual versus verbal stimuli, Haber's finding of 'pictorial superiority' led him to suggest that recognition memory for pictures may be unlimited.

A main question emerged from the early studies of memory for visual material: why is memory for pictures superior to memory for verbal material? This question has generated many models of memory which, because of the apparent differences in performance as a function of the type of stimulus material, assume separate pictorial and verbal memory systems. One of the most influential of these 'dual-code' models has been that proposed by Paivio (1971, 1975) who has suggested that there are two functionally distinct but partially interconnected memory systems, one of which processes and stores verbal (or propositional) information, and the other of which encodes and stores visual (or imaginal) information. Paivio has argued that the pictorial superiority effect is due to pictorial material being more likely than verbal material to be encoded and stored in both systems and, therefore, more effectively remembered. Other models of memory have particularly emphasised the imagery aspect of pictorial material. One of the chief proponents of imagery accounts of pictorial memory is Kosslyn (1975). The imagery debate has occupied a large area of cognitive psychology in recent years, Pylyshyn (1973, 1975) being the main opponent of imagery accounts. For a review of the major points of contention in this area see Kosslyn (1978).

Although some research has suggested that there is an independence of verbal and visual codes of the same stimuli (Bahrick & Bahrick, 1971; Bahrick & Boucher, 1968), the majority of studies which have examined an interaction between verbal and visual information have found an interdependence which suggests that there are not separate codes.

Memory for ambiguous line drawings is determined by the labels attached to those drawings (Carmichael, Hogan & Walter, 1932). If a subject verbally describes a scene at the same time as viewing it, subsequent recognition performance is boosted relative to a control condition of normal viewing (Freund, 1972; Kurtz & Hovlanc, 1953). If attempts are made to prevent verbal encoding during viewing (by forcing a subject to count backwards) subsequent memory performance is reduced although not to chance level (Freund, 1972; Loftus, 1972). Both with photographs of natural scenes

(Freund, 1972) and with nonsense forms (Daniel, 1972), subjects who generated verbal labels for the pictures at the time of presentation performed more accurately on a subsequent recognition test than subjects who did not label the pictures. Pezdek and Evans (1979) have argued that verbal labels only facilitate visual memory performance when they provide effective discrimination between otherwise complex multi-item stimuli. When verbal labels do not facilitate discrimination they may even impair recognition accuracy.

The most compelling evidence against a separate code hypothesis comes from experiments which have attempted to modify memory for pictorial material by the introduction of subsequent verbal material. In a series of studies, E.F. Loftus and her colleagues, demonstrated that the introduction of misleading verbal information following the presentation of a visual scene resulted in poorer performance on a forced-choice recognition test than did the introduction of consistent or irrelevant information (Loftus, 1975; Loftus, 1977; Loftus, Miller & Burns, 1978; Gentner & Loftus, 1979). Pezdek (1977) and Rosenberg and Simon (1977) have similarly demonstrated 'cross-modality integration'.

The cross-modality studies are important because they offer a more general kind of evidence that dichotomous approaches may be quite misleading. Dichotomous models of memory have been generated primarily by studies which have examined stimulus material in a dichotomous manner. It seems likely, in the light of the cross-modality studies that any dichotomy which does exist between memory for verbal and memory for pictorial material is not a function of separate memory systems, but rather of the stimulus material itself.

Some theorists who have been concerned with accounting for the 'pictorial superiority' effect have concentrated on the properties of pictorial material rather than the processes involved in its encoding. This emphasis has taken them away from the probably spurious verbal-visual dichotomy and provided a context for describing memory for pictorial material without its necessarily being viewed as a distinct process from memory for verbal material.

Nelson and his colleagues (Nelson & Reed, 1976; Nelson, Reed & McEvoy, 1977; Nelson, Reed & Walling, 1976) have attributed the pictorial superiority effect to the fact that pictorial material is more elaborate than verbal material and, therefore, more elaborate sensory codes are likely to be formed during the processing of pictorial materials. Several theorists have argued that pictures are remembered better than words because they are more likely to have a unique meaning (or greater specificity) than words (Durso & Johnson, 1979; Potter, Valian & Faulconer, 1977; Snodgrass & Vanderwart, 1980).

Jenkins, Wald and Pittenger (1978) proposed an account of pictorial superiority in terms of the semantic rather than featural aspects of the stimulus material. They argued that in traditional studies of memory for pictures the items are not related to each other and each item, therefore, defines a separate event. In acquisition the subject perceives a number of distinct events and during the recognition test discrimination between old and new items is made on the basis of events perceived or not perceived rather than on the basis of the distinctive features of the pictures.

The present writer argues that the impressive recognition accuracy found in studies of memory for pictures is related to the particular selection of acquisition stimuli and distractors. High recognition scores may be due to the fact that the material seen by subjects during acquisition has tended to be heterogeneous and that distractors have been highly dissimilar with respect to features.

Many of the early experiments involved a selection of stimulus material based on an intuitive notion of 'memorability'. Shepard (1967), for instance, chose sets of pictures that were "individually of high salience and memorability" and "collectively of low similarity and confusability" (p. 157). Goldstein and Chance (1970), on the other hand, presented sets of material which were not chosen for their individual salience or collective similarity. Three categories of pictures were shown to subjects - faces, inkblots and snow-crystals. Accuracy on a later recognition test was 72% for faces, 51% for inkblots, and only 36% for snow-crystals. The greater the homogeneity of the acquisition material, the poorer was memory performance. Weaver and Stanny(1978), using colour slides of outdoor scenes, found that the recognition of a single picture increased with a reduction in the similarity of the acquisition stimuli. When subjects were asked to indicate "what percentage of their responses was based on details in the pictures as opposed to a holistic analysis of the stimulus...subjects were quite consistent in judging that *details* were of greatest value" (p. 61).

When variability within the acquisition set is reduced and when distractors have many features in common with acquisition slides, recognition memory for pictures is less impressive. This thesis, therefore, argues that the featural characteristics of pictorial material must be taken into account in any interpretation of recognition performance for pictures.

Jenkins' Studies of Fusion

Jenkins, Wald and Pittenger (1978) examined recognition memory for pictorial slides sequences. They adopted "an ecological approach to the problem" in which they regarded the "event as primary" (p. 130).

> We see events as natural wholes that are, so to speak, perceived *through* the slides, rather than built up from the slides. The slides are windows through which the specifications of the event are glimpsed; they are not Tinker Toys that are used to construct some kind of event-like edifice. (Jenkins et al., 1978, p. 158).

They posed the following question:

> If a subject saw an appropriately ordered sequence of pictures that was sufficient to give him all of the necessary information for an event, would he give us evidence that he had experienced that event in its entirety? Would he, for example, falsely recognise pictures of the event that he had not seen before? Would he be able to reject pictures that were highly similar to the pictures he had seen but which violated some invariant of the event or some detail of the observation? (Jenkins et al., 1978, p. 137).

Three slide sequences of simple everyday events were used to examine this question. Two sequences were designed to depict clearly defined events in which there was a coherent progressive action-sequence. An eighteen-slide sequence showed a woman making a cup of tea and a ten-slide sequence showed a teenage girl answering the telephone. A third sequence in which ten slides showed people at a party was thought not to have a clear story line. Subjects viewed the acquisition sequences twice and were then given a standard recognition memory test in which they were asked to discriminate between original slides and distractors.

There were two kinds of distractors. The first, called 'Belongings', were part of the event depicted during acquisition. The complete 'Telephone' sequence, for example, had consisted of fourteen slides of which only ten were shown in acquisition, since every third slide in the complete sequence had been withdrawn before presentation. The four slides which had not been shown in acquisition were consistent with event and used as the Belonging distractors. A second type of distractor, called 'Controls' were inconsistent with the event depicted during acquisition. They included either some change in the appearance of an actor in the story, or the overall perspective or composition of the scene was altered. The complete recognition test material therefore consisted of Originals, Belongings, and Controls. For the Tea sequence the test series consisted of 8 Originals, 8 Belongings, and 8 Controls. For the Telephone sequence there were four slides of each test type, and the Party test material consisted of 4 Originals, 5 Belongings, and 3 Controls.

It should be noted that this experiment is a very strong test of our hypothesis. Our fundamental assumption is that if the pictures show an event taking place over time, the subjects will apprehend the event. For our first test to work successfully a further assertion is necessary, namely, that having apprehended the event, the subject will be unable to reject a picture that fits the specifications of the event he has experienced. This is in spite of the fact that the two presentations in original learning ought to ensure specific memory of the slides. Thus we must argue that specific memory for individual pictures will be outweighed by the abstract or

> general memory for the event experienced. At the same time we shall argue that some aspects of memory will be enhanced; specifically, that any picture that violates the constraints or invariants of the experienced event will be detected as new, no matter how much it resembles the original pictures in terms of its elements. (Jenkins et al., 1978, p. 139).

The recognition performance was as follows: with the Tea series, 80% of Originals were correctly accepted, 50% of Belongings were falsely accepted and 10% of Controls were falsely accepted. In the Telephone series there was 94% correct acceptance of Originals, 42% false acceptance of Belongings and only 3% false acceptance of Controls. With the Party series, 83% of Originals were correctly accepted and Belongings and Controls were falsely accepted less than 10% of the time.

The recognition performance patterns depended on which sequence was tested. Where the sequences depicted a clearly defined event, subjects correctly accepted Originals but responded to the two types of distractors in different ways. They tended to correctly reject Controls but falsely accept Belongings. However, when the sequence did not depict a clearly defined event, as in the case of the Party sequence , subjects not only correctly accepted the Originals, but also correctly rejected both types of distractor.

Jenkins argued that the first type of response pattern (in the Tea and Telephone sequences) was due to *fusion*. That is, subjects had learned the relationship between the individual slides in the sequence and had perceived them as a coherent whole event which was remembered better than the individual items of which it consisted. He reasoned that in the case of the Party sequence, where Belongings were correctly rejected, subjects had not "apprehended the event" in the first place and, therefore, had subsequently easily distinguished Originals from Belongings, since Belongings did not actually belong to any defined event.

The robustness of the 'fusion effect' was demonstrated in a classroom setting. Students were shown the Tea sequence twice at the beginning of class and then 45 minutes later were given a recognition test. This procedure was repeated on the next two days. Subjects' performance remained the same over the course of the three days. Belonging slides which were falsely accepted on the first day were also falsely accepted on the following two days, despite the presentation material being shown twice each day. When, in another similar experiment, the delay factor was removed and testing immediately followed presentation, the 'fusion effect' persisted with the slight modification that Belongings which the majority of students incorrectly judged to be Originals, became even more widely accepted over the three days, and Belongings which had been correctly rejected became more widely detected. Jenkins interpreted these findings as evidence that an event became better specified with repeated exposures.

Criticism

It is argued here that the poor discrimination between Originals and Belongings compared to Originals and Controls in the experiments described above, which Jenkins interpreted as evidence for fusion, could also be accounted for in terms of confusion with respect to featural detail.

Although it was claimed that both types of distractors were highly similar to Originals it is also the case that slides which are withdrawn from a sequential related set (Belongings) are likely to have more features in common with the remaining slides (Originals) than the Controls which not only do not come from the progressive sequence, but which were made with some specific featural alteration such as object perspective, or composition change. It is argued, therefore, that not only are Belongings semantically consistent with Originals they are also featurally more similar than the inconsistent Controls. The variables of semantic consistency and featural similarity are thus confounded and it is possible that Belongings are falsely accepted not because they are consistent with the perceived event, but because they are more similar to Originals than are Controls.

When the three sequences used are considered in terms of this interpretation, the data may easily be accounted for by reference to the features of individual slides. With the Party sequence, where Belongings were correctly rejected, they were not drawn from a progressive sequence and would not necessarily have more features in common with Originals than did the Controls. In contrast, the

Tea and Telephone sequences were not only progressive but also, these two sequences were photographed from a fixed stationary point, which defined a constant frame for Originals and Belongings (but not for Controls). The Party sequence was photographed from different angles for all three types of slides.

A study by Baggett (1975) indicated that confusion with respect to featural similarity is an important factor in studies of memory for sequential pictures. Subjects were shown four-frame cartoon sequences and later asked to discriminate between originals and distractors which were either consistent or inconsistent with the story. Although, when asked, subjects could identify distractors which were consistent, they did not falsely accept them in a recognition test. In fact there were no false recognitions at all. Baggett suggested that a possible reason for the lack of false recognitions was that with the stimulus set used, each frame was considerably different from any other with respect to features and that such gross featural differences could have enhanced discrimination.

Jenkins attempted to demonstrate that a featural-similarity account of the findings was not appropriate. He argued that if a set of Originals, which depicted an event when shown in a progressive sequence, could be presented in such a way that the event was "destroyed", then subjects would *not* falsely accept Belongings if false acceptance were indeed due to fusion. They would, however, falsely accept Belongings if false acceptance were due to featural similarity.

Since the action-sequences depicted an event which took place over time, Jenkins attempted to destroy the event by temporal disorganisation. The slides in each set were now shown randomly with respect to chronological order.

The results of this crucial test of fusion were not clear cut. Disordered presentation affected recognition performance differently depending on which sequence was examined. There was no effect of disordering the Party sequence, where subjects correctly rejected Belongings. Jenkins observed that this result was to be expected since there was no clearly defined event with the Party sequence even when they were shown in a non-random order. The important result centred on recognition performance given the two event sequences. This is where disordered presentation should have resulted in correct rejection of Belongings if their false acceptance under ordered conditions had been due to fusion. In the case of the Telephone sequence, disordered presentation resulted in correct rejection of both Belongings and Controls. However, in the case of the Tea sequence, even though the slides were presented in a random order during acquisition, subjects' performance did not differ from the ordered condition - discrimination between Originals and Belongings was much worse than between Originals and Controls.

Jenkins argued that the effect of disordering on recognition performance to the Tea sequence supported fusion and rejected an interpretation of his data in terms of the featural properties of the individual slides:

> Obviously similarity between individual pictures could not be the source of the false positives that were originally observed for the Belonging slides in this sequence. If the false positives had been due simply to picture similarity, there is no reason for the order of presentation to make any difference at all. (Jenkins et al., 1978, p. 143).

The contrary finding - that when the Tea sequence was disordered there was no change in the high false acceptance of Belongings - was attributed to the possibility that subjects had reordered the sequence and therefore been able to fuse the event. In other words, Jenkins argued that the tea sequence was so well integrated as an event that disordering did not destroy it.

Jenkins reported further studies of memory for visual events with different stimulus material, from which he found support for fusion. However, they are not discussed in any detail here since those that tested the crucial variable of disordering did not find the predicted effect (Kraft & Jenkins, 1977; Pittenger & Jenkins, 1979).

Introduction to the Present Experiments

A series of six experiments was conducted to examine and extend Jenkin's studies of memory for everyday pictorial events. At outset, the intention was to replicate the recognition performance pattern observed by Jenkins and test the validity of a fusion interpretation. It was hoped that confirmation of fusion would be established and that his paradigm would provide a way of developing a more explicit account of abstraction of pictorial events. However, the experiments described in this thesis did not support fusion.

Improvements were made both to the technical quality of the stimulus material and to the experimental procedure. Two types of stimulus material were used in the present experiments. One type consisted of six short action-sequences along the same lines as those described by Jenkins. They were all of the same length, and standardised with respect to photographic conditions and construction. This material was used in the first two experiments where an attempt was made to vary the *degree* to which material could be disordered. Six sequences allowed both for disordering individual sequences whilst keeping the separate sequences distinct, and for disordering between sequences as well as within. The second type of stimulus material consisted of one long action-sequence specifically designed to be difficult to reconstruct when disordered. It became necessary to produce a sequence of this type in order to establish whether subjects were *able* to reconstruct a particular disordered sequence.

Control slides (called 'Transformations' in the present experiments) were produced in a way which systematically attempted to maximize their similarity to Originals whilst transforming one of four specified elements within the picture along the same lines as Mandler. They were photographed at points during the original sequence and were part of that overall progression of slides but transformed in specific ways. Belongings were withdrawn from the original sequences in the manner described by Jenkins, but in the present experiments the proportions of test stimuli used by Jenkins were applied exactly, with one third Originals, one third Belongings, and one third Transformations being presented in the recognition test.

It was possible that Jenkins' results were specific to the material he used, since his Belongings were always Belongings and Originals were always Originals. Therefore, the present studies used sampling procedures in which slides from the action-sequences served as Originals and Belongings interchangeably. The use of two types of stimulus material - the short and long sequences - also provided a more general test of fusion.

A change in the analysis of the present studies concerned the measure used to assess memory performance. Jenkins relied on a direct comparison between false acceptance of Belongings and Controls. In the present experiments performance was measured in terms of relative discrimination between Belongings and Originals and Transformations and Originals. Hit rates on Originals and false alarm rates on distractors were converted into measures of d'.

One d' value gave a measure of the subjects' ability to discriminate between Originals and Belongings (d'O:B) and a second d' value gave a measure of the subjects' ability to discriminate between Originals and Transformations (d'O:T). A comparison of the two measures allowed assessment of differential performance, relatively free of response bias. This use of d' as a measure of memory performance is consistent with a number of studies of recognition memory (Connor, 1977; Franken & Rowland, 1978; Loftus & Bell, 1975; Loftus & Loftus, 1976; Mandler & Ritchey, 1977; Pezdek, 1977). In the present experiments, therefore, the effect which Jenkins interpreted as evidence of fusion - i.e. false alarms on Belongings being higher than false alarms on Controls and therefore closer to hits on Originals - is identified as present when d'O:B is significantly lower than d'O:T. In later experiments reaction times were also measured in an attempt to obtain a more sensitive index of recognition performance.

The first experiment replicated Jenkins' findings that when subjects are shown action-sequences which depict everyday events and are then given a recognition test in which there are two types of distractors, d'O:B is significantly lower than d'O:T. An attempt was made to repeat Jenkins' finding that when subjects see the action-sequences disordered during presentation they no longer falsely accept Belongings. However, randomization of the temporal order of slides in sequences did not affect recognition performance.

In the second experiment an alternative method was used to test the fusion interpretation. Subjects were shown the action-sequences within an incidental learning situation, with two types of orienting task. Some subjects were asked to perform a semantic task during acquisition, whilst others performed a non-semantic task. Since fusion depends on abstracting the meaning of a slide sequence, semantic processing of acquisition material should enhance fusion whereas non-semantic processing should impair the process. It was, therefore, predicted that if the fusion hypothesis is valid, d'O:B should be significantly lower than d'O:T following semantic processing of acquisition material, but not following non-semantic processing. This second experiment also found no support for the fusion interpretation, since non-semantic processing within an incidental learning paradigm did not attenuate subjects' tendency to falsely accept Belonging slides.

A third experiment increased the retention period between acquisition and testing. Several studies have shown that the semantic content of stimulus material is retained for longer periods than the structural content. It was therefore suggested that if fusion is a main determiner of the relative d' values, then when impairment due to delayed testing becomes obvious, Transformations should not be less accurately detected in a recognition test, since they should be no less inconsistent. Therefore impairment would be reflected in lowered d'O:B values. Conversely, if false acceptance of Belongings is primarily due to relative featural similarity, then with a longer retention interval, discrimination of both Belongings and Transformations should deteriorate - that is discrimination of the featural details of both types of distractors should be impaired. Since there was no difference between distractors in recognition performance after a longer retention interval, support for fusion was not found.

The fourth and fifth experiments returned to a more extensive investigation of the effects of presenting temporally disordered material. Before any definitive test of the effect of disordering material could be made, it now seemed essential to demonstrate that subjects who saw the sequence in a disordered version were not *able* to reconstruct the underlying event. New stimulus material was, therefore, made for the final experiments. A long action-sequence was designed in such a way that it would be difficult to reconstruct when the slides were not in a temporarily ordered sequence. The fourth experiment established that subjects who saw a disordered version of this sequence could not re-order it successfully. The fifth experiment then examined the effect of disordering on recognition memory performance. Despite careful selection of material to ensure that disordering actually destroyed the event , and despite the use of potentially more sensitive measures of recognition performance, d'O:B was significantly lower than d'O:T following Disordered Presentation.

The final experiment made a preliminary attempt to find direct evidence for the alternative featural account of relative discrimination of distractors. The aim of this experiment was to keep the meaning of the distractor set consistent with the underlying event presented in acquisition, whilst varying the degree of featural similarity between Originals and distractors. Although the experiment was unsuccessful, it was suggested that if this method were used with material specifically designed for this purpose it might provide a means of studying recognition memory for pictorial events where a more explicit distinction could be made in the test material between featural similarity and meaning.

EXPERIMENT I

Introduction

The main aim of the first experiment was to replicate the finding reported by Jenkins et al. (1978) that subjects falsely accept a significantly greater number of Belonging slides than Transformation slides on a recognition test following the presentation of a pictorial sequence which depicts an event. This is the specific response pattern from which Jenkins has inferred the process of fusion in his experiments. In the present research this response is defined in terms of d' scores and is present when d'O:T (the discrimination between Originals and Transformations) is significantly greater than d'O:B (the discrimination between Originals and Belongings).

A further aim of the first experiment was to test the fusion account by presenting some subjects with disordered acquisition sequences. In one experiment described in the Jenkins et al.(1978) paper, the individual slides which made up a sequence were shown to subjects in a random order during acquisition. Jenkins had argued that if disordering destroyed the event then fusion would not occur and subjects would correctly reject Belongings. On the other hand if a tendency to falsely accept Belongings were caused by Belongings having more features in common with Originals than had Transformations, then disordering the sequence would have no effect. That is, despite disordering, subjects would still falsely accept Belongings.

The result of Jenkins' test was ambiguous. When presentation was disordered, subjects falsely accepted the Belongings of one of his two sequences but correctly rejected the Belongings of the other sequence. Although Jenkins acknowledged that "the similarity problem is a critical one and cannot be dismissed by the results of one experiment" (1978, p. 161), he argued strongly in favour of the fusion hypothesis.

According to Jenkins, if fusion is responsible for recognition memory performance of pictorial events, disordered presentation of an event sequence should increase d'O:B so that there is no significant difference between d'O:B and d'O:T. In the present experiment, the effect of disorder was examined at two levels by showing six slide sequences in two different ways. In a 'Disordered-Within' condition, the individual slides within each sequence were disordered but the sequences were kept distinct from one another. A greater degree of disordering was used in a 'Disordered-Within-and-Between' condition, where all the slides were disordered; that is not only were the individual slides within a sequence disordered but also the sequences themselves were mixed together.

The Disordered-Within condition manipulated order to the same extent as in the Jenkins experiment which found support for the fusion hypothesis. The Disordered-Within-and-Between condition was introduced as an additional disordering manipulation in an attempt to ensure that subjects could not re-order the disordered material.

The fusion hypothesis would be supported by smaller d'O:B than d'O:T values in the Ordered condition along with no difference between the d' measures for the Disordered conditions. That is, the fusion hypothesis predicts an interaction between the variables of stimulus type and degree of ordering. An interaction would be absent if confusion of stimulus features were responsible for the d'O:B versus d'O:T difference.

Method

Subjects

Subjects were 16 women and eight men. Their ages ranged from 17 to 42 years with a mean age of 22.75 and a standard deviation of 7.10. The design was a 3 (Ordered versus Disordered-Within versus Disordered-Within-and-Between) X 2 (Belonging versus Transformation) factorial with repeated measures on the second factor. Eight subjects were randomly assigned to each of the three groups. Subjects took part in the experiment individually and all instructions were written.

Stimulus Materials

A Nikkormat camera fitted with an F2 50mm lens and ultra violet filter was used to make six sets of slides. The camera was set on a tripod so that each of the six sets of slides was taken from a fixed stationary point. Some sequences were taken by natural light whilst for others it was necessary to use electronic flash lighting. Kodachrome ASA 64 film was used throughout.

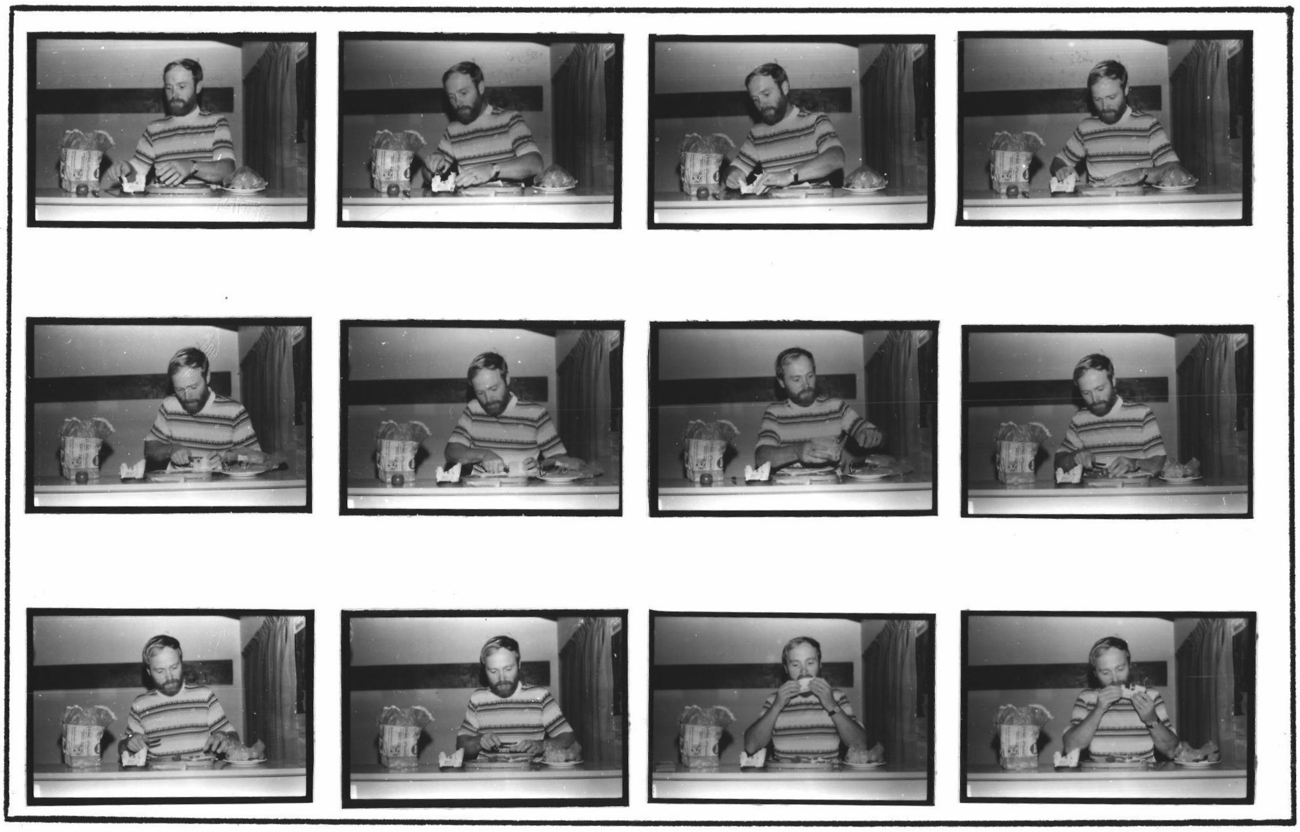
Each set consisted of 20 slides, 16 of which depicted an action-sequence in a consistent and progressive manner and four of which changed some aspect of the sequence according to a rule. Each action--sequence was a distinct activity performed by one person. The activities were as follows: (a) a woman paying a bill; (b) a woman ironing a tablecloth; (c) a woman hanging out washing; (d) a man making a sandwich; (e) a man chopping firewood; (f) a man lighting a fire. An example of one of the sets of 20 slides is shown in Figure 1.

Before photographing each activity, it was rehearsed by the actor and timed. The real-time duration of activities varied between 80 and 120 seconds. When the duration of an activity had been established, the time was divided by 16 so that each of the 16 photographs depicting the activity could be taken on cue at equal intervals throughout the action. Thus, depending on the duration of the activity, a photograph was taken every five, six, or seven seconds. In this way, six sets of 16-frame sequences were made, each one depicting an every-day activity.

The four Transformation slides changed some aspect of the activity so that although each was similar to the 16 slides which depicted the activity, it differed from them according to a specific rule. For each action-sequence, one Transformation slide was made

Figure 1: A short set of slides

Originals (Ordered-Presentation)



Belongings

Transformations

Addition	Token	Perspective	Orientation
			

according to each of these four rules: (a) Addition: An object which might reasonably appear in such a setting but which had not been included was added. For example, a hearth-brush was added to the 'man lighting a fire' sequence. (b) Token: Some change was made in the appearance of the actor. For example, the 'man making a sandwich' wore glasses. (c) Perspective: A wide-angle lens (F2 28mm) was substituted for the F2 50mm lens. (d) Orientation: A duplicate of one of the 16 slides was made and reversed with respect to left-right orientation. In photographing the scenes, the four Transformation slides were spaced at approximately equal intervals over the course of the whole activity. Addition, Token, and Perspective slides were made at the same time as the 16 slides depicting the activity. The action was 'frozen' at the appropriate moment, the change made, the photograph taken, and then the original scene was restored.

Apparatus

Slides were projected onto a white wall by means of two Kodak Carousel SAV 2000 projectors which operated in direct succession. The projected image was .45 x .69m both for acquisition and test stimuli, and subjects sat approximately 1.5m from the screen. Each slide was shown for 5 seconds and there was no interstimulus interval. Exposure duration was determined by Gerbands tachistoscopic shutters mounted in front of the projector lenses. Shutter operation was timed by standard relay and electronic programming apparatus.

Procedure

Eight randomly determined orders of the six sets of slides were generated. The first subject in each condition saw one of these orders, the second subject saw another, and so on. Four groups of evenly spaced Belongings could be withdrawn from any sequence of 16 slides depicting an activity: that is slide numbers 1, 5, 9, 13; 2, 6, 10, 14; 3, 7, 11, 15; or 4, 8, 12, 16. The group of Belongings withdrawn from a sequence was determined randomly for each of the six sequences seen by the first subject in each condition, and then again for each of the six sequences seen by the second subject in each condition, and so on. The remaining 12 Original slides were shown during acquisition. Four of these were selected at random to be shown during testing. All first subjects in each condition were shown the same Original slides during testing, as were second subjects, and so on. In this way, eight different selections of 72 slides were presented during acquisition, one for each of the eight subjects within each group.

Subjects in the Ordered condition saw the slides within each sequence in their correct chronological order. Subjects in the Disordered-Within condition were presented with randomly ordered slides within each sequence but the sequences were shown successively. Subjects in the Disordered-Within-and-Between condition saw all 72 slides completely randomized across all six sequences of 12 slides.

The test material consisted of four Originals, four Belongings and four Transformations from each of the six slide sets. The order of these 72 test slides was determined randomly with no distinction being made between the sets. An order was generated for all first subjects in each condition, another order was generated for all second subjects, and so on.

In the acquisition phase of the experiment subjects were told that they were taking part in a memory experiment and were asked to try to remember the slides which were about to be shown to them. The 72 acquisition slides were then presented. Three minutes after the end of the acquisition phase, all subjects were tested in the same way. During the interval between the end of the acquisition phase and the start of testing, they read the test instructions and were examined in their understanding of them. Subjects were then required to view the test slides and complete a form by assigning a number from 1 to 4 for each slide indicating how certain they were that they had or had not seen that slide before. "1" meant "I am certain that I have seen this slide before", "2" meant "I am fairly sure that I have seen this slide before", "3" meant "I am fairly sure that I have not seen this slide before", and "4" meant "I am certain that I have not seen this slide before". Subjects were not told about the prior probabilities of the target and distractor slides.

Results

Confidence rating data were collapsed into "Yes" and "No" responses, where 1 and 2 indicated "Yes" and 3 and 4 indicated "No". The mean (and standard deviation) Hit and False Alarm rates are given in Table 1. Hit rates for the Originals and False Alarm rates for Belongings and Transformations were converted into d' scores for each subject from tables in Green and Swets (1966). The mean (and standard deviation) d' scores for each type of distractor and for the three kinds of presentation condition are given in Table 2.

An ANOVA[4] which examined Presentation Type (between-subject) and Distractor Type (within-subject) showed no effect of Presentation, but a strong effect of Distractor, F (1,21) = 35.18, MSe = .11, $p < .001$ with d'O:T higher than d'O:B. There was no interaction.

Discussion

The main effect reported by Jenkins et al. (1978) was replicated - d'O:T was significantly higher than d'O:B. However, this response pattern was present in both Ordered and Disordered

[4] Appendix A presents summary tables of all ANOVAS reported in the thesis.

TABLE 1

Mean (and standard deviation) Hit and False Alarm proportions for Originals, Belongings, and Transformations for Ordered, Disordered-Within, Disordered-Within-and-Between presentation conditions

Presentation Type	Test Slide Type		
	Original	Belonging	Transformation
Ordered	.76 (.11)	.51 (.13)	.29 (.11)
Disordered-Within	.84 (.10)	.58 (.13)	.30 (.09)
Disordered-Within and-Between	.79 (.11)	.43 (.13)	.30 (.09)

TABLE 2

Mean (and standard deviation) d' scores for Originals to Belongings (O:B) and Originals to Transformations (O:T) for Ordered, Disordered-Within, Disordered-Within-and-Between presentation conditions

Presentation Type	Distractor Type	
	O:B	O:T
Ordered	.73 (.38)	1.35 (.68)
Disordered-Within	.85 (.40)	1.59 (.43)
Disordered-Within and-Between	1.02 (.35)	1.37 (.46)

conditions. The interaction between presentation conditions and relative d' measures predicted by a fusion hypothesis was not present and the results of this experiment did not therefore support fusion.

The more general question of the effect of order on memory is pertinent to the present result. On the whole, it has been shown that when organised material is presented to subjects, it is remembered more accurately than when it is presented in an unorganised fashion. Improvement in memory performance has been demonstrated with associative organisation (Jenkins & Russell, 1952; Bower, 1972), category organisation (Bousefield, Cohen & Whitmarsh, 1958; Cofer, Bruce & Reicher, 1966), and subjective organisation (Mandler, 1972; Tulving, 1962). The effect of organisation on memory has been demonstrated more clearly with recall than recognition performance and has been tested most extensively with verbal material (Kintsch, 1974). Studies which have shown impairment in perception and recognition of pictorial material have manipulated spatial order (Biederman, 1972; Biederman, Glass & Stacy, 1973; Biederman, Rabinowitz, Glass & Stacy, 1974; Mandler & Johnson, 1976; Mandler & Parker, 1976; Mandler & Ritchey, 1977).

The present experiment manipulated temporal order and indicated that temporal disordering of pictorial sequences does not impair recognition performance. It remains possible, however, that the manipulation of temporal order was ineffective simply because subjects in the disordered conditions spontaneously re-organised the material. Specifically, it may be the case that the slide sequences used in the present study depicted events so well specified by the individual slides that subjects were able to reconstruct the underlying events.

Three subsequent findings indicated that it was unlikely that order was ineffective because of subjective reorganisation. Firstly, a post-hoc comparison was made between the Ordered Presentation and the more extremely disorganised condition - the Disordered-Within-and-Between Presentation. An ANOVA showed that d'O:T was significantly higher than d'O:B (F, (1,21) = 31, MSe = .07, $p < .001$) but there was no effect of presentation and no interaction.

Secondly, the next experiment in this series demonstrated that disordering was not only perceived by subjects but that the intended degrees of disordering were perceived as significantly different.

Thirdly, later experiments in this series explicitly investigated the problem of subjective reorganisation and showed that reconstruction of disorganised material is unlikely to be responsible for the absence of an order effect.

It is more likely that the type of information which determined recognition performance was not affected by disordering. In her examination of the effect of disordering on recognition of pictorial material Mandler found that although memory for spatial relationships between objects was impaired by disorganisation, memory for specific object information was not affected (Mandler & Parker, 1976; Mandler & Ritchey, 1977). Mandler's results point to the possibility that specific slide information was the main determiner of recognition accuracy in the present experiment.

EXPERIMENT II

Introduction

The second experiment tested the fusion hypothesis by examining performance within an incidental rather than intentional learning paradigm. In this way, acquisition task conditions could be manipulated directly. For fusion to occur, subjects must process material semantically. If this is made difficult by requiring them to perform a non-semantic task with the material during acquisition, then fusion should be impaired. Conversely, if subjects are encouraged to engage in semantic processing during acquisition, particularly in processing which emphasises the sequential nature of the event depicted by the material, then fusion should be facilitated.

Subjects in a Semantic condition were given an orienting task specifically related to the thematic coherence of the total acquisition set - they were required to judge the logical relationship between each slide. Subjects in a Non-Semantic condition were required to concentrate on a gross global feature of each slide - they were required to judge the relative brightness of each slide during acquisition.

The fusion hypothesis would be supported if d'O:T were significantly higher than d'O:B in the Semantic condition along with no significant difference between d'O:T and d'O:B in the Non-Semantic condition. On the other hand, if d'O:T were significantly higher than d'O:B along with no interaction between the relative d' values and the acquisition task conditions, then the results would indicate that false acceptance of Belongings was due primarily to their featural similarity to Originals.

It was predicted that the overall d' values would be higher for the Semantic group compared to the Non-Semantic group since subjects who perform semantic orienting tasks during acquisition tend to give evidence of more accurate and more durable memory than subjects who perform non-semantic orienting tasks (Craik & Lockhart, 1972).

Order was varied again in the second experiment with the intention of discovering whether subjects *perceived* the different degrees of disordering intended by the arbitrary manipulation of stimulus material during presentation. The acquisition data from subjects in the Semantic group who were required to judge the logical progression of the material in the three conditions were used to assess subjects' perception of order. If the three different types of presentation represented varying degrees of order then the total number of "No" responses made to the question "Does this slide follow on logically from the previous slide?" should vary accordingly.

Method

Subjects

Subjects were 23 women and 25 men. Their ages ranged from 16 to 36 years with a mean of 20.96 and a standard deviation of 4.09. The design was a 2 (Semantic versus Non-Semantic) X 3 (Ordered versus Disordered-Within versus Disordered-Within-and-Between) X 2 (Belonging versus Transformation) factorial with repeated measures on the last factor. Eight subjects were randomly assigned to each of the 6 between-subject groups. Subjects took part in the experiment individually, and all instructions were written.

Procedure

The procedure and stimuli were the same as in Experiment 1 except that subjects were told that they were taking part in a perception experiment. Subjects in the Semantic condition were asked to judge the logical progression in the acquisition slides. They had to complete a form during presentation by marking "Yes" or "No" for the question "Does this slide follow on logically from the previous slide?" for every slide except the first. Subjects in the Non-Semantic condition were asked to judge the relative brightness levels of the slides. They had to complete a similar form by marking "Yes" or "No" for the question "Is this slide brighter than the previous slide?" for every slide except the first. No mention was made of the memory test until after the acquisition phase had been completed, when subjects were given the same recognition test instructions as in Experiment I.

Results

Confidence rating data from the test phase of the experiment were collapsed into "Yes" and "No" responses where 1 and 2 indicated "Yes" and 3 and 4 indicated "No". The mean (and standard deviation) Hit and False Alarm rates are presented in Table 3. The d' values calculated from the Hit and False Alarm rates are given in Table 4.

An ANOVA which examined Presentation Type (between-subject), Task Type (between-subject) and Distractor Type (within-subject) showed no effect of Presentation but a reliable effect of Task, $F\ (1,42) = 11.34$, MSe = .32, $p < .01$, and a strong effect of Distractor, $F\ (1,42) = 137.14$, MSe = .07, $p < .001$. No interactions were significant. Subjects in the Semantic group performed significantly better than subjects in the Non-Semantic group, but there was no significant difference between the groups with respect to distractors. For both the Semantic group and the Non-Semantic group d'O:T was significantly higher than d'O:B.

Acquisition data from the Semantic group were examined since the number of "No" responses to the question "Does this slide follow on logically from the previous slide?" gives a measure of the perceived degree of logical order in the acquisition slides where

TABLE 3

Mean (and standard deviation) Hit and False Alarm proportions for Originals, Belongings, and Transformations for Ordered, Disordered-Within, and Disordered-Within-and-Between presentation conditions and for Semantic and Non-Semantic tasks

Presentation Type	Task Type	Test Slide Type		
		Original	Belonging	Transformation
Ordered	Semantic	.84 (.12)	.57 (.12)	.38 (.08)
	Non-Semantic	.65 (.21)	.56 (.21)	.33 (.17)
Disordered-Within	Semantic	.81 (.11)	.59 (.13)	.37 (.10)
	Non-Semantic	.59 (.15)	.55 (.17)	.26 (.11)
Disordered-Within-and-Between	Semantic	.71 (.17)	.61 (.13)	.37 (.10)
	Non-Semantic	.55 (.19)	.48 (.17)	.30 (.14)

TABLE 4

Mean (and standard deviation) d' scores for Originals to Belongings (O:B) and Originals to Transformations (O:T) for Ordered, Disordered-Within, and Disordered-Within-and-Between presentation conditions and for Semantic and Non-Semantic tasks

Presentation Type	Task Type	Distractor Type	
		O:B	O:T
Ordered	Semantic	.93 (.64)	1.42 (.51)
	Non-Semantic	.19 (.48)	.98 (.42)
Disordered-Within	Semantic	.68 (.29)	1.25 (.33)
	Non-Semantic	.13 (.39)	.94 (.48)
Disordered-Within-and-Between	Semantic	.32 (.23)	.95 (.44)
	Non-Semantic	.24 (.43)	.74 (.51)

a small number would indicate a high degree of order. The mean frequencies (and standard deviation) of "No" responses were 10.75 (5.85) for the Ordered condition[5], 34.13 (7.57) for the Disordered-Within condition, and 55.38 (6.41) for the Disordered-Within-and-Between group. An ANOVA which examined Presentation Type (between-subject) showed an effect of Presentation, $F\ (2,21) = 123.93$, $MSe = 37.52$, $p < .001$. A scheffé test showed that perceived orderliness increased systematically over the three presentation types (all $ps < .001$).

Discussion

The results of this experiment did not support the fusion hypothesis. Although the Semantic groups' performance was significantly more accurate than the Non-Semantic groups', and although the Semantic groups' d'O:B scores were significantly lower than their d'O:T scores, the crucial interaction predicted by a fusion hypothesis was not present. Despite a feature-specific orienting task, the Non-Semantic groups' d'O:B scores were also

[5] In the Ordered condition at least five "No"s would be expected given that six distinct sequences were shown.

significantly lower than their d'O:T scores, and the degree of their differential discrimination of distractors was as great as that of the Semantic group.

The finding that the Semantic groups' overall recognition performance was better than that of the Non-Semantic group is consistent with the general finding of the "levels-of-processing" research. Although the early description of depth of processing has been modified (Craik & Tulving, 1975), the distinction between non-semantic and semantic processing has flourished in the literature. Few experimenters have questioned the validity of such a distinction, although as several researchers have observed (Baddeley, 1978; D'Agostino, O'Neill & Paivio, 1977; Nelson, 1977; Nelson & Vining, 1978; Postman, 1975; Wolk, 1974), the distinction lacks precision and the representative tasks depend on individual and intuitive interpretations of these types of processing.

However, the effect of differential encoding on memory performance has been demonstrated under a variety of experimental conditions. Semantic processing has been shown to result in better memory than non-semantic processing of nouns (Schulman, 1971), sentences (Rosenberg & Schiller, 1971), and faces (Bower & Karlin, 1974). The effect has been observed within both incidental (Hyde & Jenkins, 1973), and intentional (Treisman & Tuxworth, 1974) learning paradigms, and where memory has been assessed by free recall (Marslen-Wilson & Tyler, 1976), cued recall (Bobrow & Bower, 1969), and recognition (Arbuckle & Katz, 1976).

Various tasks designed to evoke semantic processing have been used. These include judging the stimuli for pleasantness (Hyde & Jenkins, 1973), familiarity (Rosenberg & Schiller, 1971), meaningfulness (Mistler-Lachman, 1974), suitability (Craik & Tulving, 1975), and contextual appropriateness (Mistler-Lachman, 1974). Similarly, a variety of non-semantic tasks have been devised and include estimating the number of 'e''s in a word (Hyde & Jenkins, 1973), crossing out vowels (Tresselt & Mayzner, 1960), and noting the case of a word (Craik, 1973).

The present experiment extended the levels-of-processing research by showing that a task which required subjects to judge the logical relationship between complex pictorial stimuli resulted in more accurate recognition performance than a task which required subjects to judge a physical relationship between stimuli.

The absence of any effect of order was relevant to this finding. Not only has it been shown that semantic tasks result in more accurate memory than non-semantic tasks, but it has also been shown that the differential encoding effect is sensitive to different degrees (or 'depths') of semantic processing (Craik, 1973; Klein & Saltz, 1976; Mistler-Lachman, 1974). Moreover the effect has been demonstrated more directly by varying the degree to which stimuli *can* be processed (Marslen-Wilson & Tyler, 1976). It was puzzling, therefore, to find that disorganisation of the logical relationships between slides did not appear to affect the performance of the Semantic group whose task was designed to elicit logical processing.

Acquisition data from the Semantic group had shown that the manipulation of chronological order was perceived by subjects. Moreover the manipulation systematically changed degrees of disordering. It might, therefore, be expected that the degree to which the material *could* be processed semantically would have been apparent in the Semantic groups' performance. Conversely, given the absence of an order effect in Experiment I, and given the featural emphasis of the non-semantic task, disordering would not have been expected to affect the Non-Semantic groups.

The interaction relevant to this question (Presentation X Task) in the main ANOVA was not significant. However, it was decided to make two separate post-hoc examinations of the effect of order. An ANOVA which examined d' scores for the Semantic group alone, found an effect of Presentation F, (2,21) = 3.55, MSe = .33, $p < .05$ and Distractor, F, (1,21) = 96.25, MSe = .04, $p < .001$, there was no interaction. A subsequent Scheffé test showed that the Ordered Presentation group's performance was significantly better than the Disordered-Within-and-Between Presentation groups' performance, $p < .001$. An ANOVA which examined the d' scores for the Non-Semantic group alone, found an effect of Distractor, F, (1,21) = 58.60, MSe = .10, $p < .001$, but no effect of order and no interaction.

These separate analyses revealed that disordering acquisition material affected recognition accuracy in the Semantic group but did not affect performance in the Non-Semantic group. Although this finding is not a strong one, it is consistent with the literature which has shown that semantic encoding effects are sensitive to the degree to which material can be processed semantically. Moreover, with respect to the question of the effects of temporal disorganisation of pictorial sequences, the findings show that when an acquisition task specifically directs subjects' attention to the chronological relationships between slides, order is an effective variable. However, since it is ineffective when the task is unrelated to chronological order (the Non-Semantic group) and since it is ineffective in the context of an intentional remembering condition (Experiment I), it seems unlikely that apprehension of the logical relationship between individual slides in a pictorial sequence (or fusion) is the primary determinant of recognition memory.

EXPERIMENT III

Introduction

The third experiment tested fusion by varying the retention interval between acquisition and test. Not only has the majority of research investigating duration of memory shown that performance deteriorated over time, but more specifically, several studies have revealed differential rates of decay for the deep and surface structure of the material.

Although some studies suggest that long-term retention of the featural details of verbal material is greater than has been commonly thought, (Bates, Masling & Kintsch, 1978; Keenan, McWhinney & Mayhew, 1977; Kintsch & Bates, 1977), there is considerable evidence that the decay rate of the surface structure of sentences is faster than the decay rate of the semantic structure (Anderson, 1974; Bartlett, 1932; Graesser & Mandler, 1975; Kintsch, 1974; Kintsch & Keenan, 1973; Mistler-Lachman, 1974; Sachs, 1967, 1974).

Several researchers have argued that pictorial material has equivalent surface and deep structures (Craik & Lockhart, 1972; Nelson, Reed & McEvoy, 1977; Rafnel & Klatzky, 1978), which appear to have similarly differential decay rates. Memory for pictorial material is more accurate when the stimulus elements are presented in a meaningful context as opposed to no context or a meaningless context (Bower, 1970; Epstein, Rock & Zuckerman, 1960; Mandler &

Ritchey, 1977; Paris & Mahoney, 1974; Posner & Keele, 1970). Furthermore, it has been shown that memory for pictures which are meaningful as opposed to ambiguous or nonsensical is more durable (Baggett, 1975; Bower, Karlin & Dueck, 1975; Klatzky & Rafnel, 1976; Rafnel & Klatzky, 1978).

The general principle which emerges from this research is that memory for the gist of stimulus material is more durable than memory for the details. This principle suggested a way of testing the fusion hypothesis. If memory were impaired by increasing the retention interval, then it could be assumed that the impairment would be due primarily to a loss of memory for the details of the material. That is, at the point at which performance became significantly worse in delayed testing, subjects would be responding predominantly on the basis of their memory for the event. And the nature of the impairment could be used specifically to examine Jenkins' argument that at immediate testing a lower d'O:B than d'O:T was due to fusion.

If fusion is responsible for the difference in the relative d' values at immediate testing, then when recognition memory is impaired by delay, the impairment should not be reflected in a lowered d' O:T value, since there is no reason to assume that Transformations would become any *less* inconsistent with the event. Therefore, memory impairment should be due primarily to a lowered d'O:B value. On the other hand, if the difference between d'O:B and d'O:T at immediate testing is attributable to the relative featural similarity between distractor types and Originals, then when recognition memory is impaired by delay, the impairment should be reflected in a lowered d' value for both types of distractors, since memory for the featural detail of both Transformations and Belongings should be equally affected by delay. Three retention intervals were used to examine the fusion hypothesis in this way: immediate testing, and delays of 24 hours and one week.

Method

Subjects

Subjects were 22 women and 26 men. Their ages ranged from 16 to 41 years with a mean age of 23.78 and a standard deviation of 6.55. The design was a 3 (Immediate versus 24 Hours versus One Week delay) X 2 (Belonging versus Transformation) factorial with repeated measures on the second factor. Sixteen subjects were randomly assigned to each of the three groups. Subjects took part in the experiments individually, and all instructions were written.

Procedure

The procedure and stimuli were the same as in Experiment I, except for the following changes: only ordered sequences were shown during acquisition; slides were shown for four seconds each during acquisition; and subjects were asked only to indicate "Yes" or "No" during testing to the question "Have you seen this slide before?".

Retention interval was now varied so that subjects in the delayed conditions returned either 24 hours or one week later for the test phase.

Results

The mean (and standard deviation) Hit and False Alarm rates for each retention interval are shown in Table 5. The d' values calculated from these are given in Table 6.

An ANOVA which examined Retention Interval (between-subject) and Distractor Type (within-subject), showed a reliable effect of Retention, $F(2,45) = 3.36$, MSe = .39, $p < .05$, and a strong effect of Distractor, $F(1,45) = 19.92$, MSe = .49, $p < .001$, with d'O:T higher than d'O:B. There was no interaction. Subsequent Scheffé tests showed that performance was less accurate after a retention interval of one week than after 24 hours ($p < .01$) or on immediate testing ($p < .025$). There was no significant difference in recognition performance between immediate or 24-hour delayed testing.

TABLE 5

Mean (and standard deviation) Hit and False Alarm proportions for Originals, Belongings, and Transformations for Immediate, 24 Hours, and One Week retention intervals

Retention Interval	Test Slide Type		
	Original	Belonging	Transformation
Immediate	.77 (.11)	.59 (.09)	.32 (.12)
24 Hours	.81 (.14)	.63 (.14)	.38 (.10)
One Week	.77 (.09)	.67 (.13)	.44 (.17)

TABLE 6

Mean (and standard deviation) d' scores for Originals to Belongings (O:B) and Originals to Transformations (O:T) for Immediate, 24 Hours, and One Week retention intervals

Retention Interval	Distractor Type	
	O:B	O:T
Immediate	.59 (.44)	1.25 (.64)
24 Hours	.62 (.47)	1.26 (.47)
One Week	.27 (.39)	.89 (.52)

Discussion

There was a significant impairment in recognition performance after a delay of one week, but no deterioration after a delay of 24 hours. Memory for complex photographic slide sequences appears to be stable over the course of at least one day, but has deteriorated after a week. This finding is consistent with studies which have examined the effects of retention interval on memory for pictorial line drawings (Baggett, 1975; Mandler & Ritchey, 1977). Although performance was significantly impaired after a week, the difference between d'O:B and d'O:T was virtually identical for all three conditions: (d'O:T - d'O:B - .66 for Immediate, .64 for 24-Hour delay, .62 for One Week delay). Since impairment over time was reflected in both distractor discrimination values, this experiment did not support the argument that fusion is responsible for the lower d'O:B relative to d'O:T at Immediate testing

EXPERIMENT IV

Introduction

The fourth experiment returned to the question of the effect of disordering material during presentation. In the first experiment in this series, the difference between d'O:B and d'O:T measures for subjects who viewed the disordered sequences was no different from that for subjects who viewed ordered sequences. The absence of this interaction between stimulus type and order, also apparent in the second experiment, was interpreted as a serious failure to find support for the fusion hypothesis. However, a criticism of this test of fusion is that subjects may reconstruct the original event from the disordered sequence. Indeed, when disordering has not demonstrated fusion in previous studies, this argument has been offered consistently (Kraft & Jenkins, 1977; Jenkins et al., 1978; Pittenger & Jenkins, 1979).

For stories presented verbally, Baker (1978) found no evidence that subjects reconstructed story order when shown disordered stories, although the majority of research has indicated that subjects are likely to re-order stories into their natural schematic order (Bower, Black & Turner, 1979; Kintsch, Mandel & Kozminsky, 1977; Mandler, 1978; Stein & Glenn, 1978). Therefore, the primary aim of the fourth experiment was to challenge the re-ordering criticism directly by producing a pictorial action-sequence which could not be re-ordered. A test of the fusion hypothesis via disordering could therefore be made with greater confidence.

New material was produced for this experiment. A long slide sequence was specifically devised to make successful reconstruction of the disordered story unlikely. The stimulus set, whose story was a series of encounters between men and women which resulted in new friendships, is shown in Figure 2. The exact order of the encounters between individuals was crucial to an understanding of the story. The intention was to ensure that the presence of any two people in a slide seen out of its temporal context was devoid of the implication of their connection - a factor crucial to the meaning of the story.

The first phase of the experiment attempted to establish whether subjects who saw the acquisition slides in a random order were *able* to reconstruct the underlying story. Two experimental groups - Ordered Presentation and Disordered Presentation were asked to write an account of the story contained in the set of slides they saw. The Disordered Presentation group were specifically asked to attempt to reconstruct the story. These accounts were later assessed by independent judges who had seen the ordered sequence. They were rank-ordered and rated for their closeness to the actual story depicted by the ordered version.

The subjects were also given the standard recognition test after writing their accounts. In this experiment the recognition test not only recorded subjects' decisions about test slides but also their reaction times for those decisions. This additional

Figure 2: The long set of slides

Originals (Disordered – Presentation)

Figure 2: continued (i)

Originals (Disordered - Presentation)

Figure 2: continued (ii)

Originals (Disordered-Presentation)

Figure 2: continued (iii)

Belongings

Figure 2: continued (iv)

Transformations

measure of performance was introduced since it seemed possible that reaction time might provide a more sensitive measure of recognition performance.

Method

Subjects

Subjects were 16 women and eight men. Their ages ranged from 17 to 29 years with a mean age of 19.25 and a standard deviation of 2.69. The design was a 2 (Ordered versus Disordered) X 2 (Belonging versus Transformation) factorial with repeated measures on the second factor. Twelve subjects were randomly assigned to each of the two groups. Eleven women and one man were judges for the rank ordering phase of the experiment. Their ages ranged from 17 to 28 years with a mean age of 21.42 and a standard deviation of 3.60. Subjects and judges took part in the experiment individually and all instructions were written.

Stimulus Material

A Nikkormat camera fitted with an F2 50mm lens and ultra violet filter was used to make one set of slides. The camera was set on a tripod so that each slide was taken from a fixed stationary point.

The sequence was taken by natural light and Kodachrome ASA 64 film was used.

The total slide set used in the experiment consisted of 120 slides, of which 96 depicted an action-sequence in a consistent and progressive manner and of which 24 changed some aspect of the sequence according to a rule. The action-sequence involved four actors and the setting was a park bench. The story was a series of encounters between men and women which resulted in new friendships. The outline of the story is as follows:

A young man is seated on a park bench. He opens a bag of fruit and begins to eat. An older man arrives and sits on the other end of the bench. The two men converse for a while and then the older man begins to read a magazine. A young woman appears and arranges her hair then sits on the bench between the two men. She lights a cigarette. The young man strikes up a conversation with her and offers her some fruit which she accepts. They talk and finally leave together. The older man watches them go and then continues to read his magazine. An older woman arrives and sits on the bench. She begins to feed some birds whilst engaging in conversation with the older man. The older man leaves after a while but forgets his magazine. The woman picks up the magazine and is reading it when he returns to collect it. He sits down again and they talk and then both feed the birds and finally leave together.

Before photographing the action-sequence it was rehearsed by the actors and timed. The real-time duration of the sequence was approximately 11 minutes and each of the 96 slides depicting the sequences was taken on cue at equal intervals of 7 seconds throughout the action.

The twenty-four Transformations used in this experiment changed some aspect of the actual action in such a way that although each was similar to the 96 slides which depicted the sequence, each differed from them according to one of four rules. Six Transformations were made for each of the four rules which were: (a) Addition: An object which might reasonably appear in such a setting but which has not been included was added. For example, the young man was shown with a briefcase beside him. (b) Token: Some change was made in the appearance of an actor. For example, the older woman was shown wearing a cardigan. (c) Perspective: A wide-angle lens (F2 28mm) was substituted for the F2 50mm lens. (d) Orientation: Duplicates of six of the 96 slides were made and reversed with respect to left-right orientation.

In photographing the action-sequence Transformations were spaced at approximately equal intervals over the course of the whole activity. The rule chosen when the time came to make a Transformation was determined randomly with the constraint that there should be six slides made for each rule. Addition, Token, and Perspective slides were made at the same time as the 96 slides depicting the activity. The action was 'frozen' at the appropriate moment, the change made, the photograph taken and then the original scene was restored.

Apparatus

Slides were projected onto a white wall by means of two Kodak Carousel SAV 2000 projectors which operated in direct succession. The projected image was .45 x .69m for both acquisition and test stimuli, and subjects sat approximately 1.5m from the screen. Each slide was shown for 4 seconds and there was no interstimulus interval. This was the case for both acquisition and test presentation. Exposure duration was determined by Gerbrands tachistoscopic shutters mounted in front of the projector lenses. Shutter operation was timed by solid state programming apparatus.

For the memory phase of the experiment, pen and paper were provided for subjects to write their accounts of the story. When they later performed the recognition memory test, electronic programming apparatus was used to record subjects' decisions and reaction times. Subjects had to press one of two mounted micro-switches on each trial to indicate their decision. The apparatus was programmed so that latencies of greater than 4 seconds were defined as a 'No Response' trial. The criterion for accepting a subject's data was set at no more than 10 No Response trials. For the rank-ordering (and rating) phase of the experiment, pen and paper were provided for judges to make notes.

Procedure

Randomisation: Four groups of evenly spaced Belonging slides could be extracted from the sequence of 96 slides: that is slide numbers 1, 5, 9, 13, ... etc. (Set A); slide numbers 2, 6, 10, 14, ... etc. (Set B); slide numbers 3, 7, 11, 15, ... etc. (Set C); or slide numbers 4, 8, 12, 16, ... etc. (Set D). There were, therefore, four sets of 72-frame action-sequences or Original slides each with a corresponding set of twenty-four Belongings.

In the Acquisition phase of the memory experiment Subjects 1, 5 and 9 in both the Ordered and Disordered Presentation groups saw Set A Originals; Subjects 2, 6 and 10 saw Set B; Subjects 3, 7 and 11 saw Set C; and Subjects 4, 8 and 12 saw Set D. Subjects in the Ordered Presentation condition saw slides in the correct chronological order and subjects in the Disordered Presentation condition saw them in a random order (the same random order for each set).

In the Test phase of the memory experiment, the seventy-two slides for each subject in both experimental groups consisted of twenty-four of the Original slides they had seen during acquisition, the corresponding set of twenty-four Belonging slides, and all twenty-four Transformation slides. Original slides for the test were chosen randomly without replacement so that, for example, in the case of Set A - Subjects 1, 5 and 9 in both groups - Subjects 1 saw twenty-four randomly selected Original slides, Subjects 5 saw twenty-four Original slides randomly selected from the remaining forty-eight, and Subjects 9 saw the remaining twenty-four. In this way, it was assured that during testing over the course of the experiment for each group each Original was sampled three times, each Belonging was sampled three times, and each Transformation slide was sampled twelve times. The seventy-two test slides were presented in a random order which was the same for corresponding subjects in both experimental conditions.

In the rank-ordering (and rating) phase of the experiment, judges saw all 96 action-sequence slides. Although this meant that they were not seeing the exact combination that any one subject had seen, it made it possible for the judges to rank-order and rate all twenty-four stories produced by the subjects in the memory phase of the experiment. Other methods, such as different judges for different sets, or the same judges seeing all four sets, presented considerable practical problems either in terms of procedure or subsequent data analysis. Therefore, although the method used was not ideal, it was considered to be the most practical.

Memory phase of the experiment: The memory phase of the experiment consisted of three parts - acquisition, story writing, and recognition test.

In the acquisition part, subjects were told that they were taking part in a memory experiment and were asked to try to remember the slides which were about to be shown to them. The 72 Originals were then presented.

Immediately after presentation, subjects were asked to write the story. The main body of the instructions were the same for both groups of subjects:

> 'Write down what happened in the story as simply but as accurately as possible. Be sure that you identify any characters you refer to in some way (for example,

by dress, or age, or hair, etc.). You have ten minutes in which to do this .

For the Ordered Presentation group this was prefaced by:
All of the slides together that you have just seen depicted a simple story. You saw them in their correct order. On the paper provided, try to describe the story depicted by the slides you have just seen'.

For the Disordered Presentation group the preface was:
'All of the slides together that you have just seen depicted a simple story. You saw them jumbled up. On the paper provided, try to describe the story depicted by the slides you have just seen, as if you had been shown them in their correct order''.

If subjects had not finished writing the story after ten minutes they were given extra time in blocks of two minutes. Three subjects each required an extra four minutes - two from the Ordered group and one from the Disordered group. Other subjects found ten minutes adequate.

After the subjects had written their story, they were then given a set of instructions explaining the recognition phase of the experiment. Subjects were asked to press one of the two buttons in front of them according to whether they thought they had seen the test slide before or not. They were asked to decide as quickly and as accurately as possible. They were requested to use their index fingers to respond and the decision represented by each button was counterbalanced across presentation groups. Half the subjects used the left button to indicate "seen before" and half used the right button to indicate "seen before". They were examined in their understanding of the instructions and were trained in the procedure before the test began. Test slides were then presented and subjects' responses were automatically recorded.

Rank-ordering and rating phase of the experiment: The rank-ordering and rating phase of the experiment took place after all data from the Memory phase had been collected. The stories written by the subjects were typed so that problems of individual handwriting would not interfere with assessment. The judges were given instructions which explained exactly what was going to occur (except that they were not forewarned of the rating task): They were told that they would be shown a slide sequence which depicted a simple story, after which they would be able to make notes, if they chose to do so, about what had occurred in the story. They would then be given a set of 24 descriptions of the story which they would be asked to read through once each in the order in which they were given them. (Each judge saw a different random order of the stories). Next they would be shown the slide sequence a second time, and again have the opportunity to make notes if they wished. When they were ready, they would be taken to another room where there would be a long table and plenty of space to sort out the stories. Their instructions would be to rank-order the stories so that the top story in their final selected pile was "the one which came closest to the story they had seen", the second one, the next close, and so on, so that the bottom story in the pile was the one which came least close to the story they had seen.

They were told that when viewing the slide sequences it was important for them to remember who did what in the story and in what order. They were advised to pay careful attention to this when asking themselves how close a description came to the story and that details of dress, or age, or hair, etc. were important only in as much as they were ways to identify the characters in the story. They would be given as much time as they needed to complete the rank-ordering task.

On average the entire rank-ordering phase of the experiment took each judge approximately two hours to complete. When they had finished the rank-ordering and the order was recorded by the experimenter, they were asked to assign a number between 1 and 10 to each story, where 10 indicated that it was extremely close to the original and 1 indicated that it was not at all close to the original.

Results

Rank-ordering and rating data: The rank-order assigned to each individual story by each of the twelve independent judges was averaged over the stories produced by each experimental group. The mean rank-order for the set of stories produced by the Ordered Presentation group was 7.83 (SD = 4.86), and for the stories produced by the Disordered Presentation group the mean was 17.17 (SD = 5.38), $t\ (23) = 6.20$, $p < .001$. The difference in mean ranks was considered to be convincing evidence that the judges decided that stories written by subjects who had seen an Ordered Presentation were closer to the original than stories written by the Disordered Presentation group.

If each of the 12 stories from the Ordered group had been assigned a rank-order between 1 and 12 and each of the 12 stories from the Disordered group had been assigned a rank-order between 13 and 24 (i.e. perfect division of groups on the dimension of 'closeness to the original') the mean rank-order for Ordered groups' stories would be 6.5 and the mean rank-order for Disordered groups' stories would be 18.5. Random assignment of rank-order (i.e. no difference between the two groups) would theoretically produce a mean rank-order for both groups which approached 12.5 as the number of trials increased. The actual mean ranks obtained were close to those associated with exact separation. The degree of inter-judge agreement was reasonably high, according to Kendall's Coefficient of Concordance which was W = .72.

The rank-ordering results, therefore, showed that there was a clear division between the two experimental groups. Although the Disordered Presentation group was specifically asked to re-order the randomly presented slides and attempt to reconstruct the underlying event they produced stories that were judged to be significantly less close to the actual story than stories written by the Ordered Presentation group.

It was possible that even if judges categorized stories from the Disordered Presentation group as relatively less close to the original than stories from the Ordered Presentation group, the Disordered Presentation groups' stories might in fact still be very close to the original. Therefore, the rating task provided supplementary data to assess whether this was the case.

The mean rating for stories from the Disordered Presentation group was 2.97 (SD = 1.54), significantly lower than that for stories from the Ordered Presentation group 6.27 (SD = 1.39), t (23) = 66.33, $p < .001$, where a score of 1 had indicated that the story was not at all close to the original and a score of 10 indicated extreme closeness. Therefore, not only were the two groups of stories distinguished by judges, but moreover the stories produced by the Disordered Presentation group were not close to the original story.

Recognition test data: The mean (and standard deviation) proportion of Hits and False Alarms for each Presentation condition are shown in Table 7. The Hit and False Alarm rates for each subject were converted into d' values and the means (and standard deviations) are given in Table 8.

TABLE 7

Mean (and standard deviation) Hit and False Alarm proportions for Originals, Belongings, and Transformations for Ordered and Disordered presentation conditions

Presentation Type	Test Slide Type		
	Original	Belonging	Transformation
Ordered	.80 (.10)	.68 (.15)	.15 (.08)
Disordered	.76 (.12)	.65 (.16)	.13 (.08)

TABLE 8

Mean (and standard deviation) d' scores for Originals to Belongings (O:B) and Originals to Transformations (O:T) for Ordered and Disordered presentation conditions

Presentation Type	Distractor Type	
	O:B	O:T
Ordered	.40 (.52)	1.99 (.43)
Disordered	.33 (.29)	2.03 (.74)

An ANOVA which examined Presentation Type (between-subject) and Distractor Type (within-subject) showed no effect of Presentation but a strong effect of Distractor, $\underline{F}$ (1,22) = 141.04, MSe = .23, $\underline{p}$ < .001 with d'O:T higher than d'O:B. There was no interaction.

Reaction time data are presented in Table 9 where the means (and standard deviations) are shown separately for "Yes" and "No" responses. Since "No" RTs tend to be slower than "Yes" RTs, Yes and No responses were treated separately in order to avoid confounding of the experimental variables with response type. Similarly, correct and incorrect responses were analysed separately. It was, therefore, not possible to compare RTs for all three types of test slide together. Accordingly, analysis of the reaction time data was restricted to the following: Two ANOVAs examined RTs to distractors separately for "Yes" and for "No" responses; and two $\underline{t}$-tests examined RTs to Originals separately for "Yes" and "No" responses.

The data were adjusted for missing cells (Kirk, 1966, pp. 146-147) since one subject in the Disordered Presentation condition correctly rejected all Transformations and one subject in the Disordered Presentation condition correctly accepted all Originals and falsely accepted all Belongings.

An ANOVA of the "Yes" (Incorrect) RT data which examined Presentation Type (between-subject) and Distractor Type (within-subject) showed no main effects and no interaction. When subjects falsely accepted distractors there was no significant difference

TABLE 9

in RTs between Belongings and Transformations. There was also no difference between the Ordered Presentation and Disordered Presentation conditions.

A similar analysis was made of the "No" (Correct) RT data. There was no effect of Presentation but a strong effect of Distractor, F (1,21) = 42.89, MSe = 63143, p < .001 with Transformations rejected more quickly than Belongings. There was no interaction.

For "No" (Incorrect) responses to Originals there was no significant difference between the mean RTs of Ordered Presentation and Disordered Presentation subjects (t (22) < 1). Similarly there was no significant difference between the mean "Yes" (Correct) RTs of Ordered Presentation and Disordered Presentation subjects (t (21) < 1).

Discussion

The stimulus material used in this experiment could not be reorganised successfully when presented in a disorganised sequence. Even when specifically asked to try to reconstruct the disordered slides into the underlying ordered story subjects were unable to do so. Independent judges ranked stories written by the Disordered Presentation subjects as less close to the actual story than stories written by subjects from the Ordered Presentation group and rated them not close to the original story. Although these findings are not consistent with the majority of linguistic reconstruction experiments, it must be pointed out that this material did not have a familiar and strong thematic structure (Van Dijk, 1977) which could be reconstructed easily. However, it should also be noted that subjects in the Ordered Presentation condition were able to describe the underlying story quite accurately.

The main effect reported by Jenkins et al. (1978) and demonstrated in the previous three experiments in this study was replicated with the long slide sequence. But the crucial test offered by comparison of Ordered versus Disordered Presentations showed that the argument that Belonging slides are falsely accepted because of their consistency with the abstracted event does not seem to be tenable.

Despite the fact that subjects were *unable* to reconstruct the underlying story when given a disordered presentation of acquisition slides, the d'O:B values for this group were significantly lower than the d'O:T values and their recognition performance did not differ from that of the subjects who saw ordered acquisition slides. If subjects are not able to form a representation of the event in the first place then Belongings do not belong to any event and false acceptance of Belongings cannot be due to fusion. It is more likely, as the earlier experiments indicated, that Transformations are more frequently rejected on a recognition test than Belongings, not because they are inconsistent with the meaning of the event seen during acquisition or with the relationships which define the event, but because they have fewer features in common with Originals than do Belongings.

Reaction time data analysis showed that not only were Transformations more frequently rejected than Belongings, they were also rejected more quickly. However, although Belongings were more frequently accepted, they were not accepted more quickly than were Transformations.

These results support a featural interpretation of the d' data and are most easily explained in terms of a memory-scanning model of recognition memory. With such models, the subject is described as comparing a test item to items stored in memory and the subject's reaction time on a given test item is thought to reflect the number of comparisons made, the time required for each comparison, and the extent to which comparisons may be made simultaneously. Although the times involved can be assessed only approximately, a tentative explanation can be offered, along the same lines as Biederman's interpretation of RT data in his contextual studies.

When the subject views an individual test slide, if there is no diagnostic feature in the slide which strikes him as new, then a search may take place either until a match with an Original slide is found, or a mis-match is made, or until the search is terminated without finding a match. It is also possible that these processes operate consecutively - a vigilance for new features and an automatic search for a match, the search being terminated if a new feature is spotted or when the matching process is resolved (either correctly or incorrectly).

The Transformations have fewer features in common with Originals than do Belongings. Each of the four types of Transformations offers a potentially new feature - whether it be the addition of an object, a token change in dress, a left-right orientation reversal, or a perspective change. New features in the Belonging slides would not be available to the same extent, since they are the same as Originals in terms of the basic components of the picture, being different primarily in terms of composition. Belongings and Originals would also, therefore, tend to have the same degree of specific detail and general familiarity, whereas with the introduction of a featural change, Transformations would have greater specific information (Loftus & Bell, 1975). In the case of the correct rejections of the distractors, where Transformations were rejected more quickly than Belongings, it seems likely that a new feature was spotted in the Transformations which resulted in faster rejection. However, in the case of Belongings, the search-for-match process would take precedence and the search time involved would account for the longer RTs. When the subject falsely accepts a distractor slide, it is presumably on the basis of a match (or a decision in favour of that). So when the subject accepts Transformations and Belongings it is as a result of the same process - and the RTs are not significantly different.

EXPERIMENT V

Introduction

In the previous experiment, it was demonstrated that subjects who saw a set of disordered slides during acquisition were not able to reconstruct those slides to form the event they depicted when presented in their correct chronological order. A subsequent recognition memory test had shown that with this material the performance of subjects who had seen a disordered version did not differ from that of subjects who had seen the ordered sequence. It was concluded that the fusion hypothesis was not supported by the results which were accounted for by the relative featural detail of the two types of distractors. The aim of the present experiment was to confirm that the results of the recognition memory test in the previous experiment were not influenced by the intervening writing task.

Apart from the absence of the writing task, this experiment was an exact replication in terms of design, method, randomisation, and subject pool. It was assumed, therefore, without further demonstration, that subjects who saw the ordered acquisition slides were able to form a coherent story consistent with the one actually depicted, whereas subjects who saw the disordered version were not. This experiment was a more direct test of the fusion interpretation in that subjects were given the recognition test immediately after acquisition.

Consistent with the previous rationale, it is argued that if the differential performance on distractors is due to fusion then the difference between d'O:B and d'O:T should be greater for the Ordered Presentation than for Disordered Presentation. If performance is determined by the featural context of the slides then there should be no difference in performance between the two groups.

Having established in the previous experiment that subjects were not able to reconstruct this pictorial sequence, it was decided to make a specific examination of rejection rates to the four Transformation Types in Experiment V. Kraft and Jenkins (1977) found that orientation reversals were detected easily when subjects had been presented with acquisition slides in a consistent way with respect to orientation. They varied not only whether the presentation was consistent with respect to orientation, but also whether it was consistent with respect to temporal (or chronological) order. They found that recognition of orientation reversal was high regardless of whether subjects had seen slide sequences in an ordered or disordered version but argued that subjects "in the random order context condition were able to construct coherent events even though the slides were in random order" (p. 399). The present experiment, therefore, tested whether order of presentation affects identification of orientation reversal distractors. In this experiment all subjects saw acquisition slides in a consistent left-to-right orientation, but temporal order was varied.

Method

Subjects

Subjects were 22 women and two men. Their ages ranged from 17 to 41 years with a mean age of 24.67 and a standard deviation of 7.89. The design was a 2 (Ordered versus Disordered) X 2 (Belonging versus Transformation) factorial with repeated measures on the second factor. Twelve subjects were randomly assigned to each of the two groups. Four subjects were dropped from this experiment - two because of equipment failure, and two because they did not reach the pre-determined criterion of no more than 10 no response trials. These subjects were replaced. Subjects took part individually and all instructions were written.

Procedure

Stimulus material, apparatus, and randomisation were the same as in Experiment IV. Similarly, procedure was the same except that subjects were not given the story-writing task. Other instructions were identical to those used in Experiment IV.

Results

The same analyses were performed on the data from this experiment as on the data from Experiment IV. Additional analyses were performed where the two experiments (IV and V) were treated as a between-subject factor. The mean (and standard deviation) proportion of Hits and False Alarms for each Presentation condition are shown in Table 10. The Hit and False Alarm rates for each subject were converted into d' values and the means (and standard deviations) are given in Table 11.

An ANOVA which examined Presentation Type (between-subject) and Distractor Type (within-subject) showed no effect of Presentation but a strong effect of Distractor, $\underline{F}$ (1,22) = 85.05, MSe = .39, $\underline{p} < .001$ with d'O:T higher than d'O:B. There was no interaction. An ANOVA which examined Experiment IV versus V (between-subject factor), Presentation Type, and Distractor Type, showed a reliable effect of Experiment, $\underline{F}$ (1,44) = 7.39, MSe = .57, $\underline{p} < .01$, no effect of Presentation, a strong effect of Distractor, $\underline{F}$ (1,44) = 211.65, MSe = .31, $\underline{p} < .001$, and no interactions. Subjects in the present experiment performed with greater accuracy than subjects in the previous experiment. The mean d's for the Ordered condition were 1.71 and 1.19 for Experiments V and IV respectively, and for the Disordered condition, 1.50 and 1.18 for Experiments V and IV respectively.

Reaction time data are presented in Table 12 where the means (and standard deviations) are shown separately for "Yes" and "No"

TABLE 10

Mean (and standard deviation) Hit and False Alarm proportions for Originals, Belongings, and Transformations for Ordered and Disordered presentation conditions

Presentation Type	Test Slide Type		
	Original	Belonging	Transformation
Ordered	.86 (.13)	.67 (.20)	.14 (.13)
Disordered	.79 (.13)	.58 (.11)	.12 (.12)

TABLE 11

Mean (and standard deviation) d' scores for Originals to Belongings (O:B) and Originals to Transformations (O:T) for Ordered and Disordered presentation conditions

Presentation Type	Distractor Type	
	O:B	O:T
Ordered	.83 (.39)	2.59 (1.05)
Disordered	.72 (.53)	2.29 (.94)

TABLE 12

Mean (and standard deviation) RTs to Originals, Belongings and Transformations for Ordered and Disordered presentation conditions

Presentation Type	Test Slide Type		
	Original	Belonging	Transformation
Given that the subject responded "YES"			
Ordered	1823 (440)	1996 (519)	1892 (555)
Disordered	1620 (424)	1690 (431)	1877 (607)
Given that the subject responded "NO"			
Ordered	2186 (516)	2075 (581)	1578 (377)
Disordered	1943 (481)	1888 (504)	1433 (343)

responses. Two ANOVAs examined Presentation Type (between-subject factor) and Distractor Type (within-subject factor) for "Yes" (Incorrect) RTs, and separately for "No" (Correct) RTs. Two t-tests examined Presentation Type for "Yes" (Correct) RTs, and separately for "No" (Incorrect) RTs. The RT data were adjusted for missing cells (Kirk, 1966, pp. 146-147) since two Ordered Presentation and two Disordered Presentation subjects correctly rejected all Transformations and three Ordered Presentation and one Disordered Presentation subject correctly accepted all Originals and falsely accepted all Belongings.

An ANOVA on the "Yes" (Incorrect) RTs showed no effect of Presentation or Distractor and no interaction. When subjects falsely accepted distractors, there was no significant difference in RTs between Belongings and Transformations. There was also no difference between the Ordered Presentation and Disordered Presentation conditions. This result confirmed the finding in Experiment IV. An ANOVA which examined Experiment IV versus Experiment V (a between-subject factor), Presentation Type, and Distractor Type, showed no main effects and no interactions. Therefore, the two experiments were not significantly different with respect to "Yes" RT data for distractors.

Similar analyses were made of the "No" (Correct) RTs. Again the result confirmed the findings of Experiment IV. There was no effect of Presentation but a strong effect of Distractor, $\underline{F}\,(1,22) = 33.07$, MSe = 82227, $\underline{p} < .001$ with Transformations rejected more quickly than Belongings. There was no interaction. An ANOVA which introduced the two experiments as an additional between-subject factor, showed no effect of Experiment, no effect of Presentation, and a strong effect of Distractor, F (1,43) = 74.45, MSe = 72907, $p < .001$ (Transformation being rejected faster than Belongings). There were no interactions.

For "No" (Incorrect) responses to Originals, there was no significant difference between the mean RTs of Ordered Presentation and Disordered Presentation subjects, (t (18) < 1). An ANOVA which examined Experiments IV versus V (a between-subject factor) and Presentation Type (a between-subject factor) showed no main effects and no interaction.

Similarly there was no significant difference between the mean "Yes" (Correct) RTs of Ordered Presentation and Disordered Presentation subjects, (t (22) < 1). An ANOVA which introduced Experiment IV versus V as a between-subject factor showed no effects of either Experiment or Presentation and no interaction.

Responses to the Transformations - of which there were four types, Addition, Orientation, Perspective, and Token - were analysed in greater detail. The mean (and standard deviation) proportion correct responses are given in Table 13.

An ANOVA which examined Presentation Type (between-subject) and Transformation Type (within-subject), showed no effect of presentation, but a significant effect of type of Transformation, F (3,66) = 5.90, MSe = 268, $p < .01$. There was no interaction. A subsequent Sheffé test showed that Orientation slides were significantly better detected than Token and Addition, $p < .001$.

TABLE 13

Mean (and standard deviation) proportion of correct rejections of Addition, Orientation, Perspective and Token Transformation types for Ordered and Disordered presentation conditions

Presentation Type	Transformation Type			
	Addition	Orientation	Perspective	Token
Ordered	.85 (.13)	.93 (.13)	.94 (.11)	.75 (.26)
Disordered	.76 (.24)	.98 (.06)	.90 (.25)	.85 (.18)

Perspective slides were better detected than Token and Addition, $p < .01$. There was no significant difference between Perspective and Orientation, and none between Token and Addition.

Discussion

The main effect reported by Jenkins et al. (1978) and demonstrated in the previous experiments in this series was again replicated. The d'O:T values were significantly higher than d'O:B values. Moreover, as in Experiment IV, performance did not differ between the two experimental groups. Despite seeing a disordered story (which had been shown in Experiment IV to be not re-orderable) subjects exhibited d' values for both O:B and O:T which did not differ from those of subjects who had seen an ordered presentation.

The difference between the two experiments was that in Experiment IV subjects had a writing task before their recognition test which introduced a delay factor of about 14 minutes and which effectively represented a recall test. Since the third experiment in this series had shown that performance was not affected by a delay of 24 hours, it seemed unlikely that delay was a significant factor. It would appear, therefore, to be the task itself which impaired performance in Experiment IV compared to the present experiment. It is likely that the nature of the task (or possibly simply being given a task) produced interference. It has been argued that memory for surface form is impaired when subjects are given a non-featural task to perform (Hunt & Elliott, 1980). The writing task in Experiment IV required subjects to form a verbal and predominantly semantic account of the presentation material, whereas the subsequent recognition test examined memory for discrete pictorial items.

The results of the RT analyses confirmed the d' findings - that there was no difference in performance between subjects who saw slides in an Ordered Presentation and subjects who saw a Disordered Presentation. The distractor RT analysis adds weight to the d' analysis - not only were Transformations rejected more frequently than Belongings, they were rejected more quickly. However, although Belongings were accepted more frequently than Transformations, they were not accepted more quickly. An interpretation of this result was given in the discussion section of the previous experiment.

The finding that Orientation slides were the most frequently detected Transformations independently of presentation order is congruent with Kraft and Jenkins (1977). However, although Kraft and Jenkins' Disordered Presentation group may have reconstructed the event, subjects in the present experiment could not. It seems likely, therefore, that a major factor in the ease with which subjects detect Orientation reversals is consistency of orientation during presentation.

In summary, this experiment, which was a replication of Experiment IV (with the exception that there was no recall test) showed that there was no difference in response pattern between subjects who saw an ordered presentation of the stimulus material and subjects who saw a disordered presentation. Since subjects in Experiment IV, who saw a disordered version were unable to reconstruct the story (and it may be assumed that the same was true for subjects in the present experiment), it can be concluded that fusion was not responsible for the low d'O:B compared to d'O:T, since Disordered Presentation subjects could not have abstracted the underlying event.

EXPERIMENT VI

Introduction

Previous experiments in this series had found no support for Jenkin's fusion interpretation of differential detection of distractor types, despite systematic tests of his hypothesis. However, neither had specific evidence been found for the alternative hypothesis which argued that d'O:B values were low compared to d'O:T values because Belongings had more features in common with Originals than did Transformations, and were, therefore, more confusable. A more direct attempt was now made to show that the surface structure of stimulus material is a major determinant of recognition memory.

The main problem to overcome in examining the effects of the meaning and features of stimulus material is to achieve a separation of the variables with no confounding. Following the procedure of Jenkins et al. (1978), the main way in which meaning was varied while holding physical features constant was to include Transformations, the meaning of which was inconsistent with that of the event by virtue of the changed detail. But as it has been argued here, that change also represented a change in physical features. However, the task of holding meaning constant while varying physical features was equally problematic since any change in features may also produce a change in meaning. Although some research has produced this kind of distinction with scaled material (Mandler & Stein, 1974; Posner, 1969) or arbitrarily labelled material (Rafnel &

Klatzky, 1978), systematic distinction and quantification of the variables did not seem possible with complex pictorial stimuli involving many featural details, colours, and spatial and temporal relationships.

Despite these difficulties, it was considered to be essential to attempt to vary features whilst holding meaning constant, if only by an approximation, to find more specific evidence of a featural account. This was done in the final experiment by excluding Transformations and varying the featural similarity of Belongings to Originals.

By definition, the meaning of Belongings is consistent with the overall event, and their meaning is therefore held constant. On the assumption that featural similarity between a single pair of slides would vary inversely with their chronological distance in the story sequence, featural similarity of Belongings was manipulated by choosing sets of Belongings which were immediately adjacent to Original slides of the acquisition set and sets of Belongings which were nonadjacent to the Originals. By selecting different 24-slide sequences of Originals and presenting all 96-slides in the recognition test, the variable of adjacency was examined within-subject, and adjacent 'Near' slides for some subjects were nonadjacent 'Far' for others and vice versa.

According to Jenkins et al. (1978) notion of fusion, d' measures for the discrimination between Originals and Near Belongings (d'O:N) should not differ from d's for the discrimination between Originals and Far Belongings (d'O:F). But if recognition performance is based on the featural similarity between Belongings and Originals, d' (O:F) should be greater than d' (O:N).

Method

Subjects

Subjects were six women and ten men. Their ages ranged from 18 to 51 years with a mean age of 25.63 and a standard deviation of 11.77. The experiment was a within-subject design so that all subjects served in both experimental (Near and Far distractor) conditions.

Stimulus Material and Apparatus

The stimulus material was the same as that used in Experiments IV and V except that Transformations were not used. The apparatus was also the same.

Procedure

Four sets of evenly spaced Originals could be withdrawn from the overall sequence of 96 slides: that is slide numbers 1, 5, 9, 13 ... etc. (Set A); numbers 2, 6, 10, 14 ... etc. (Set B);

numbers 3, 7, 11, 15 ... etc. (Set C); and numbers 4, 8, 12, 16 ... etc. (Set D). There were therefore four sets of 24-frame action-sequences of Originals each with a corresponding set of Belongings. From the remaining set of 72 Belongings there were 24 Far slides for each set of Originals and 48 Near slides - half of which occurred before and half of which occurred after the related Original slide. Half the subjects saw Nears which occurred before the Originals and half the subjects saw Nears which occurred after the Originals. Thus the design was counterbalanced so that every slide was used an equal number of times as an Original, Near and Far. Subjects 1, 2, 9, and 10 saw Set A in acquisition (where on testing subjects 1 and 2 saw Nears which occurred before Originals and subjects 9 and 10 saw Nears which occurred after Originals). Subjects 3, 4, 11, and 12 saw Set B; subjects 5, 6, 13, and 14 saw Set C; and subjects 7, 8, 15, and 16 saw Set D.

In the data analysis, slide numbers 1, 2, 3, 4, and numbers 93, 94, 95, 96 were not included, since with Set A (where slide number 1 was an Original) there was no Near which occurred before the Original and also with this set slide number 96 was not a valid Near. Similarly with Set D (where slide number 96 was an Original) there was no Near which occurred after that original and also with this set, slide number 1 was not a valid Near. Therefore, although during the course of the experiment all slides were tested, responses to the first and final four slides from the set of 96 were not included in analysis. Thus on the recognition test, subjects were scored out of 22 responses for each kind of slide - Originals, Nears, and Fars. In acquisition, slides were shown in their correct chronological order and in the test were shown in different random orders for each subject.

In the acquisition phase of the experiment subjects were told that they were taking part in a memory experiment and were asked to try to remember the slides which were about to be shown to them. The 24 Original slides were then presented. As in the previous experiments, slides were shown for 4 seconds each, with no inter-stimulus interval.

Immediately after presentation, subjects were given the same recognition memory test instructions as in Experiments IV and V. Subjects' decision and reaction time were recorded automatically for each of the 72 test slides.

Results

The mean (and standard deviation) proportion Hits for Originals was .75 (.14) and False Alarms for Nears was .45 (.19) and for Fars was .41 (.17). The mean (and standard deviation) d' scores for Originals to Nears (d'O:N) was .89 (.41) and for Originals to Fars (d'O:F) was .96 (.36). These d' scores were not significantly different (t (15) < 1). Reaction time data are presented in Table 14. As in previous experiments RT data were analysed separately for the "Yes" and "No" responses, and for the

TABLE 14

Mean (and standard deviation) RTs to Originals, Nears and Fars

Test Slide Type		
Original	Near	Far
Given that the subject responded "YES"		
1880 (293)	2059 (539)	2098 (468)
Given that the subject responded "NO"		
2154 (439)	2035 (408)	1970 (300)

Correct and Incorrect responses. There was no significant difference in either "Yes" or "No" RTs for Nears versus Fars (in both cases $\underline{t}$ (15) < 1).

Discussion

No differences in responses to Nears and Fars were found with either d' or RT measures. If the two types of distractor did differ with respect to similarity to Originals, then this did not affect performance.

It was possible that with the material used, chronological distance did not in fact vary featural similarity. Although the sequence of slides which depicted the story was progressive and continuous, the action within these slides had not been designed specifically to ensure that a slide which was Far from a given original was necessarily less similar to it than one which was Near. However, further experiments along the lines suggested in this final experiment, but with material which systematically varied featural similarity by chronological distance within a consistent event,might establish more explicity the role of featural detail in memory for picture stories.

GENERAL DISCUSSION

The experiments in this thesis investigated a fusion account of recognition memory for pictorial sequences. Jenkins et al. (1978) had found that after viewing temporally ordered picture stories, subjects tended to discriminate well between original slides and Transformation distractors which were inconsistent with the story, compared to their ability to discriminate between Originals and Belonging distractors which were consistent with the story. Jenkins interpreted the differential response to distractor types as evidence of *fusion*, where recognition performance was based on the comparison of new slides to the abstracted schema of the total event depicted in acquisition. He had defended this interpretation against an alternative account which argues that recognition performance is determined by the relative featural similarity between distractor types and original slides.

This thesis does not dispute the influence of the overall meaning of stimulus material on memory, nor the particular phenomenon of abstraction of the relationships between a series of related items presented over time. However, it argues that the featural composition of stimulus material should be an integral part of any interpretation of recognition performance unless the effect of featural similarity is explicity controlled.

It is clear that both the meaning and features of stimulus material contributes to performance in studies of perception and memory. Numerous studies provide converging evidence that meaning is a crucial factor in determining the degree of memory accuracy and durability (Bower & Karlin, 1974; Craik & Lockhart, 1972; Kintsch, 1974; Mistler-Lachman, 1974; Sachs, 1967, 1974). However, there is also evidence that the degree of featural similarity between distractors and acquisition material in a recognition memory paradigm affects the degree of recognition accuracy (Baggett, 1975; Goldstein & Chance, 1970; Hunt & Elliott, 1980; Weaver & Stanny, 1978).

It is also evident that the variables of features and meaning interact in determining performance. Biederman demonstrated that object identification and scene recognition were more accurate when scenes were presented in a meaningful, organised way. He also demonstrated that in labelling tasks of briefly presented scenes, identification of discrete objects in the scene influenced perceptual decisions. Extensions of Biederman's work by Palmer and others confirmed that both meaningful context and the specific featural elements of objects in scenes are equally relevant to perceptual processing. Interpretations of data proved to be inadequate without incorporating an interaction of context and specific features in the account. Mandler's studies of memory for pictorial scenes showed that both the discrete objects in a scene and their spatial relationships were relevant to recognition performance.

The contention of this thesis is that memory is determined by an interaction of the meaning and features of stimulus material, but in order to examine the nature of this interaction the two main variables must be distinguishable. Such distinction must be explicit to the experimenter, if not to the subject. When the distinction is not explicit, interpretation of the data may be misleading. This has been the case particularly with abstraction paradigms which rely on recognition memory performance to demonstrate the integration of relationships between discrete elements in stimulus material.

When distractors vary both memory and features relative to the acquisition material, apparent apprehension of the total meaning of the acquisition set can also be explained in terms of the featural structure of the stimulus material. Bransford and Franks' classic linguistic abstraction studies have been criticised precisely on these grounds and the present thesis has criticised Jenkins' pictorial abstraction studies on the same grounds.

It was argued that in the fusion experiments reported by Jenkins, et al. (1978), distractors which were consistent with the story depicted by Original slides also had more features in common with Originals than did the inconsistent distractors. The variables of semantic consistency and featural similarity were therefore confounded.

The methodological difficulty of demonstrating temporal integration through recognition memory performance is considerable. At some stage during the procedure there must be an effective separation of the meaning and features of materials. Ideally, in a recognition paradigm where inferences are made on the basis of responses to types of distractors this should be achieved by varying the test material directly along the dimension of meaning whilst holding featural properties constant, or vice versa. Other procedures for separating the crucial variables indirectly are possible through manipulation of *the acquisition material, the acquisition conditions* or *the test conditions*. Because of the inherent difficulties in achieving a direct separation of meaning and features with complex pictorial material, the present research used each of the indirect procedures before finally attempting a direct manipulation of the test material.

The first experiment replicated Jenkins' findings of differential discrimination of Belongings and Transformations. It also tested the fusion interpretation of these results by manipulating the *acquisition material* in an attempt to keep featural properties constant whilst varying the overall semantic structure of the stimulus material. Performance following Ordered Presentation of acquisition material was compared with two Disordered Presentation conditions - disordering of six separate slide sequences (the Disordered-Within condition) and disordering of all six sequences together (the Disordered-Within-and-Between condition). Despite evidence in the second experiment that the manipulation produced the intended effect of disrupting semantic structure, d'O:B was significantly lower than d'O:T in the Disordered Presentation conditions, and the d'O:B measure relative to d'O:T did not differ across presentation conditions.

The second experiment tested fusion by manipulating the *acquisition task conditions* so that the semantic and featural properties of the material might be separated during acquisition processing. Within an incidental learning paradigm, some subjects were given a semantic task designed to facilitate fusion whereas others were given a non-semantic task designed to inhibit fusion. Although the semantic task group gave a more accurate performance overall, d'O:B was significantly lower than d'O:T in both experimental conditions, and the degree of difference between the two d' measures was not affected by the type of acquisition task.

A third experiment tested fusion by manipulating the *test conditions* so that the contribution of the semantic and featural properties of the material to memory might be separated by performance impairment induced by delayed testing. Although memory was significantly impaired after a retention interval of one week, both d'O:B and d'O:T were reduced - a result inconsistent with a fusion account.

The results of the first three experiments, therefore, did not support a fusion account of the low d'O:B compared to d'O:T measure. In both experimental procedures where a fusion interpretation would have been supported by an attenuation of the response pattern, this did not occur - manipulation of the acquisition material by varying presentation order did not affect the response and nor did manipulation of the acquisition conditions where varying orienting tasks did not produce the expected difference in response patterns. The manipulation of test conditions in the third experiment did not produce the more pronounced pattern at impairment which would have been consistent with a fusion interpretation. Apart from examining the specific question of fusion, the first three experiments relate to more general questions concerning memory for pictorial material, by their manipulation of the variables of organisation, incidental learning tasks, and retention interval.

The effects of the incidental learning task and retention interval variations were consistent with related literature. The superiority of memory following semantic processing compared to non-semantic processing during acquisition was demonstrated convincingly with temporally ordered picture stories where the semantic task was relevant to the logical structure of the story and the non-semantic task was related to a gross featural characteristic of individual slides. These results were consistent with the majority of levels-of-processing studies (Bower & Karlin, 1974; Craik & Lockhart, 1972; Craik & Tulving, 1975; Mistler-Lachman, 1974). A more detailed analysis of the results of the second experiment also supported studies which have shown that the degree to which material can be processed in depth affects performance.

A retention interval of one day did not impair recognition performance compared to immediate testing, whereas a one week delay significantly reduced accuracy. These findings were consistent with other studies which have used pictorial material of a homogeneous nature. Baggett (1975) found that memory for the featural details of simple cartoon stories was not impaired by a 3-day retention interval. Mandler and Parker (1976) reported impairment of accuracy in memory for descriptive information about pictorial scenes when testing was delayed for a week.

The variable which has been not only the least well explored with complex picture sequences but also the most central to the test of fusion in the present studies is organisation. A generally accepted principle has been that organisation of to-be-remembered material increases the accuracy and duration of memory (Kintsch, 1977; Tulving, 1972). Classic experiments demonstrated the power of this principle with category organisation (Bousefield, Cohen & Whitmarsh, 1958) and associative organisation (Jenkins & Russell, 1952). In the absence of experimenter imposed organisation, subjects themselves tend to organise stimulus material in order to enhance memory (Mandler, 1972). Organisation manipulations are primarily effective with recall procedures (Kintsch, 1970), although Mandler (1972) has argued that recognition performance can be improved if sufficiently strong manipulations of organisation are used.

The effect of organisation on memory for pictorial material is less well documented. It appears, in general, that spatial organisation of pictorial material is an effective variable, but its effects depend on the particular task given to the subject and the specific material used. Using photographic material of complex every-day scenes, Biederman showed that accuracy is impaired and latency increased in object identification tasks and labelling tasks, given briefly presented spatially disorganised as opposed to organised scenes. Mandler's studies used simpler hand-drawn scenes which were presented for several seconds, and assessment of performance was in the context of recognition memory tests. In general she found that differences in recognition accuracy given spatially-disorganised compared to organised scenes depended on the type of information elicited by the recognition task. Memory for the identity or figurative appearance of objects in a scene was unaffected by disorganised presentation. Disorganisation impaired memory for the spatial relationships between objects but enhanced memory for the position or presence of any single object.

The effects of temporal disorganisation of series of related pictures seem to be elusive. Although Jenkins et al. (1978) reported better discrimination of Belongings given a disordered compared to ordered version of one of his photographic sequences of every-day events, other studies have not found temporal organisation to be an effective variable (Allen, Siegel & Rosinski, 1978; Kraft & Jenkins, 1977; Pittenger & Jenkins, 1979).

Nor did the first experiment in the present research obtain any effect of organisation. Data from the second experiment suggested that this was because subjects were not basing their recognition responses primarily on the temporal structure of the material. The disordering procedure used in the first two experiments was designed to vary the *degree* of disruption of semantic structure, and acquisition data from the second experiment supported this assumption when subjects' judgements of the logical order of slides proved to be systematically related to the intended degrees of disruption. Moreover, when subjects were explicity directed to attend to the temporal structure of the acquisition material during processing they gave less accurate recognition performances in the most Disordered Presentation condition than in the Ordered condition, whereas subjects who were directed to attend to the surface structure of the acquisition material were not affected by order of presentation.

The absence of an effect of temporal disordering has tended to be explained by arguments that subjects spontaneously reconstruct the disordered sequences. Although Baker (1978) did not find subjective re-organisation of temporal relationships in simple verbal stories, there has been strong evidence that subjects re-order disordered stories into their correct schematic order (Bower, Black & Turner, 1979; Kintsch, Mandel & Kozminsky, 1977; Mandler, 1978; Stein & Glenn, 1978; Stein & Nezworski, 1978). The reconstruction explanation of absence of temporal organisation effects was crucial to the present research, since if subjects reconstruct disordered stories, disordering is not a critical test of fusion. The fourth and fifth experiments therefore examined the variable of temporal organisation more directly.

The fourth experiment specifically attempted to establish that subjects who saw a disordered presentation of acquisition material were *not able* to reconstruct the underlying event. New material was designed to minimise the possibility of reconstruction. This long action-sequence was shown to subjects either in its correct chronological order or in a random order. They were asked to provide written accounts of the slide sets and attempt to describe the underlying event. Independent judges later decided that the stories produced by the Disordered Presentation group were not close to the actual event and were significantly less so than the accounts produced by the Ordered Presentation group. This experiment principally demonstrated that the story in this particular pictorial sequence could not be reconstructed from a disordered presentation. It also showed that disordering did not impair overall recognition performance, and more particularly that significantly low d'O:B values compared to d'O:T values were no different for the ordered and disordered groups. Moreover reaction time measures also showed no effect of order.

A fifth experiment confirmed the recognition memory results of the previous experiment without the inclusion of the recall-writing task. There was no evidence of an effect of temporal organisation, and no support for the fusion account of recognition memory of pictorial sequences.

The results of the experiments which attempted to separate meaning from features by indirect methods in order to examine the fusion hypothesis, did not support a fusion account and were consistent with a featural account of this type of recognition performance. The final experiment therefore attempted to vary the *test material* directly along the dimension of featural similarity to acquisition material, whilst holding the meaning of the individual test slides consistent with the overall event shown during acquisition. Although this experiment was unsuccessful, it was suggested that with material designed explicitly for the purpose, it might be possible to examine variations in featural similarity of distractors via their chronological distance from an ordered acquisition set, whilst holding overall meaning constant.

In summary, the present research did not find any evidence to suggest that recognition memory for pictorial sequences is determined by a representation of the total event depicted by the sequence. The main results indicated that recognition performance is determined by the degree of featural similarity between distractors and original slides, and this effect can be demonstrated in procedures which vary acquisition material, acquisition task, and retention interval. A conclusive demonstration of the effect of featural similarity requires systematic variation of features without a confounding of the variable of meaning. The present research attempted this demonstration and suggested that with complex photograph slide sequences of every-day events, further research could be directed towards examining the effects of featural similarity by developing stimulus material which varied magnitude of featural change in relation to chronological distance.

APPENDIX A: SUMMARY TABLES OF ANOVAS

This appendix presents summary tables of all ANOVAS referred to in the experiments. They are listed in the order in which they appeared and are referenced by Experiment (and page number).

Summary of ANOVA relating to data in Experiment I (p.77)

Source	SS	df	MS	F
Presentation Type	.29	2	.15	.47
S/Presentation Type	6.67	21	.32	
Distractor Type	3.87	1	3.87	35.18
Presentation Type x Distractor Type	.31	2	.16	1.45
Distractor Type x S/Presentation Type	2.38	21	.11	

Summary of ANOVA relating to data in Experiment I (p.81)

Source	SS	df	MS	F
Presentation Type	.10	1	.10	.38
S/Presentation Type	5.39	21	.26	
Distractor Type	2.17	1	2.17	31
Presentation Type x Distractor Type	.24	1	.24	3.43
Distractor Type x S/Presentation Type	1.56	21	.07	

Summary of ANOVA relating to Experiment II (p.85)

Source	SS	df	MS	F
Presentation Type	1.64	2	.82	2.56
Task Type	3.63	1	3.63	11.34
Presentation Type x Task Type	.78	2	.39	1.22
S/Presentation Type x Task Type	13.34	42	.32	
Distractor Type	9.60	1	9.60	137.14
Presentation Type x Distractor Type	.07	2	.04	.57
Task Type x Distractor Type	.11	1	.11	1.57
Presentation Type x Task Type x Distractor Type	.23	2	.12	1.71
Distractor Type x S/Presentation Type x Task Type	3.01	42	.07	

Summary of ANOVA relating to Experiment II (p.88)

Source	SS	df	MS	F
Presentation Type	9300	2	4650	123.93
S/Presentation Type	788	21	37.52	

Summary of ANOVA relating to data in Experiment II (p.91)

Source	SS	df	MS	F
Presentation Type	2.34	2	1.17	3.55
S/Presentation Type	6.82	21	.33	
Distractor Type	3.85	1	3.85	96.25
Presentation Type x Distractor Type	.04	2	.02	.05
S/Presentation Type x Distractor Type	.92	21	.04	

Summary of ANOVA relating to data in Experiment II (p.91)

Source	SS	df	MS	F
Presentation Type	.08	2	.04	.13
S/Presentation Type	6.51	21	.31	
Distractor Type	5.86	1	5.86	58.60
Presentation Type x Distractor Type	.25	2	.13	1.30
S/Presentation Type x Distractor Type	2.09	21	.10	

Summary of ANOVA relating to data in Experiment III (p.96)

Source	SS	df	MS	F
Retention Interval	2.61	2	1.31	3.36
S/Retention Interval	17.77	45	.39	
Distractor Type	9.76	1	9.76	19.92
Retention Interval x Distractor Type	.01	2	0	0
S/Retention Interval x Distractor Type	22.13	45	.49	

Summary of ANOVA relating to data in Experiment IV (p.120)

Source	SS	df	MS	F
Presentation Type	0	1	0	0
S/Presentation Type	6.88	22	.31	
Distractor Type	32.44	1	32.44	141.04
Presentation Type x Distractor Type	.03	1	.03	.13
S/Presentation Type x Distractor Type	5.10	22	.23	

Summary of ANOVA relating to data in Experiment IV (p.120)

Source	SS	df	MS	F
Presentation Type	225000	1	225000	.82
S/Presentation Type	5741000	21	273381	
Distractor Type	43000	1	43000	.53
Presentation Type x Distractor Type	1000	1	1000	.01
S/Presentation Type x Distractor Type	1714000	21	81619	

Summary of ANOVA relating to data in Experiment IV (p.122)

Source	SS	df	MS	F
Presentation Type	60000	1	60000	.21
S/Presentation Type	6086000	21	289810	
Distractor Type	2708000	1	2708000	42.89
Presentation Type x Distractor Type	35000	1	35000	.55
S/Presentation Type x Distractor Type	1326000	21	63143	

Summary of ANOVA relating to data in Experiment V (p.129)

Source	SS	df	MS	F
Presentation Type	.49	1	.49	.60
S/Presentation Type	18.12	22	.82	
Distractor Type	33.17	1	33.17	85.05
Presentation Type x Distractor Type	.11	1	.11	.28
S/Presentation Type x Distractor Type	8.54	22	.39	

Summary of ANOVA relating to data in Experiment V (p.129)

Source	SS	df	MS	F
Experiment	4.21	1	4.21	7.39
Presentation Type	.28	1	.28	.49
Experiment x Presentation Type	.21	1	.21	.37
S/Experiment x Presentation Type	25	44	.57	
Distractor Type	65.61	1	65.61	211.65
Experiment x Distractor Type	0	1	0	0
Presentation Type x Distractor Type	0	1	0	0
Experiment x Presentation Type x Distractor Type	.14	1	.14	.45
Distractor Type x S/Experiment x Presentation Type	13.64	44	.31	

Summary of ANOVA relating to data in Experiment V (p.133)

Source	SS	df	MS	F
Presentation Type	310000	1	310000	.52
S/Presentation Type	10753000	18	597389	
Distractor Type	20000	1	20000	.21
Presentation Type x Distractor Type	254000	1	254000	2.71
S/Presentation Type x Distractor Type	1689000	18	93833	

Summary of ANOVA relating to data in Experiment V (p.133)

Source	SS	df	MS	F
Experiment	118000	1	118000	.28
Presentation Type	531000	1	531000	1.26
Experiment x Presentation Type	6000	1	6000	.01
S/Experiment x Presentation Type	16493000	39	422897	
Distractor Type	62000	1	62000	.71
Experiment x Distractor Type	3000	1	3000	.03
Presentation Type x Distractor Type	139000	1	139000	1.59
Experiment x Presentation Type x Distractor Type	114000	1	114000	1.31
Distractor Type x S/Experiment x Presentation Type	3403000	39	87256	

Summary of ANOVA relating to data in Experiment V (p.133)

Source	SS	df	MS	F
Presentation Type	333000	1	333000	.97
S/Presentation Type	7552000	22	343273	
Distractor Type	2719000	1	2719000	33.07
Presentation Type x Distractor Type	5000	1	5000	.06
S/Presentation Type x Distractor Type	1809000	22	82227	

Summary of ANOVA relating to data in Experiment V (p.134)

Source	SS	df	MS	F
Experiment	334000	1	334000	1.05
Presentation Type	339000	1	339000	1.07
Experiment x Presentation Type	56000	1	56000	.18
S/Experiment x Presentation Type	13637000	43	317140	
Distractor Type	5428000	1	5428000	74.45
Experiment x Distractor Type	0	1	0	0
Presentation Type x Distractor Type	7000	1	7000	.10
Experiment x Presentation Type x Distractor Type	21000	1	21000	.29
Distractor Type x S/Experiment x Presentation Type	3135000	43	72907	

<u>Summary of ANOVA relating to data in Experiment V (p.134)</u>

Source	SS	df	MS	F
Experiment	47000	1	47000	.22
Presentation Type	39000	1	39000	.18
Experiment x Presentation Type	415000	1	415000	2
S/Experiment x Presentation Type	8487000	39	217615	1.91

Summary of ANOVA relating to data in Experiment V (p.134)

Source	SS	df	MS	F
Experiment	166000	1	166000	1.15
Presentation Type	261000	1	261000	1.81
Experiment x Presentation Type	35000	1	35000	.24
S/Experiment x Presentation Type	6356000	44	144455	

Summary of ANOVA relating to data in Experiment V (p.134)

Source	SS	df	MS	F
Presentation Type	11	1	11	.02
S/Presentation Type	11980	22	545	
Transformation Type	4743	3	1581	5.90
Presentation Type x Transformation Type	1210	3	403	1.50
S/Presentation Type x Transformation Type	17714	66	268	

REFERENCES

ALLEN, G.L., SIEGEL, A.W., & ROSINSKI, R.R. The role of perceptual context in structuring spatial knowledge. Journal of Experimental Psychology: Human Learning and Memory, 1978, 4, 617-630.

ALLIK, J.P., & SIEGEL, A.W. Facilitation of sequential short-term memory with pictorial stimuli. Journal of Experimental Psychology, 1974, 103, 567-573.

ANDERSON, J.R. Verbatim and propositional representation of sentences in immediate and long-term memory. Journal of Verbal Learning and Verbal Behaviour, 1974, 13, 149-162.

ANDERSON, J.R., & BOWER, G.H. Human Associative Memory, Washington, D.C., Winston, 1974.

ANTES, J.R. Recognising and localising features in brief picture presentations. Memory and Cognition, 1977, 5, 155-161.

ARBUCKLE, T.Y., & KATZ, W.A. Structure of memory traces following semantic and nonsemantic orientation tasks in incidental learning. Journal of Experimental Psychology: Human Learning and Memory, 1976, 2, 362-369.

BADDELEY, A.D. The trouble with levels: a re-examination of Craik and Lockhart's framework for memory research. Psychological Review, 1978, 85, 139-152.

BAGGETT, P. Memory for explicit and implicit information in picture stories. Journal of Verbal Learning and Verbal Behaviour, 1975, 14, 538-548.

BAHRICK, H.P., & BAHRICK, P. Independence of verbal and visual codes of the same stimuli. Journal of Experimental Psychology, 1971, 91, 344-346.

BAHRICK, H.P., & BOUCHER, B. Retention of visual and verbal codes of the same system. Journal of Experimental Psychology, 1968, 78, 417-422.

BAKER, L. Processing temporal relationships in simple stories: effects of input sequence. Journal of Verbal Learning and Verbal Behaviour, 1978, 17, 559-572.

BARTLETT, F.C. Remembering: A Study in Experimental and Social Psychology, Cambridge, England: Cambridge University Press, 1932.

BATES, E., MASLING, M., & KINTSCH, W. Recognition memory for aspects of dialogue. Journal of Experimental Psychology: Human Learning and Memory, 1978, 4, 187-197.

BIEDERMAN, I. Perceiving real-world scenes. Science, 1972, 177, 77-80.

BIEDERMAN, I. On processing information from a glance at a scene: some implications for a syntax and semantics of visual processing. In S. Trev (Ed), User-oriented Design of Interactive Graphics Systems. Association for Computing Machinery, Inc., 1977.

BIEDERMAN, I., GLASS, A.L., & STACY, E.W., JR. Searching for objects in real-world scenes. Journal of Experimental Psychology, 1973, 97, 22-27.

BIEDERMAN, I., RABINOWITZ, J.C., GLASS, A.L., & STACY, E.W., JR. On the information extracted from a glance at a scene. Journal of Experimental Psychology, 1974, 103, 597-600.

BOBROW, S.A., & BOWER, G.H. Comprehension and recall of sentences. Journal of Experimental Psychology, 1969, 80, 455-461.

BOBROW, D.G., & NORMAN, D.A. Some principles of memory schemata. In D.G. Bobrow & A. Collins (Eds), Representation and Understanding: Studies in Cognitive Science, New York: Academic Press, 1975.

BOUSEFIELD, W.A., COHEN, B.H., & WHITMARSH, G.A. Associative clustering in the recall of words of different taxonomic frequencies of occurrence. Psychological Reports, 1958, 4, 39-44.

BOUSEFIELD, W.A., ESTERSON, J., & WHITMARSH, G.A. The effects of concomitant coloured and uncoloured pictorial representations on the learning of stimulus words. Journal of Applied Psychology, 1957, 41, 165-168.

BOWER, G.H. Analysis of a mnemonic device. American Scientist, 1970, 58, 496-510.

BOWER, G.H. A selective review of organizational factors in memory. In E. Tulving & W. Donaldson (Eds), Organisation and Memory. New York: Academic Press, 1972.

BOWER, G.H., BLACK, J.B., & TURNER, T.J. Scripts in Memory for Text. Cognitive Psychology, 1979, 11, 177-220.

BOWER, G.H., & KARLIN, M.B. Depth of processing pictures of faces and recognition memory. Journal of Experimental Psychology, 1974, 4, 751-757.

BOWER, G.H., KARLIN, M.B., & DUECK, A. Comprehension and memory for pictures. Memory and Cognition, 1975, 3, 216-220.

BRANSFORD, J.D., & FRANKS, J.J. The abstraction of linguistic ideas. Cognitive Psychology, 1971, 2, 331-350.

BRANSFORD, J.D., & FRANKS, J.J. The abstraction of linguistic ideas: A review. Cognition, 1972, 1, 211-249.

BRUNER, J.S. On perceptual readiness. Psychological Review, 1957, 64, 123-152.

CARMICHAEL, L., HOGAN, H.P., & WALTER, A.A. An experimental study on the effect of language on the reproduction of visually perceived forms. Journal of Experimental Psychology, 1932, 15, 73-86.

COFER, C.N., BRUCE, D.R., & REICHER, G.M. Clustering in free recall as a function of certain methodological variations. Journal of Experimental Psychology, 1966, 71, 858-866.

CONNOR, J.M. Effects of organisation and expectancy on recall and recognition. Memory and Cognition, 1977, 5, 315-318.

CORTIS PARK, D., & WHITTEN II, W.B. Abstraction of linguistic, imaginal, and pictorial ideas. Journal of Experimental Psychology: Human Learning and Memory, 1977, 3, 525-538.

CRAIK, F.I.M. A "levels of analysis" view of memory. In P. Pliner, L. Krames, and T.M. Alloway (Eds), Communication and Affect: Language and Thought, New York: Academic Press, 1973.

CRAIK, F.I.M., & LOCKHART, R.S. Levels of processing: a framework for memory research. Journal of Verbal Learning and Verbal Behaviour, 1972, 11, 671-684.

CRAIK, F.I.M., & TULVING, E. Depth of processing and the retention of words in episodic memory. Journal of Experimental Psychology: General, 1975, 104, 268-294.

D'AGOSTINO, P.R., O'NEILL, B.J., & PAIVIO, A. Memory for pictures and words as a function of level of processing: depth or dual coding? Memory and Cognition, 1977, 5, 252-256.

DALLETT, K., & WILCOX, S.G. Remembering pictures vs remembering descriptions. Psychonomic Science, 1968, 11, 139-140.

DANIEL, T.C. Nature of the effects of verbal language on recognition memory for form. Journal of Experimental Psychology, 1972, 96, 152-157.

DAVIES, G., MILNE, J.E., & GLENNIE, B.J. On the significance of "double encoding" for superior recall of pictures to names. Quarterly Journal of Experimental Psychology, 1973, 25, 413-423.

DURSO, F.T., & JOHNSON, M. Facilitation in naming and categorizing repeated pictures and words. Journal of Experimental Psychology: Human Learning and Memory, 1979, 5, 449-459.

EPSTEIN, W., ROCK, I., & ZUCKERMAN, C.B. Meaning and familiarity in associative learning. Psychological Monographs, 1960, 74, (4, Whole No. 491).

FLAGG, P.W. Semantic integration in sentence memory? Journal of Verbal Learning and Verbal Behaviour, 1976, 15, 492-504.

FRANKEN, R.E., & ROWLAND, G.L. Organisational factors in picture recognition. Acta Psychologica, 1978, 42, 441-452.

FREUND, R.D. Verbal and non-verbal processes in picture recognition (Doctoral dissertation, Stanford University, 1971). Dissertation Abstracts International, 1972, 32, 4886-B. (University Microfilms No. 72-5912, 73).

FRIEDMAN, A. Framing pictures: the role of knowledge in automatized encoding and memory for gist. Journal of Experimental Psychology: General, 1979, 108, 316-355.

GARNER, W.R. Uncertainty and Structure as Psychological Concepts. New York: Wiley, 1962.

GENTNER, D., & LOFTUS, E.F. Integration of verbal and visual information as evidenced by distortions in picture memory. American Journal of Psychology, 1979, 92, 363-375.

GIBSON, J.J. The Senses Considered As Perceptual Systems. Boston: Houghton Mifflin, 1966.

GOLDSTEIN, A.G., & CHANCE, J.E. Visual recognition memory for complex configurations. Perception and Psychophysics, 1970, 9, 237-240.

GRAESSER, A. II, & MANDLER, G. Recognition memory for the meaning and surface structure of sentences. Journal of Experimental Psychology: Human Learning and Memory, 1975, 104, 238-248.

GREEN, D.M., & SWETS, J.A. Signal Detection Theory and Psychophysics. New York: Wiley, 1966.

HABER, R.N. Information Processing Approaches to Visual Perception. New York: Holt, Rinehart & Winston, 1969.

HEAD, H. Studies in Neurology, London: H. Frowde, 1920.

HOCK, H.S., GORDON, G.P., & CORCORAN, S.K. Alternative processes in the identification of familiar pictures. Memory and Cognition, 1976, 4, 265-271.

HOCK, H.S., GORDON, G.P., & WHITEHURST, R. Contextual relations: The influence of familiarity, physical plausibility, and belongingness. Perception and Psychophysics, 1974, 16, 4-8.

HOCK, H.S., ROMANSKI, L., GALIE, A., & WILLIAMS, C.S. Real-world schemata and scene recognition in adults and children. Memory and Cognition, 1978, 6, 423-431.

HUNT, R.R., & ELLIOTT, J.M. The role of non-semantic information in memory: orthographic distinctiveness effects on retention. Journal of Experimental Psychology: General, 1980, 109, 49-74.

HYDE, T.S., & JENKINS, J.J. Recall for words as a function of semantic, graphic, and syntactic orienting tasks. Journal of Verbal Learning and Verbal Behaviour, 1973, 12, 471-480.

JAMES, C.T., & HILLENGER, M.L. The role of confusion in the semantic integration paradigm. Journal of Verbal Learning and Verbal Behaviour, 1977, 16, 711-721.

JENKINS, J.J. Remember that old theory of memory? Well, forget it! American Psychologist, 1974, 29, 785-795.

JENKINS, J.J., NEALE, D.C., & DENO, S.L. Differential memory for picture and word stimuli. Journal of Educational Psychology, 1967, 58, 303-307.

JENKINS, J.J., & RUSSELL, W.A. Associative clustering during recall. Journal of Abnormal and Social Psychology, 1952, 47, 818-821.

JENKINS, J.J., WALD, J., & PITTENGER, J.B. Apprehending pictorial events: An instance of psychological cohesion. In C.W. Savage (Ed), Minnesota Studies in the Philosophy of Science, (Vol. 9). Minneapolis: University of Minnesota Press, 1978.

KATZ, S., & GRUENEWALD, P. The abstraction of linguistic ideas in "meaningless" sentences. Memory and Cognition, 1974, 2, 737-741.

KEENAN, J., MacWHINNEY, B., & MAYHEW, D. Pragmatics in memory: A study in natural conversation. Journal of Verbal Learning and Verbal Behaviour, 1977, 16, 549-560.

KINTSCH, W. Recognition and free recall of organised lists. Journal of Experimental Psychology, 1968, 78, 481-487.

KINTSCH, W. Learning, Memory and Conceptual Processes, New York: Wiley, 1970.

KINTSCH, W. The Representation of Meaning in Memory. Hillsdale, N.J.: Erlbaum, 1974. (second edition of Learning, Memory and Conceptual Processes, 1970).

KINTSCH, W. Memory and Cognition, New York: Wiley, 1977.

KINTSCH, W., & BATES, E. Recognition memory for statements from a classroom lecture. Journal of Experimental Psychology: Human Learning and Memory, 1977, 3, 150-159.

KINTSCH, W., & KEENAN, J. Reading rate and retention as a function of the number of propositions in the base structure of sentences. Cognitive Psychology, 1973, 5, 257-274.

KINTSCH, W., MANDEL, T.S., & KOZMINSKY, E. Summarizing scrambled stories. Memory and Cognition, 1977, 5, 547-552.

KIRK, R.E. Experimental Design: Procedures for the Behavioural Sciences. California: Brooks/Cole, 1966.

KLATZKY, R.L., & RAFNEL, K. Labelling effects on memory for nonsense pictures. Memory and Cognition, 1976, 4, 717-720.

KLEIN, K., & SALTZ, E. Specifying the mechanisms in a levels-of-processing approach to memory. Journal of Experimental Psychology: Human Learning and Memory, 1976, 2, 671-679.

KOSSLYN, S.M. Information representation in visual images, Cognitive Psychology, 1975, 7, 341-371.

KOSSLYN, S.M. Imagery and Internal Representation in E.H. Rosch & B.B. Lloyd (Eds). Cognition and Categorisation, Pototmac, Md: Erlbaum, 1978.

KRAFT, R.N., & JENKINS, J.J. Memory for lateral orientation of slides in picture stories. Memory and Cognition, 1977, 5, 397-403.

KURTZ, K., & HOVLANC, C.I. The effect of verbalisation during observation of stimulus objects upon accuracy of recognition and recall. Journal of Experimental Psychology, 1953, 45, 157-164.

LOFTUS, E.F. Leading questions and the eye witness report. Cognitive Psychology, 1975, 7, 560-572.

LOFTUS, E.F. Shifting human colour memory. Memory and Cognition. 1977, 5, 696-699.

LOFTUS, E.F., MILLER, D.G., & BURNS, H.J. Semantic integration of verbal information into visual memory. Journal of Experimental Psychology: Human Learning and Memory, 1978, 4, 19-31.

LOFTUS, G.R. Eye fixations and recognition memory for pictures. Cognitive Psychology, 1972, 3, 525-551.

LOFTUS, G.R., & BELL, S.M. Two types of information in picture memory. Journal of Experimental Psychology: Human Learning and Memory, 1975, 104, 103-113.

LOFTUS, G.R., & LOFTUS, E.F. Human Memory. The Processing of Information. New Jersey: Erlbaum, 1976.

MACKWORTH, N.H., & MORANDI, A.J. The gaze selects information details within pictures. Perception and Psychophysics, 1967, 2, 547-551.

MANDLER, G. Organisation and recognition. In E. Tulving & W. Donaldson (Eds), Organization of Memory. New York: Academic Press, 1972.

MANDLER, J.M. A code in the node. Discourse Processes, 1978, 1, 14-35.

MANDLER, J.M., & JOHNSON, N.S. Some of the thousand words a picture is worth. Journal of Experimental Psychology: Human Learning and Memory, 1976, 2, 529-540.

MANDLER, J.M., & PARKER, R.E. Memory for descriptive and spatial information in complex pictures. Journal of Experimental Psychology: Human Learning and Memory, 1976, 2, 38-48.

MANDLER, J.M., & RITCHEY, G.H. Long-term memory for pictures. Journal of Experimental Psychology: Human Learning and Memory, 1977, 3, 386-396.

MANDLER, J.M., & STEIN, N.L. Recall and recognition of pictures by children as a function of organisation and distractor similarity. Journal of Experimental Psychology, 1974, 102, 657-669.

MARSLEN-WILSON, W., & TYLER, L.K. Memory and levels of processing in a psycholinguistic context. Journal of Experimental Psychology: Human Learning and Memory, 1976, 2, 112-119.

MINSKY, M. Frame-systems: A framework for representing knowledge. In P.H. Winston (Ed), The Psychology of Computer Vision, New York: McGraw-Hill, 1975.

MISTLER-LACHMAN, J.L. Depth of comprehension and sentence memory. Journal of Verbal Learning and Verbal Behaviour, 1974, 13, 98-106.

NAVON, D. Forest before trees: the precedence of global features in visual perception. Cognitive Psychology, 1977, 9, 353-383.

NEISSER, U. Cognitive Psychology. New York: Appleton-Century-Crofts, 1967.

NEISSER, U. Cognition and Reality. San Francisco: W.H. Freeman & Co., 1976.

NELSON, D.L., & REED, V.S. On the nature of pictorial encoding: a levels-of-processing analysis. Journal of Experimental Psychology: Human Learning and Memory, 1976, 2, 49-57.

NELSON, D.L., REED, V.S., & McEVOY, C.L. Learning to order pictures and words. A model of sensory and semantic encoding. Journal of Experimental Psychology: Human Learning and Memory, 1977, 3, 485-497.

NELSON, D.L., REED, V.S., & WALLING, J.R. Pictorial superiority effect. Journal of Experimental Psychology: Human Learning and Memory, 1976, 2, 523-528.

NELSON, T.O. Repetition and depth of processing. Journal of Verbal Learning and Verbal Behaviour, 1977, 16, 151-171.

NELSON, T.O., METZLER, J., & REED, D.A. Role of details in the long-term retention of pictures and verbal descriptions. Journal of Experimental Psychology, 1974, 102, 184-186.

NELSON, T.O., & VINING, S.K. Effect of semantic versus structural processing on long-term retention. Journal of Experimental Psychology: Human Learning and Memory, 1978, 4, 198-209.

NEUMANN, P.G. An attribute frequency model for the abstraction of prototypes. Memory and Cognition, 1974, 2, 241-248.

NICKERSON, R.S. A note on long-term recognition memory for pictorial material. Psychonomic Science, 1968, 11, 58.

NORMAN, D.A., & RUMELHART, D.E. Memory and knowledge. In D.A. Norman & D.E. Rumelhart (Eds), Explorations in Cognition. San Francisco: Freeman, 1975a.

NORMAN, D.A., & RUMELHART, D.E. Reference and comprehension. In D.A. Norman & D.E. Rumelhart (Eds), Explorations in Cognition. San Francisco: Freeman, 1975b.

PAIVIO, A. Imagery and Verbal Processes. New York: Holt, Rinehart, and Winston, 1971.

PAIVIO, A. Perceptual comparisons through the mind's eye. Memory and Cognition, 1975, 3, 636-647.

PAIVIO, A., & CSAPO, K. Concrete-image and verbal memory codes. Journal of Experimental Psychology, 1969, 80, 279-285.

PAIVIO, A., & CSAPO, K. Picture superiority in free recall: Imagery or dual coding? Cognitive Psychology, 1973, 5, 179-206.

PALMER, S.E. The nature of perceptual representation: An examination of the analog/propositional controversy. In R.C. Schank & B. Nash - Webber (Eds), Theoretical issues in natural language processing. Arlington, Va: Tinlap Press, 1975a.

PALMER, S.E. The effects of contextual scenes on the identification of objects. Memory and Cognition, 1975b, 3, 519-526.

PALMER, S.E. Visual perception and world knowledge: Notes on a model of sensory cognitive interaction. In D.A. Norman & D.E. Rumelhart (Eds), Explorations in Cognition. San Francisco: Freeman, 1975c.

PALMER, S.E. Hierarchical structure in perceptual representation. Cognitive Psychology, 1977, 9, 441-475.

PALMER, S.E. Fundamental aspects of cognitive representation. In E.H. Rosch & B.B. Lloyd (Eds), Cognition and Categorisation. Potomac, Md: Erlbaum, 1978.

PARIS, S.G., & MAHONEY, C.J. Cognitive integration in children's memory for sentences and pictures. Child Development, 1974, 45, 633-642.

PERFETTI, C.A., & LESGOLD, A.M. Discourse comprehension and sources of individual differences. In M. Just & P. Carpenter (Eds), Cognitive Processes in Comprehension. Hillsdale, N.J.: Erlbaum, 1977.

PEZDEK, K. Cross-modality semantic integration of sentence and picture memory. Journal of Experimental Psychology: Human Learning and Memory, 1977, 3, 515-524.

PEZDEK, K. Recognition memory for related pictures. Memory and Cognition, 1978, 6, 64-69.

PEZDEK, K., & EVANS, G.W. Visual and verbal memory for objects and their spatial locations. Journal of Experimental Psychology: Human Learning and Memory, 1979, 5, 360-373.

PIAGET, J., & INHELDER, B. Memory and Intelligence. New York: Basic Books, 1973.

PICK, A.D. Improvement of visual and tactual form discrimination. Journal of Experimental Psychology, 1965, 69, 331-339.

PITTENGER, J.B., & JENKINS, J.J. Apprehension of pictorial events: the case of a moving observer in a static environment. Bulletin of the Psychonomic Society, 1979, 13, 117-120.

POSNER, M.I. Abstraction and the process of recognition. In G.H. Bower & J.T. Spence (Eds), The Psychology of Learning and Motivation, (Vol.3), New York: Academic Press, 1969.

POSNER, M.I., & KEELE, S.W. On the genesis of abstract ideas. Journal of Experimental Psychology, 1968, 77, 353-363.

POSNER, M.I., & KEELE, S.W. Retention of abstract ideas. Journal of Experimental Psychology, 1970, 83, 304-308.

POSTMAN, L. Verbal learning and memory. Annual Review of Psychology, 1975, 291-335.

POTTER, M.C., VALIAN, V.V., & FAULCONER, B.A. Representation of a sentence and its pragmatic implications: Verbal, imagistic, or abstract? Journal of Verbal Learning and Verbal Behaviour, 1977, 16, 1-12.

PYLYSHYN, Z.W. What the mind's eye tells the mind's brain: A critique of mental imagery. Psychological Bulletin, 1973, 80, 1-24.

PYLYSHYN, Z.W. Do we need images and analogues? In R.C. Shank & B. Nash-Webber (Eds), Theoretical issues in natural language processing. Arlington, Va: Tinlap Press, 1975.

RAFNEL, K.J., & KLATZKY, R.L. Meaningful-interpretation effects on codes of nonsense pictures. Journal of Experimental Psychology: Human Learning and Memory, 1978, 4, 631-636.

REITMAN, J.S., & BOWER, G.H. Storage and later recognition of examples of concepts. Cognitive Psychology, 1973, 4, 194-206.

ROSENBERG, S., & SCHILLER, W.J. Semantic coding and incidental sentence recall. Journal of Experimental Psychology, 1971, 90, 345-346.

ROSENBERG, S., & SIMON, H.A. Modelling semantic memory: Effects of presenting semantic information in different modalities. Cognitive Psychology, 1977, 9, 293-325.

SACHS, J. Recognition memory of syntactic and semantic aspects of connected discourse. Perception and Psychophysics, 1967, 2, 437-444.

SACHS, J. Memory in reading and listening to discourse. Memory and Cognition, 1974, 2, 95-100.

SCHULMAN, A.I. Recognition memory for targets from a scanned word list. British Journal of Psychology, 1971, 62, 335-346.

SHANK, R.C., & ABELSON, R.P. Scripts, plans and knowledge. International Joint Conference on Artificial Intelligence. USSR, Tbilisi, Georgia: Advance Papers of the Fourth International Joint Conference, 1975.

SHEPARD, R.N. Recognition memory for words, sentences, and pictures. Journal of Verbal Learning and Verbal Behaviour, 1967, 6, 156-163.

SINGER, M., & ROSENBERG, S.T. The role of grammatical relations in the abstraction of linguistic ideas. Journal of Verbal Learning and Verbal Behaviour, 1973, 12, 273-284.

SNODGRASS, J.G., & VANDERWART, M. A standardised set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. Journal of Experimental Psychology: Human Learning and Memory, 1980, 6, 174-215.

STANDING, L. Learning 10,000 pictures. Quarterly Journal of Experimental Psychology, 1973, 25, 207-222.

STANDING, L., CONEZIO, J., & HABER, R.N. Perception and memory for pictures: Single trial learning of 2,500 visual stimuli. Psychonomic Science, 1970, 19, 73-74.

STEIN, N.L., & GLENN, C.G. An analysis of story comprehension in elementary school children. In R. Freedle (Ed), Multi-disciplinary perspectives in discourse comprehension. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1978.

STEIN, N.L., & NEZWORSKI, T. The effects of organization and instructional set on story memory. Discourse Processes, 1978, 1, 177-194.

STERNBERG, S. High-speed scanning in human memory. Science, 1966, 153, 652-654.

TREISMAN, A., & TUXWORTH, J. Immediate and delayed recall of sentences after perceptual processing at different levels. Journal of Verbal Learning and Verbal Behaviour, 1974, 13, 38-44.

TRESSELT, M.E., & MAYZNER, M.S. A study of incidental learning. Journal of Psychology, 1960, 50, 339-347.

TULVING, E. Subjective organization in free recall of "unrelated" words. Psychological Review, 1962, 69, 344-354.

TULVING, E. Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds), Organisation and Memory, New York: Academic Press, 1972.

VAN DIJK, T.A. Semantic macro-structures and knowledge frames in discourse. In M. Just & P. Carpenter (Eds), Cognitive Processes in Comprehension, Hillsdale, N.J.: Erlbaum, 1977.

WEAVER, G.E., & STANNY, C.J. Short-term retention of pictorial stimuli as assessed by a probe recognition technique. Journal of Experimental Psychology: Human Learning and Memory, 1978, 4, 55-65.

WHITE, K.G. Abstracting ideas from sentences. New Zealand Psychologist, 1974, 3, 74-84.

WINOGRAD, T. Understanding Natural Language, New York: Academic Press, 1972.

WOLK, S. The influence of meaningfulness upon intentional and incidental learning of verbal material. Memory and Cognition, 1974, 2, 189-193.