How sketches work: a cognitive theory for improved system design

This item was submitted to Loughborough University’s Institutional Repository by the/an author.

Additional Information:

- A Doctoral Thesis. Submitted in partial fulfillment of the requirements for the award of Doctor of Philosophy of Loughborough University. NOTE: Images removed for copyright purposes.

Metadata Record: https://dspace.lboro.ac.uk/2134/7418

Publisher: © Jonathan C. Fish

Please cite the published version.
This item is held in Loughborough University’s Institutional Repository (https://dspace.lboro.ac.uk/) and was harvested from the British Library’s EThOS service (http://www.ethos.bl.uk/). It is made available under the following Creative Commons Licence conditions.

For the full text of this licence, please go to: http://creativecommons.org/licenses/by-nc-nd/2.5/
HOW SKETCHES WORK
A COGNITIVE THEORY FOR IMPROVED SYSTEM DESIGN

by

JONATHAN C. FISH

A Doctoral Thesis
Submitted in partial fulfilment of the requirements
for the award of

Doctor of Philosophy of the Loughborough University of Technology

February 20th 1996
(Corrected 14th November 1996)

Copyright © by Jonathan C. Fish 1996
HOW SKETCHES WORK:  
A Cognitive Theory for Improved System Design

ABSTRACT

Evidence is presented that in the early stages of design or composition the mental processes used by artists for visual invention require a different type of support from those used for visualising a nearly complete object. Most research into machine visualisation has as its goal the production of realistic images which simulate the light pattern presented to the retina by real objects. In contrast sketch attributes preserve the results of cognitive processing which can be used interactively to amplify visual thought. The traditional attributes of sketches include many types of indeterminacy which may reflect the artist's need to be "vague".

Drawing on contemporary theories of visual cognition and neuroscience this study discusses in detail the evidence for the following functions which are better served by rough sketches than by the very realistic imagery favoured in machine visualising systems.

1. Sketches are intermediate representational types which facilitate the mental translation between descriptive and depictive modes of representing visual thought.
2. Sketch attributes exploit automatic processes of perceptual retrieval and object recognition to improve the availability of tacit knowledge for visual invention.
3. Sketches are percept-image hybrids. The incomplete physical attributes of sketches elicit and stabilise a stream of super-imposed mental images which amplify inventive thought.
4. By segregating and isolating meaningful components of visual experience, sketches may assist the user to attend selectively to a limited part of a visual task, freeing otherwise over-loaded cognitive resources for visual thought.
5. Sequences of sketches and sketching acts support the short term episodic memory for cognitive actions. This assists creativity, providing voluntary control over highly practised mental processes which can otherwise become stereotyped.

An attempt is made to unite the five hypothetical functions. Drawing on the Baddeley and Hitch model of working memory, it is speculated that the five functions may be related to a limited capacity monitoring mechanism which makes tacit visual knowledge explicitly available for conscious control and manipulation. It is suggested that the resources available to the human brain for imagining non-existent objects are a cultural adaptation of visual mechanisms which evolved in early hominids for responding to confusing or incomplete stimuli from immediately present objects and events. Sketches are cultural inventions which artificially mimic aspects of such stimuli in order to capture these shared resources for the different purpose of imagining objects which do not yet exist.

Finally the implications of the theory for the design of improved machine systems is discussed. The untidy attributes of traditional sketches are revealed to include cultural inventions which serve subtle cognitive functions. However traditional media have many short-comings which it should be possible to correct with new technology. Existing machine systems for sketching tend to imitate non-selectively the media bound properties of sketches without regard to the functions they serve. This may prove to be a mistake. It is concluded that new system designs are needed in which meaningfully structured data and specialised imagery amplify without interference or replacement the impressive but limited creative resources of the visual brain.

Jonathan Fish, 20th February 1996
# CONTENTS

## LIST OF ILLUSTRATIONS

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
</tr>
</tbody>
</table>

## PREFACE

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
</tr>
</tbody>
</table>

The Origins and Evolving Purpose of this Study. 12

Overview of Contents. 15

Acknowledgements. 21

A Note on Typefaces. 22

## 1. WHAT IS A SKETCH ?

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
</tr>
</tbody>
</table>

1.1 The Problem of Definition. 23

1.2 The Problem of Defining Function. 26

1.3 The Problem of Defining Attributes. 27

1.4 Three Visualizing Methods for Experiment in the Arts. 31

1.5 Sketching from the Motif and from Imagination. 39

1.6 Sketching in Context.
   1.6.1 "Les Demoiselles d'Avignon": A Fine Art example. 41
   1.6.2 The "Yamazaki Serving Collection": A Design Example. 43

1.7 Summary and Conclusions. 44

## 2. SKETCH ATTRIBUTES AND THEIR INFLUENCE.

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>46</td>
</tr>
</tbody>
</table>

2.1 Three Kinds of Indeterminacy.
   2.1.1 Generality 46
   2.1.2 Ambiguity 47
   2.1.3 Vagueness 47
   2.1.4 Discussion 48

2.2 The Origins of Sketch Indeterminacies. 48

2.3 Indeterminacies in Fine Art; The Beholder's Share. 55

2.4 Photography, Realism and the Sketch. 58

2.5 Sketches and Written Notes. 63

2.6 On Being the Right Size. 64

2.7 Discussion and Conclusions. 65

## 3. SKETCHES AND REPRESENTATIONAL TYPE.

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>68</td>
</tr>
</tbody>
</table>

3.1 Introduction 68
3.2 Description versus Depiction: Defining the attributes. 68
3.3 Indices, Icons and Symbols: The Sagacity of C. S. Peirce. 72
3.4 Visual Representation: A Continuum of Type. 74
   3.4.1 Replicas and Examples. 74
   3.4.2 Working Models and Simulations. 75
   3.4.3 Static Three-D Models 75
   3.4.4 Holograms, Television and Photography. 75
   3.4.5 Figurative Art and Animated Graphics. 75
   3.4.6 Line Drawings and Illustrations. 76
   3.4.7 Sketches 77
   3.4.8 Maps and Diagrams 77
   3.4.9 Pictorial Notation Systems, Graphic Signs. 78
   3.4.10 Natural Language 79
   3.4.11 Pure Notation Systems. 80
   3.4.12 Mathematical Symbol Systems. 80
3.5 Intermediate types as Translation Systems. 81
3.6 Cognitive Representation: Is there also a Continuum of Type? 83
   3.6.1 The Neural Medium. 83
   3.6.2 Different Approaches. 85
   3.6.3 Cognitive Depiction. 86
   3.6.4 Cognitive Description. 87
   3.6.5 Intermediate or Mixed Types and Cognitive Type translation. 88
3.7 Evidence for Sketch Mediation of Type Translation. 91
3.8 Conclusions concerning Representational Type. 93
3.9 Summary 95
4. SKETCH PERCEPTION AND RECOGNITION. 96
4.1 Introduction 96
4.2 Sketching and Theories of Perception. 96
   4.2.1 The Classical Problems of Perception 96
   4.2.2 The Gibsonian Theory of "Direct Perception" and its Influence. 98
   4.2.3 The Gestalt School and its Influence. 100
   4.2.4 Helmholtz's Theory of "Unconscious Conclusions". 102
   4.2.5 The Ideas of Ernst Gombrich. 105
   4.2.6 Perception as a Multi-Stage process: Towards a Synthesis 105
4.3 Sketches and Object Recognition. 109
   4.3.1 How Sketches Exploit Recognition Mechanisms. 109
   4.3.2 Feature theories 110
   4.3.3 Structural Descriptions 111
   4.3.4 Theories of Recognition by Components 112
   4.3.5 Object Position, Size and orientation. 114
4.4 Does Sketch Recognition include a Depictive Match? 117
   4.4.1 Categories versus Exemplars. 119
   4.4.2 Schemata: The Importance of Context. 120
   4.4.3 Memory Retrieval 121
   4.4.4 Towards a Model of Sketch Recognition. 122
LIST OF ILLUSTRATIONS.
(To be found in order directly after page 355)

I. VILLARD D'HONNECOURT 356
Swan. Middle of the 13th century. Paris, Bibliotheque Nationale, Ms Fr 19093, folio 7.

II. LEONARDO DA VINCI 357
a. Study for the Battle of Anghiari, about 1503, Venice, Accademia.
b. Neptune, about 1504, Windsor Castle, Royal Library. Sketch b may have been inspired by sketch a despite the different subjects (Gombrich, 1966)

III. LEONARDO DA VINCI 358

IV. ALEXANDER COZENS 359
a. A demonstration "blot" from "A New Method of Landscape"
b. One of three landscapes derived from the above "blot"

V. MAX ERNST 360
"Fire in the Forest" ("Frottage" 1966)
Ernst invented his "frottages" under the influence of Leonardo's "Trattato" (Spies, 1969).

VI. CHARLES JAMES, C. 361
1964 X-Ray Sketch showing relationship of garment to the body. (Coleman 1983)

VII. CHARLES JAMES 362
Sketch for an opera cape. Illustrates the need to preserve multiple alternatives at early stages of design.

VIII. YVES ST LAURENT 363

IX. KARL LAGERFELD 364
Ideas sketches smothered in written notes. Descriptive to depictive translation in progress (Piaggi, 1986)

X. MATISSE 365
Nude. 1901-03 (Elderfield 1984)

XI. PICASSO 366
Sketches for "Les Demoiselles d'Avignon" (Glimcher and Glimcher 1986)
b. Sketch-book No 44, p 24

XII. PICASSO 367
Sketches for "Les Demoiselles d'Avignon" (Glimcher and Glimcher 1986)
b. (Lower) Sketch-book No 44, p 10
XIII. PICASSO
"Les Demoiselles d'Avignon" 1907 (244 x 234 cms)

XIV. ROBERT WELCH
Yamazaki Serving Collection. (1981)
Sketch for a Covered Bowl

XV. ROBERT WELCH
Yamazaki Serving Collection. (1981)
A sketch for the ice-bucket, jug and tray

XVI. ROBERT WELCH
Yamazaki Serving Collection. (1981)
A drawing of critical importance. Robert Welch was selected to design the
Collection on the basis of this idea sketch.

XVII. ROBERT WELCH
Yamazaki Serving Collection. (1983)
Finished tableware (Welch 1986)

XVIII. CEZANNE
Sketch of a Landscape (National Gallery Prague)
Demonstrates Global precedence to object detail

XIX. CEZANNE
Two studies of the head of the artist's son. From a sketch-book circa 1880
(Cezanne 1951)
Object precedence to global structure. (see 6.8.3).

XX. REMBRANDT
Study for the painting "The Rape of Ganymede" Pen and brush. About
1635. Dresden, print Room.

XXI. REMBRANDT
"The Rape of Ganymede" 1635. Dresden. (Goldscheider 1960)

XXII. GOYA
Procession. Black crayon 1811-24 (Gassier 1975)

XXIII. MILLET
"A Heath" The Louvre. (Bacou 1975)

XXIV. MILLET
"The Curtain of Trees" Black crayon. The Louvre 1922.

XXV. THEODORE ROUSSEAU
"Twisted tree " with part misplaced. (Schema experiment 4.3.8)

XXVI. THEODORE ROUSSEAU
"Twisted tree " original with figure correctly placed.

XXVII. SPATIAL FREQUENCY FILTERS.
a.(Top) Photograph of Groucho Marx besides its low and high frequency
filtered derivatives (Ginsburg 1986).
a. Seurat drawing "Woman reading" Undated (Kahn, 1971)
XXVIII. INGRES

XXIX. HENRI DE TOULOUSE-LAUTREC.
Rearing horse. Pencil, 11.4 X 14.5 cm. Museum of Czech literature, Karasek gallery, Prague

XXX. SHELDON TETEWSKY
Degraded scene (from Biederman 1981)

XXXI. AMBIGUOUS FIGURES
a) (Upper) "My wife or Mother in Law" (Gregory 1970)
b) (Lower) "Rabbit or Duck?" (from Chambers & Reisberg 1985)

XXXII. THE HIDDEN MAN.
From Porter (1954)

XXXIII. CAMILLE COROT
Landscape with Tall Trees.
Charcoal with Indian ink. Private collection, Prague (Kesnerova & Spielman 1969)

XXXIV. MATISSE.
Reclining Nude, 1938 (Elderfield 1984)

XXXV. TURNER
Watercolour sketch. (Wilkinson, 1975)
One of the Stages of a Sunset. 9.5 x 13.5 inches

XXXVI. MICRDYNAMICS INC. (BROCHURE 1994)
Machine assisted variations on a theme.

XXXVII. HAROLD COHEN
Drawing generated by "Aaron" software.(1987)

XXXVIII. SOLID MODELLING
DIAGRAMS

DIAGRAM I  (facing pages 67 and 249)
Five functions of the sketch related to Hypotheses 1-5.

DIAGRAM II (facing page 156)
A diagram showing neural connections between various visual regions of the macaque cortex. (Modified from Van Essen 1985)

DIAGRAM III. (facing page 162)
An outline model of processes used in sketch hybrid image formation suggesting a symmetry and shared resources between perceptual processing and memory retrieval processes (see 5.3.7)

DIAGRAM IV. (facing page 252)
Is translation of representational type a unifying link between the cognitive functions of the sketch?

DIAGRAM V. (facing page 253)
Are all the functions of the sketch ways of supporting indeterminacy?

DIAGRAM VI. (facing page 255)
The functions of the sketch as support for visual consciousness. Is this a unifying principle?

DIAGRAM VII. (facing page 258)
A model of sketch support for the cognitive control of working memory resources.

DIAGRAM VIII. (facing page 295)
An outline schema for a sketching system in which mind and machine collaborate.
PREFACE

The Origins and Evolving Purpose of this Study.

The subject matter of this research was chosen after a discussion between the author and Professor Stephen Scrivener, who was then on the staff of the Human Computer Interface Unit at Leicester, but who moved almost immediately afterwards with Professor Edmonds to Loughborough University. Stephen knew from our correspondence that I had a background in both Natural Science (Biochemistry and Computing) and Fine Art (painting) and that I had a special interest in Art and Science relationships. I was at that time teaching Computing, Visual Perception and Colour Science to undergraduate students of Fine Art and Fashion design. At the time of our first meeting I was interested in designing new more meaningful computer data structures for drawings. However my attention was drawn to one particular topic on a list of possible research problems that Stephen had drawn up. It started "Artists and designers often complain that Computer Systems make it hard to be vague". Because sketches are often vague drawings, I agreed, without fully realising what I had "taken on" that I should research the design of improved sketching systems for artists. I had hoped that I would only have to spend at most a year's preliminary study into just exactly what a sketch was (I had only the haziest of ideas). I would then be free to devote the rest of my research time to implementing and testing some new software ideas for a drawing system which would "allow artists to be vague", whatever that might mean. Hopefully my ideas on new data structures for drawings would then come in useful. I mention this in the preface, to answer the possible criticism that I have "shirked" the task of the actually testing some of the implications of my theoretical ideas in machine software. In fact I am an enthusiastic computer graphics programmer and it was only the realisation that truly innovative and useful software for artists will only be possible when we have a better understanding of how sketches interact with cognition that prevented me from trying my hand at implementing some of the many "sketching system" ideas this study has suggested to me.

So it was, at first, a disappointment when I discovered how little helpful research or literature was available on the sketch. It became clear that because sketches are incomplete representations which assist and interact with an artist's private thoughts, any understanding of their function would mean studying the mental processes available to artists for visual thought and invention. It was also clear that there would be no easy solution to the problem of finding machine representations and algorithms to manipulate the hidden subtleties of sketch
attributes. Stephen gave me a lead by pointing out Ernst Gombrich's article "Leonardo's Method of Working out Compositions". This paper and many of Gombrich's other writings, with which I was already familiar, became a starting point for my enquiries. I acknowledge a great debt to the insights of Ernst Gombrich into the nature of representation in art. However his most relevant essays, and his unequalled book "Art and Illusion" were written in the sixties. Since then a minor revolution has occurred in the study of the "visual brain" and a great deal more relevant published research into the mechanisms of visual cognition is now ripe for application to the visual arts and (in my opinion) to the breath-taking possibilities of machine image handling. Thus I have had to study (as a scientist and artist but not as a qualified psychologist) a huge literature on visual cognition, including perception, object recognition, imagery, attention, working memory, creativity and the neuroscience that is increasingly beginning to support these topics. It dawned on me gradually that the task of applying recent progress in these subjects to the visual arts in order to understand the mysterious attributes of untidy sketches was far too great a task for one thesis. Reluctantly I was persuaded that the task of actually implementing a system or part of a system would have to take second place to that of achieving the prerequisite for designing such a system - a plausible theory as to "How Sketches Work". I then believed and still believe that without at least a theory of sketch attributes it is impossible to design appropriate machine systems to replace the humble sketch book. I had also concluded from my experience of teaching that there was something fundamentally wrong with the design of existing machine sketching (paint) systems for artists, powerful and fascinating though such systems can be. I put their design faults down to a lack of any explicit theory about the functions they support and the consequent attempt to simulate "blindly" traditional media - a design philosophy I call for reasons explained in Chapter 9 the "Imitation Bronze" approach.

Thus this study is unashamedly theoretical. Its claim to original content depends mainly upon the ideas and hypotheses (such as those at the beginnings and ends of Chapters 3 to 7) which are derived from the investigation of some consequences of applying cognitive psychology and neuroscience to the visual arts and to the media - traditional and new - currently available to support visual invention. However because cognitive science is in such a state of debate and uncertainty I have been forced, in order to frame a theory of the sketch, to choose between rival theories of cognition and to explain, with reviews of the evidence, my reasons for doing so. In addition to the hypotheses directly concerning sketches I have offered hypotheses or models with more general implications for cognitive theory. Thus on at least three occasions I have felt justified in going beyond the
published literature in order to document a theory or at least a theory variant of my own. I found this was necessary to support my five cognitive "hypotheses" of sketch function. However I have no right to make any strong claims about the general applicability of these theory variants except in so far as they help to explain the cognitive properties of sketches. These theories for which I claim some originality are mentioned in the review of contents which follows.

I find to my alarm that I have been influenced by over 1070 books and papers on art theory and cognitive science. The examiners will be relieved to learn that I have managed to reduce the final number of references to a bibliography of about half that number. Nevertheless a great proportion of the text reviews the work of others. I felt this was necessary to support the hypotheses (numbered one to five) I was presenting and I have throughout tried to keep in mind the ultimate purpose of the research - to improve computer systems for artists. Thus besides supporting a theory of the sketch, the cognitive psychology is intended to provide useful "relevance pointers" for other visual system designers who are not psychologists. All the ideas which are not my own I have acknowledged and carefully referenced. Whenever possible I have quoted (briefly) in italics some of the actual words of an author whose ideas are important to my central thesis. In some cases this includes ideas with which I do not agree but which have been influential in the design of machine visualising systems (Chapter 4).

An important part of the research consisted in simply looking at the sketches (and writings when available) of artists and designers. Clearly the subject matter demands illustrations. I would have liked to have used several hundred illustrations of published sketches. For practical reasons I settled on a mere 38. Because I have frequently had to refer to the same sketch in different places in the text the illustrations are placed together behind the bibliography. I make no apology for discussing sketches by Fine Artists and by Designers under the same headings. I do this not because I wish to underestimate the great differences of purpose and method in the different visual arts but because I am claiming that in the early stages of visual invention there are similarities of sketch function common to all the visual arts which are important, interesting and in need of a better understanding.

Three short papers have already appeared summarising some of the ideas behind this research (Fish and Scrivener 1990, Fish 1991, 1994). However this text is the first attempt to give substance to those early versions of the theory and to document it in more detail.
Overview of Contents

The text is divided into nine chapters and sub-sections classified in decimal notation. It is not always easy to force the products of a massively parallel processor (the brain) into the serial straight-jacket of prose. Chapters 3 to 7 describe logically parallel strands of thought and are interdependent. I have tried to cope with this by making numbered cross-references to previous paragraphs. Occasionally however I know I have erred by repeating myself as an idea introduced in say Chapter 4 is developed in Chapters 7 and 9. I apologise for these lapses.

The core of the thesis is summarised by the following five hypotheses about sketch function which are discussed in Chapters 3 to 7 respectively:

**Hypothesis One.** (Translation of Representational Type)

Sketches have attributes which make them intermediate between depictive (picture-like) and descriptive (language-like) representations. Such attributes are used to facilitate the bidirectional translation between descriptive and depictive modes of thought.

**Hypothesis Two.** (Inventive or unexpected perceptual retrieval)

Human perception and object recognition processes evolved to enable accidentally impoverished stimuli fast access to long term memory for perceptual completion. Sketch attributes exploit such unconscious processes providing memory search and retrieval cues that improve the availability of tacit visual knowledge for invention.

**Hypothesis Three.** (Support for superimposed mental imagery)

It is proposed that sketches are percept-memory hybrids. The incomplete, physical attributes of sketches act as a stimulus for percepts which invite completion from memory and imagination. This results in the generation of transient mental images which after a matching process are transformed and spatially superimposed upon an internal representation of the sketch.

**Hypothesis Four** (Selective attention to Visual Components)

Sketches amplify inventive thought by isolating and representing separately those attributes of visual experience that are of special relevance to a particular task. This assists the user to attend selectively to a limited part of the task, freeing otherwise shared components of cognitive capacity and reducing the complexity of preparatory visual processing.

**Hypothesis Five.** (Conscious monitoring of visual thought)

Sequences of sketching acts support the conscious awareness of one’s own cognition. This assists creativity, providing voluntary control over highly practised mental processes which can otherwise become stereotyped. Unforeseen percepts
from untidy or accidental stimuli can elicit unconscious processes which break the mould of habitual thought, whilst a temporal record of recent ideas makes it easier to change one’s mind at appropriate stages.

Chapter 1, "What is a Sketch?" limits the scope of the subject matter to the class overlap between three kinds of representation - those which represent two dimensional views of real or imagined objects, those "used for imagining something else" and those which possess certain "privileged" attributes of untidy indeterminacy (the "vagueness" mentioned in Stephen Scrivener’s quotation). I further limit the subject by considering only sketches intended to represent appearance and thus to thought processes used in art and design where imagined appearance is of primary importance. This last requirement is often inseparable, of course from sketch attributes in which function and appearance overlap. Although mostly common sense, the discussion which follows is necessary in order to put sketching in its functional context. I discuss two examples of sketching in practice, one from fine art and the other from design, to illustrate briefly how sketches are used.

Chapter 2 explores the origins and nature of the indeterminate attributes of sketches mentioned in Chapter 1. It summarises the history of the influence of Leonardo's recommended use of deliberate indeterminacy to "rouse the mind to new inventions" in the Trattato, and includes a section on the rather surprising influence of photography on sketch style. The absorption of sketch attributes into exhibited Fine Art "to enhance the beholder's share" towards the end of the nineteenth century is also briefly discussed. I consider this relevant because it helps us to understand how sketches work and to understand why the style of much clever but over determinate "pseudo realistic" computer graphics is alien to the interests of many artists who feel an affection for this long tradition of imaginative understatement. The Chapter owes a debt to the work of others, especially Ernst Gombrich and Aaron Scharf. However the point that Leonardo's recommendations about indeterminacy imply two separate functions - visual concept innovation and visual refinement of an existing idea is my own extension of Gombrich's treatment. The application of Black's distinction of three kinds of indeterminacy to sketch attributes may also be a "find". The treatment of the importance of sketch size as an attribute related to size in biology and later in chapter 6 to research on global or local precedence in selective attention is my own.

Chapter 3 considers the evidence for Hypothesis One. The chapter originally included a long section on the nature of representation which I decided had to be cut. I take it for granted following Palmer (1978) that the concept of a
"representation" is only meaningful as part of a "representational system" which is an ordered triple of the represented world, the representing world and an interpretive process which can map one onto the other. The discussion considers several complications to the "analog" versus "propositional" debate and makes some new proposals. Under the influence of the philosopher C. S. Pierce I consider that the way in which a representation is generated is almost as important to sketch theory as how it is interpreted. Before reading Pierce I had conceived a distinction between "semantically" and "non-semantically" generated representations but later gave way to Pierce's related notion of "contingent" and "non-contingent" representations. The idea of treating descriptive (a term I prefer to "Propositional") or depictive (preferred to "analog") representational types as instances within a multidimensional continuum (or within a prismatic conceptual space) with intermediate types such as sketches having a special function of "catalysing" type translation, is my own. I have of course acknowledged a debt to others here especially to C. S. Pierce and to psychologists such as Stephen Palmer, but the "continuum" treatment with the inclusion of a "sparsity - richness" dimension is my own. I really do not know if the significance of intermediate representational types for mental type translation has been seriously discussed before. The idea is mentioned briefly in Fish and Scrivener, 1991.

Chapters 4 and 5 discuss evidence for two further but related functions of sketches summarised by Hypotheses Two and Three. Chapter 4 examines the hypothesis that "Leonardo type" indeterminacies exploit automatic recognition mechanisms which give the user inventive access to her or his unconscious visual memory (Hypothesis Two). It includes a discussion of the relevance of the three main classes of picture perception theory to sketches. I give reasons for preferring a "late construction" theory of picture perception in contrast particularly to James Gibson's "invariant pick-up" theory. The value of different theories of recognition to an understanding of sketch attributes is then reviewed. Both Hoffman,'s and Biederman's theories of "recognition by parts" (which are related) are used to explain how sketch incompletions may trigger automatic but useful memory retrieval mechanisms. The evidence for and explanatory value of the theory of context schemata is also discussed in relation to sketch indeterminacies. Finally I suggest a further stage to traditional accounts of perception as a process of "going beyond the information given".

Chapter 5 discusses in some detail the background evidence for Hypothesis Three - that sketches are hybrid images, partly sensory information from the eyes and partly superimposed imagery generated from memory. I review the experimentally determined properties of mental imagery within the context of a
model proposed by Kosslyn and Schwartz and consider the evidence that imagery is an important mental tool for visual invention. The hybrid-image theory of conscious perception is developed as a variant of Marcel's "late-construction" theory of perception. (Chapter 5) The idea of a hybrid between a percept and a superimposed mental image is not new and I have reviewed and fully acknowledged the precedents of this idea. The idea that mental images owe their origins to perceptual completion has been hinted at by Kosslyn and others but not developed in the way I have attempted here. Nor, I claim have the evolutionary implications of the hybrid-image theory of perception been pursued before. Every "hint" I have found that others are thinking in the same direction I have quoted and acknowledged. The idea that normal percepts are late spatial hybrids between stimulus information and an unconscious mental image seems so obvious to me that I would not be surprised if it has been suggested before though I have not come across any clear statement of it in my reading. However the model mechanism for hybrid-image formation with a triple buffer and selective "gate" and the comparison with binocular fusion or rivalry is entirely my own idea. I also claim to have noticed some rather elegant hybrid-image explanations, based on this model, for well known experimental results in perception and imagery. Amongst other fascinating relevant "finds" which fit the hybrid-image model are the work showing the non-reversibility of mental images of ambiguous drawings (Chambers and Reisberg, 1985) (the percept component is needed) and the PET studies by Kosslyn et al. (1993) showing that both perceptual and mental imagery versions of a common task activate the same parts of topographically mapped primary visual cortex and that incomplete or confused visual stimuli activate these visual parts of the brain more than their completed counter-parts. A discovery which surprised its authors.

Chapter 6 discusses the different types of selective attention that sketch attributes support according to Hypothesis Four. Although I know of no systematic treatment of sketches in this context, the idea that sketches support selective attention seems to be too obvious not to have been thought of before. However I believe this chapter contains some real inter-disciplinary "finds" about sketch function that invite further research. One is the relevance of Garner's distinction between Integral and Separable perceived dimensions to the attentional needs of artists. Another is the published work on the ability to attend selectively to different spatial frequencies and the close resemblance of some sketches to spatial frequency filtered images. Yet another is the fascinating series of papers investigating the factors which influence global or local precedence in perception and attention. This relates to the problem of how the visual system decides that
certain arrangements of stimuli constitute an "object" and my theory that sketch attributes can be used to bias the visual system into selecting the candidate "object" within a hierarchy of possible structure according to interest. Coincidentally this work provides a useful extra piece in the jig-saw puzzle raised in Chapter 3 about how an artist or designer determines the appropriate "size" for a sketch.

Chapter 7 examines the theory that sketches and the act of sketching support the users' conscious awareness of their own thought processes (Hypothesis 5). But why, (the reader may ask) should they need such visual support? This was the most ambitious and difficult chapter to write since I knew that in discussing consciousness I was treading on ground that has long been regarded as taboo in cognitive science. Recently however the climate has changed, as I show in my review of a number of theories of "consciousness" and of the distinction between automatic and voluntary mental processes. The debt that this chapter owes to authors such as Johnson-Laird, Norman and Shallice, Marcel and to Baars is fully acknowledged. Drawing on the theories of these authors but with a different emphasis I develop the idea of consciousness as a "Cognitive Monitor" which records the initiation and the consequences of cognitive acts for mental "back-tracking" and control. If this hypothesis about consciousness is not entirely new, I believe that at least the discussion of its application to the theory of the sketch and the specification of what it must do is new.

The idea of documenting such a possible function of sketching came from a number of diverse sources such as a remark by Philip Rawson (1969) that one should always be able to tell where a drawing began and ended ("Why? and How?" I asked) and an idiosyncratic little essay by Douglas Hofstadter entitled "On the Seeming Paradox of Mechanizing Creativity". Hofstadter points out the importance of conscious awareness for mastering those over-learned unconscious habits of thought that so often inhibit inventiveness. It seemed obvious that original invention presupposes the ability to remember consciously a chain of thought and to back-track in order to try new directions in a conceptual hierarchy of possibilities. How do sketches support this type of short term episodic memory and how might this support be improved? This led therefore to the need to include some discussion of the nature of creativity and the relative roles that controlled and automatic patterns of thought play in it. Whilst acknowledging (with quotations) my debt to the literature on "Creativity", especially the little book by Margaret Boden (1992), I have outlined some sketch relevant ideas of my own on this subject as well.

Chapter 8 discusses the inter-dependencies between the five proposed functions of the sketch with the purpose of finding some common unifying
principle worthy of the title "A Cognitive Theory of the Sketch". The attempt is not very successful. Perhaps this is because contemporary knowledge of the brain and of those cognitive processes I am claiming are crucial to an understanding of the sketch are still in such a rudimentary state. I take some comfort also from Minsky's (1985) suggestion that perhaps the search for a unified principle to explain cognition is simply "physics envy". After examining and rejecting the possibility that mental translation is a unifying principle of sketch function I prefer instead the possibility that all the five functions are related by the way they support the Cognitive Monitor. This is posited to control and monitor as slave systems the components of the Baddeley-Hitch-Logie model of working memory (summarised in Logie 1995). When this cognitive system is used for manipulating imaginary objects I call it metaphorically the "Inner Artist" which has as slave systems an "Inner Hand" (Mental image generator) an "Inner Eye" (Covert attention), an "Inner Sketch-Pad" (Spatial memory) and for descriptive thought an "Inner Voice" and an "Inner Ear" (Verbal working memory components). I am careful to avoid the infinitely recursive "homunculus" trap suggested by the "Inner artist" metaphor. This is really posited to be a limited capacity neural mechanism which can control and record a wide range of other processes used for descriptive and depictive visual problem solving. It is envisaged purely as an operating system, very complex and poorly understood but in no sense a Cartesian "mind within the brain". It is activated automatically and without awareness by certain types of data inconsistency which may occur during the continual processes of comparison between new sensory information and stored knowledge. The reason why the creative brain needs the perceptual support of untidy sketch stimuli is that the cognitive resources available to this "Inner Artist" evolved not to invent novel imaginary objects but to use memory in order to respond quickly and appropriately to confusing or unexpected stimuli from real objects. Sketches are artificial substitutes for such confusing or incomplete "real" events which awaken the "inner artist". This is the nearest to a unified theory of the sketch I can develop. The temporal sequence of real events is important. That is why the Inner Artist expects the sequence of sketch actions to be so also.

Chapter 9 discusses some implications of the above Cognitive Theory of the Sketch, incomplete though it is. The chapter is really an invitation to others to test the theory and to examine how new media can be used to improve on the ancient and, surely, very primitive medium of paper and pigment. The theory provides good reasons why many artists and designers continue to use sketch pads and pencils in preference to powerful but inappropriate machine "paint" systems in the early stages of visual invention. However it also suggests that paper and pigment is
far from an ideal sketching medium and that in many respects it provides rather poor support for the Inner Artist. I therefore outline as a broad framework a possible improved sketching system as a "Mind-Machine Synergism" in which machine processes and internal data structures are designed specifically to support not the media bound attributes of the sketch (as in the "Imitation Bronze" approach) but the five cognitive functions of the Inner Artist. This includes a collaboratively generated "Structure Map" of the sketch subject matter and machine assisted mechanisms for implementing the one to many mappings implicit in the different types of "vagueness" represented by sketch attributes. I am well aware of the difficulties in implementing such a system but consider that it signposts some quite realistic targets for further research.

Acknowledgements.

Because of its inter-disciplinary subject matter this essay owes a great deal to many writers from many different fields. The ideas expressed have been influenced by far too many authors and research workers to acknowledge them all here. In art theory I have been especially influenced by the writings of Ernst Gombrich and Rudolf Arnheim and in cognitive psychology by those of Stephen Kosslyn, Stephen Palmer, Michael Posner, Ann Treisman, Ronald Finke, Bernard Baars, Margaret Boden, Alan Baddeley, Robert Logie, Irvin Biederman, Glyn Humphreys, Jerome Bruner, Richard Gregory, Martha Farah and literally several hundred other authors all of whom I have carefully acknowledged in the text. In general if a statement or idea is not referenced it can be assumed to be my own.

I must thank CGCHE who paid all my post-graduate fees at Loughborough and allowed me the time for research whilst I was employed full time as a member of their staff. Some of the research referenced has formed part of a team research project - "Improving Computer Systems for Fashion Designers" funded by the Science and Engineering Research Council of Great Britain. I am grateful to my collaborators, Paivi Makirinne-Crofts, Sohrab Sadaat and William Godwin, on this project which has introduced me to the ethos of Fashion design and given me the opportunity to visit some design companies abroad.

I would like to thank Robert Welch for discussing his design methods with me and for letting me use some of his sketches. I would also like to thank Harold Cohen for corresponding with me in such a friendly way about his work and especially Gabriela Goldschmidt who corresponded with me about two of my papers and who sent me copies of her own complementary papers investigating the sketching methods used by architects.

Finally I must acknowledge my debt to Stephen Scrivener, now Director of Research at Derby University who suggested this project to me so long ago and
who has continually given me patient advice and encouragement even though he must have sometimes thought I would never finish it!

A Note on Typefaces.

I have deviated slightly from normal practice in the way I have used typefaces. "Italics" are used inside quotation marks, not for emphasis, but to indicate that the text is an exact quotation by another author from a reference given. This allows me to use normal typeface inside quotation marks for hypothetical questions, metaphors (i.e. expressions not to be taken literally) and for when I am quoting another part of this text. I have used underlined text, rather than the usual italics, for special emphasis of a word or phrase of the kind that one normally makes in speech by intonation, to clarify one's meaning.

Bold typeface is used for words or phrases which have a specialised or technical meaning which is explained either in the accompanying text or in a reference which has just been quoted. Bold typeface is also used for Chapter and Section Headings which are, I hope, easily distinguished from technical terms.
1. WHAT IS A SKETCH?

"The process of drawing is before all else the process of putting the visual intelligence into action, the very mechanics of taking visual thought. Unlike painting and sculpture it is the process by which [artists or designers] make clear to [themselves], and not to the spectator, what they are doing. It is a soliloquy before it becomes communication." (after Michael Ayrton, 1959 with "an artist" altered to "artists or designers" and "himself" altered to "themselves").

1.1 The Problem of Definition.

Even when restricted to the visual arts, as here, the word "sketch" is difficult to define without loss of usefulness. When defined too strictly many words lose part of their meaning in ways which damage their power to communicate or to assist in thought about the ideas with which they are associated. One way by which such words acquire their meaning is by reference to a family of naturally occurring instances of a vague class linked often by a common origin. This can result in a network of overlapping "typical" attributes which are not necessarily possessed by every exemplar of the set. "Sketch" (and "to sketch") are examples of such words. A definition may cover typical sketches but there are always cases which defy the definition. The very fact that a word is hard to define suggests that it refers to something poorly understood and stimulates thought about the nature of the object or concept to which it refers.

For example the word "sketch" is commonly used to refer to those rough drawings made by artists to represent the general outline of a preliminary idea. However a study of working practice shows that, whilst sketches often represent the outlines of visual structure without the detail they may also do the reverse and represent the details of a part without the structure (e.g. see Picasso's studies for Guernica, documented by Arnheim (1962)). Moreover colour studies with little drawing are often made by painters which can equally well be called sketches. Moreover, sketches are just as widely used by sculptors, and designers or, for that matter, by anybody who wishes to think or communicate visually.

For the purposes of this thesis I limit the meaning of the word "sketch" to include only two-dimensional visual representations of some real or imaginary scene or object. Thus models and maquettes used by designers for purposes similar to the sketch are not included. A non-figurative painting cannot be a sketch but a representation of a non-figurative painting can be. Also excluded here, are works using incompletion and ambiguity intended for exhibition rather than as ideas for further development (ruling out the Oxford dictionaries last sentence). However the attributes of such works are of relevance, due to their historical links with similar
attributes in real sketches. I am further limiting my attention to sketches which represent or are primarily concerned with visual appearance as opposed to the representation by visual means of the structure or function of something. There exists a fuzzy continuum of sketch types from those which might be used (typically by engineers or scientists) to think visually about a mechanism or function and are not concerned with the appearance of what is represented to those which are only concerned with appearance. The majority of sketch types probably fall between these two extremes, with visual appearance forming a part but not the whole of their content. For example, sketches made by architects of the ground plans and elevations of buildings explore the implications of spatial relationships for human use simultaneously with appearance. In sketches for the decorative arts appearance is the primary but not the only function. Thus a Fashion designer may use the same sketch to think about and represent the visual profile of a garment and the non visual spaces between the users body and the fabric (Illustration VI). Here I will concentrate on sketches whose function is to represent the possible visual appearance of something as seen from a particular viewpoint. In other words they will be viewer-centred representations (as defined by Marr, 1982). It is necessary to distinguish conceptually sketch functions which represent visually thought about non-visual ideas as diagrams often do, from those which concern thought about the visual appearance of what is represented. This does not mean that it will always be easy or even possible to separate these two types of function in individual sketches. So this enquiry will be limited in scope not only to a subset of a subset of the set of visual objects which can legitimately be called sketches, it will also be limited within this subset to only those attributes which concern visual appearance. The reason for this constraint is that I believe that progress in understanding how we perceive, recall, manipulate and generate the appearance of real or non-existent objects within the mind, is central to understanding sketches. Cognitive psychology has proved to be a fertile source of ideas to explain what would otherwise be mysterious attributes of familiar sketches. I shall also claim that ideas suggested by this investigation of sketches have implications for existing theories of visual imagery and perception.

As already mentioned, the sketches discussed here are limited to two-dimensional representations of objects or scenes as seen or imagined from a particular viewpoint. It is necessary to further distinguish such sketches from other viewer-centred images such as photographs, drawings made for their own sake and paintings. The interesting constraints on meaning are those which determine how sketches differ from more finished and permanent kinds of visual representation.
1. What is a Sketch?

Three possible, fuzzy "semantic descriptions" of the word "sketch" are listed below.

1. Sketches are representations used to assist in the visualisation and mental manipulation of some unfinished or imaginary object. Unlike other drawings, illustrations, or photographs which can be valued for their intrinsic merits, sketches should only be understood and valued in relation to the objects they represent and the ideas to which they contribute.

2. Sketches are drawings, colour studies or illustrations which possess the attributes of untidyness, incompleteness and ambiguity. They typically give the appearance (often false) of spontaneity and often occur in multiple series related to a common subject matter. In contemporary art these attributes are valued for their power to suggest in the mind of the viewer visual content which is not explicit.

3. Sketches are representations used to assist the creation and experimental manipulation of an unfinished or imaginary object. They accomplish this by virtue of privileged attributes which may appear superficially to be superfluous or unhelpfully untidy. However, these attributes have concealed cognitive functions which amplify the user's capacity for visual invention.

The first meaning above, attempts to isolate a function or set of typical functions which is unique to sketches and which distinguishes them from the other types of representation mentioned. This description concentrates attention on the working practices of artists and how sketches are used. The second meaning attempts to define the sketch in terms of its characteristic attributes. This approach assumes that whilst sketches may share some of their attributes with other representations such as non-sketch drawings, illustrations, maps, diagrams etc., they also possess some distinguishing typical attributes not usually found in other types of representation. I shall refer to these candidate defining attributes of the sketch as "privileged" sketch attributes. As hinted earlier I have used the word "typical" repeatedly to indicate that I regard both these approaches as leading at best to fuzzy semantic descriptions rather than clear definitions. The third hybrid description of what a sketch is encapsulates those properties of sketches that are of special interest to this study. It represents, of course, only a subset of the possible meanings. This adopted meaning has many conceptual flaws. However it serves the ultimate goal of the study, which is to design new machine sketching systems genuinely useful to artists and designers. Treating artists sketches as images which serve a particular function leads to the study of how artists work and think. How is a new visual object conceived, manipulated, refined and completed? Treating
sketches as images with particular attributes leads to the study of the images created by artists, what features they possess and how they can be classified. What are the privileged attributes of sketches and how did they originate? The third hybrid description leads, in addition, to the study of the relationship between the functions of the sketch and the attributes it possesses. This relationship is far from clear. The familiar attributes of sketches are constrained not only by function but also by time and skill and by the limitations of the traditional media with which they are usually associated. Many of the attributes have evolved gradually over centuries of trial and error by many artists. They possess cognitive functions the origins of which are hard to trace and which may work by mechanisms that their users themselves do not understand. This is exactly the problem which must be faced by the designers of new sketching media.

1.2 The Problem of Defining Function.

The first meaning, in terms of a defining function seemed appropriate to me at the start of this research. A sketch, I thought, could be any visual representation which was used in the design or invention of something else. True sketches were of interest only to the extent that they helped their user conceive, create and refine the object to which they referred. This had the advantage or eliminating those objects sometimes called sketches in museums, as for example the Oxford Dictionary's "a drawing or painting of unpretentious nature". However it gradually became apparent that this description was inadequate. It is difficult to circumscribe the possible functions of the sketch. It is clear that there are other secondary functions which although they may be less use in characterising the sketch may be impossible to separate from the primary function. Thus besides the private and personal function of assisting the visual thought of their user, sketches have two other functions.

1. Sketches may be used for recording visual information and visual thought for later use even though this use may be unspecified. They may also be stored to record stages of one design or composition for specified later use in a different task. Whilst the primary function is concerned with a task immediately active in working memory, the recording function must support long term memory.

2. It is clear also that especially in design, sketches are used to communicate visual ideas to others. However it does not follow that because sketches record a private dialogue within the user's mind they must also be useful for communicating to others.

Both communicating and recording functions are logically distinct from the private visualising function even if they are often combined with the latter in
1. What is a Sketch?

practice. It may be due to a limitation of traditional media that these two additional functions should so often be mixed with the primary purpose of sketches. It is therefore important to separate the three functions conceptually and to consider to what extent their requirements may conflict with one another. However convenient it may be to use rough sketches for all three functions at once it may prove advantageous with new media to separate them. The attributes which best assist the generation of visual ideas are not necessarily those which best record or communicate these ideas. It may then be necessary to distinguish more clearly and formally the "visualising sketch" from the "recording sketch" or the "communicating sketch". However at present, in sketches as commonly used by artists and designers, several functions are confused and it is premature to attempt to define the sketch by function alone. Moreover it is impossible to constrain meaningfully the possible attributes of "sketches" from a definition of function. This is because we need a better understanding of the subtle mechanisms by which sketch attributes interact with their users' minds.

1.3 The Problem of Defining Attributes.

The second description of the sketch in terms of special or privileged attributes avoids the problem of the mixed function of many sketches. However, although an attribute definition may be useful for museum keepers and art lovers it seems quite inappropriate as a guide-line for designing new sketching systems. Moreover it is very difficult to select from the many typical and familiar attributes of sketches those which are obligatory. Possible distinguishing attributes may be divided into two groups, (a) those associated with the sketching medium and the way sketches are generated and (b) those associated with the media-independent, culturally understood meanings of sketch attributes. (A distinction understood by Leonardo when he compared drawing to poetry see 2.1)

Sketches are usually, but not always smaller than other images. The media on which they are usually produced is cheap and portable. The vast majority are made on paper with pencil, crayon, ink or water-colour. However painters traditionally make working studies, called sketches in oil paint on canvas. Commonly, sets of sketches, often related to the same subject, are collected together in "sketch books". The crucial and inter-related questions raised by the traditional media of sketches are

- To what extent are the desired attributes of the sketch constrained or handicapped by the medium?
- To what extent have the medium and its properties been influenced or selected by the desired attributes of the sketch?
1. What is a Sketch?

Typical visual attributes are listed below.

- Sketches are commonly drawings and use lines to represent form. The lines used have a number of culturally acquired meanings. Arnheim (1974) has distinguished three:
  
  **Object lines** which are perceived as one-dimensional objects (or lines in the thing represented)

  **Hatch lines** usually in parallel clusters, which are used either to represent shadows or the shape and orientation of contiguous surfaces.

  **Contour Lines** where a line embracing an area is seen as enclosing a visual object. Arnheim refers here to the segregation of simple images into figure and ground, much investigated by the Gestalt psychologists (Koffka, 1938, Kohler, 1947, Ellis, W. 1938). However it is necessary to subdivide such lines into different types. For example, contour lines may represent the "self-occluding contour" (profile) of a three-dimensional object or they may represent edges of adjacent surfaces at different orientations within an object or they may represent the boundary of a region of contrasting colour within a common surface.

  A detailed discussion as to how lines represent is beyond the scope of this study. However the question will be referred to again within the context of sketch perception.

  Not all sketches are drawings. Sketches can also be paintings or colour studies with no lines and only vague indication of form (e.g. Turner's colour sketches, (Illustration XXXV).

- Sketches produced early in the process of design or composition may have superimposed written notes. Although this attribute is an important clue to sketch function it does not distinguish sketches from diagrams and scientific illustrations (Illustrations IX, XIV, XXIII).

- Most of the attributes which can claim to be special to sketches are forms of visual indeterminacy, many of which appear, perhaps misleadingly to be accidental.

*Indeterminacies of under-representation*

  For example

  - Incompletion, where the drawing fades away to an empty space or a contour line is missing, fragmented or only suggested (Illustration X)
1. What is a Sketch?

- Limitation of the visual attributes, features or parts which are represented. Some attributes or features of the represented object or scene are emphasised and represented in more detail than others which may be represented vaguely or not at all (examples to be discussed later).

Indeterminacies of over-representation

for example

- Multiple alternative contour lines ("pentimenti")
- Suggestive scribbles and cross-hatching
- Confused smudges and incomplete erasure
- Blobs and accidental flow patterns of paint or ink
- Mysterious dark shadows (e.g. Leonardo's "sfumati")

(see Chapter 2 for examples and discussion)

When these and other attributes of sketches are considered without reference to their possible function or to their historical origins it is difficult to distinguish those which might be genuinely intrinsic to sketches and those which are accidental and unimportant or even undesirable consequences of the context in which sketches are generated and the traditional media used to create them.

In this preliminary listing of functions and attributes I have tried to avoid the temptation to suggest sketch attributes from observed functions or to deduce functions from observed attributes. I shall now consider what attributes sketches and sketching systems might be expected to possess if their primary functions are those I have suggested.

1.4 Experimentation and the Need to Visualise in the Arts.

The primary functions of the sketches discussed here are to visualise a nonexistent object or to predict in advance the visual consequences of changes to an existing object. The two functions are frequently combined as when one imagines a non-existent object and then also attempts to imagine the effects of changes to that object. Sketching is not essential to the creation of a work of art or visual design. The designer "R.T." for example claimed never to sketch and so also do a small minority of Fashion designers (Bates et Al, 1988). Painters also frequently work directly on the canvas without preliminary studies. In order better to understand the role of visual monitoring in the creation of visual objects it is worth considering briefly how new visual objects might in principle be created "blind" without alteration until at least the late prototype or detailed refinement stage of development.

It is conceivable that a set of rules for design or composition might be so rigorous that they can be applied systematically without any need to visualise the
designed object before it is finished. In the visual arts of Western Culture today, such a way of working is rare. Exceptions can be found, for example in the work of the "Systems" artists of the nineteen seventies, who applied rules involving number systems to the structural composition of abstract paintings. However such rules usually included an element of surprise or randomness and it is not clear that advance visualisation was not used to select from the large number of compositional permutations such number systems typically generated. (See The "Chance and order" drawings by Kenneth Martin, K. 1975). (Another exception is Harold Cohen’s sixteen year old expert system "Aaron". Using only a few random parameters (sometimes only one) this program uses Cohen’s rules to generate non interactively an infinite possible number of drawings. Cohen described Aaron as a "program designed to investigate the cognitive principles underlying visual representation" (Cohen 1974). Until Cohen publishes in detail his "rules" it will not be clear how far he has succeeded in his objectives. It is certainly an ambitious and interesting attempt by an artist to code explicitly knowledge related to personal visual style. (see Illustration XXXVII and Chapter 9.3))

The difficulty of generating interesting visual art or design by any form of rule, however sophisticated the knowledge on which it is based, is linked to cultural context. In ancient Egyptian tomb painting for instance the descriptive function of visual representation appears to have been strictly defined (Abbate 1972). The style of representation, similar in some respects to Egyptian pictographic writing, was constrained by culturally imposed standards, equivalent to rules. Despite its frequent beauty, Egyptian art allowed little scope for stylistic novelty or individuality. Both Egyptian and medieval Islamic art (Critchlow 1976) were constrained by visual conventions determined by the philosophy and the cultures they served. The religious art of both cultures demonstrate that freedom from cultural constraints and individuality of style are not essential for the development of durable and memorable art. In contrast to the relative permanence of earlier cultures, our own post-renaissance Western culture is in a state of accelerating change, perhaps to its own detriment. This can be traced to the growth of knowledge and the technology which results from such growth. Arnold Hauser (Hauser 1951, volume 4, Chapter 4) writing of the effects of technology on impressionism made an ominous deduction. "For the rapid development of technology not only accelerates the change of fashion, but also the shifting emphases in the criteria of aesthetic taste; it often brings about a senseless and fruitless mania for innovation, a restless striving for the new for the mere sake of novelty." and a sentence later "The continual and increasingly rapid replacement of old articles in everyday use by new ones leads however to a diminished affection
for material and soon also for intellectual possessions, too, and readjusts the speed at which philosophical and artistic revaluations occur to that of changing fashion."

Whether one agrees with this evaluation of the effects of technology or not, it is clear that the degree to which an artist or designer needs to visualise in advance the multiple possible appearances of an object before it is finished, is greatly increased by the need to experiment. This need in turn is determined by the degree to which the context within which an artist works, causes her or him to value individuality and novelty in a work as opposed to those tacit values acquired by learnt tradition and cultural context. The need to experiment is also dependent, of course, on the type of art that is practised, and in design by the client and the market. It is often easier commercially to design by precedent, modifying existing successful prototypes, than to indulge in a time-consuming search for visual originality. Another consideration is the degree to which the tacit expert knowledge possessed by artists, as the result of experience, can be applied automatically to new but related problems and used to decrease experiment or visual manipulation.

Whether or not the word "experimentation", when applied to the visual arts, has a similar meaning to its use in the sciences, is a question discussed by Gombrich (Gombrich 1982, Chapter 9 "Experiment and Experience in the Arts"). He concludes that the word cannot be applied in the visual arts exactly as it is in the sciences due to the problem of how the products of such experimentation might be assessed. However there is a sense in which artists may claim to be experimenting. Gombrich makes the point that historically the meaning of experiment in the arts depends upon the variety of functions and the variety of cultures that an art has served. When an artist manipulates perceived or imagined appearance, this is a kind of experimentation which differs from that of science in that the criteria for assessing the results are private and rarely explicit. The need for this form of experiment depends upon the task and its cultural context. The higher the value placed upon originality and invention and the lower the value placed upon past experience or stylistic convention, the greater is the need for such experiment with continuous refinement and correction.

1.4 Three Visualizing Methods for Experiment in the Arts.

The commonest working method used in fine art and in those types of design where appearance is critical, consists of visual experiment (as described above), in a series of stages, by a process of gradual development by step-wise refinement with repeated back-tracking when a line of thought is discarded. Typically an initially vague idea is constrained by privately imposed propositions and, in design especially, by criteria defined by the function and imposed by the clients. However carefully conceived the initial idea, the number of possible
alternatives in the early stages of visual thought will be enormous. Therefore continuous visual comparison is necessary in order to select from the many alternative ideas and modifications of ideas that are possible. Three ways of doing this are available to artists

1. The visual object itself (or a prototype of the object) can be roughly fabricated, manipulated in stages, altered, completed, destroyed and recreated whilst it is continuously inspected. This is a popular image of the fine artist at work in the studio or before the motif. It is obviously only possible if the medium in which the work is created is easily manipulated and relatively cheap. Painting in oils is the obvious example. Paint can be easily scraped off while wet and replaced or painted over when dry. A painting can be developed in stages on the canvas in such as way that the artist's attention is drawn either to global structure (broad washes and loose outlines) or to fine detail. This way of working out an idea directly on the canvas, became fashionable with the Impressionists who painted in front of the subject with the object of capturing spontaneous and fleeting visual experience. It may also be slow and laboured. Emile Bernard, watching Cezanne at work saw him carefully place a single brush stroke of paint on his painting, stare at the painting intently for twenty minutes, then very carefully scrape the brush stroke off (Perruchot, 1961). Clay is another medium which has the properties necessary for rapid manipulation and alteration. A potter for example can test many shapes of pot at the wheel. Not only can changes of profile be quickly adjusted but a skilled potter can throw many related pots in a few minutes for visual comparison. Indeed speed is necessary as the clay becomes too soft and wet to hold its shape unless manipulated quickly.

It is clear that visual experiment by manipulation of the desired object itself, is limited to a few specially adapted media. The scope of this limitation expanded dramatically during the industrial revolution which brought about the much discussed separation between the designing and the making of our cultural artifacts (Lawson, 1992). When a visual prototype is too large and expensive or when the designed object is intended for industrial manufacture then this method of "experiment by making" is impractical. A painter would think twice before working on a large mural without having already completed many preliminary visual experiments and decisions (see Picasso's "Guernica", Arnheim 1962). It is prohibitively expensive for an architect to "experiment" by erecting large buildings only to see them torn down and rebuilt. When this happens it is a disaster. A large sculpture in marble cannot be made in many versions in order to select the best design. However, even when the chosen medium is designed for easy manipulation, it must be noted that direct visual experiment on a single object is
1. What is a Sketch?

not an easy way of working. Consider for example Cezanne's single brush-mark experiment mentioned above. In some respects it resembles a scientific experiment. Cezanne compares his "sensations" (his own word) before and after making a single change, holding all other factors, which might influence this visual sensation, constant. In a scientific experiment however there would be for comparison, in addition to the manipulated object, a control object in which everything is kept exactly the same except for the variable of interest. In other words Cezanne should have had at least two identical paintings for comparison. Since this is impractical the comparison necessary for making every small spatial or chromatic "before and after" decision must rely upon the vagaries of memory.

2. The second method by which artists and designers can conduct visual experiments on non-existent objects is simply to use visual imagination. The capacity to form mental images of remembered or imagined objects and to mentally rotate, re-scale, recombine their parts and manipulate them in other ways has been widely investigated in the last twenty years (see Chapter 5). Interesting research has been done on the possibilities of using mental imagery to invent new visual objects (Finke and Slayton, 1988, Finke, 1990). However there is astonishingly little recent published work on the role of mental imagery in the visual arts. In a review entitled "Imagery in the Arts" Lindauer (1983) reports that most studies of imagery in artists agree that "sketches are tried out in the mind's eye before they emerge onto the canvas" (Cohen, 1976). In studies based on interviews most artists reported that they had such images both in real life and in their work (Henderson and Lindauer 1976, Bates et Al, 1988). It is therefore safe to assume that artists use mental imagery regularly to assist in visual decisions in both the early inventive stages and the later development stages of visual creation. The point at issue however is how far it is possible for an artist to rely on his or her mental visualising with "the mind's eye". The ability to form mental images and their vividness varies widely between individuals (Marks, 1972, Kosslyn et Al 1984). However most people report that their mental imagery is difficult to hold, is limited in clarity and field of view and is often vague or lacks detail (Kosslyn, 1980 see 5.2.7). It has also been shown that it is difficult to form a mental image of attributes of a previous percept that were not consciously attended to (Reed & Johnsen, 1975, Kunen & May, 1980). In short there is a limit on the ability to mentally control and manipulate an imagined object. It is especially the unexpected interactions and relationships which occur in actual perception, which are likely to be missed in a mental image. For example startling colour contrast and after-image effects occur in colour perception (see Albers, 1971), which do not occur in mental images (Finke, 1980, 1989, Finke and Kurtzman, 1981b). It is necessary to
perceive conjunctions of real colours in order to detect the unexpected subtleties of colour contrast effects. Moreover, as will be discussed later, accidental and unexpected form or colour, which results from the vagaries of physical media, stimulates inventive discovery for the very reason that it is outside mental control.

In many ways, the advantages and disadvantages of visual manipulation by perceiving physical media are complementary to those of mental imagery. Relationships perceived in real vision are vivid, detailed, stable and accurate. The skilful manipulation of physical media enables the discovery of unexpected visual relationships impossible with imagery alone. On the other hand the use of physical media for controlled visual comparison is either prohibitively slow or too difficult. In contrast mental manipulation is as fast as thought and unconstrained by physical limitations. But it is fleeting and a sequence of mental manipulations cannot be easily recorded. They must be remembered, and human memory is notoriously fickle. It is clear that perceptual and imagery manipulation need each other. To manipulate a physical medium without imagining in advance the myriad possibilities that need never be implemented would be useless. Cezanne’s single brush-mark could have been selected from about ten million noticeably different colour shades (Judd and Wyszecki 1963). At twenty minutes each it would take 380 years to try out all the possibilities of a single brush-stroke. Even when tightly constrained to a small range of shades and to the already determined spatial structure of the composition, the number of possible permutations of the many brush-strokes which make a painting is astronomical. Similar limitations apply to the experimental testing of small variants of a line made while drawing. It must be assumed that, like other artists working directly with a physical medium, Cezanne imagined and rejected many hundreds of brush-strokes for every one that he actually "tested" on the canvas. These would be fleeting images quickly forgotten.

Some Fashion designers claim that, once they have seen and handled a fabric, they can mentally visualise a complete garment in three-dimensions without sketching or modelling toile (Bates et Al 1988). Despite the rare claims made by a few however, it is difficult to believe that inventive and authoritative art or design is often created by manipulating images in the mind only, without simultaneous testing the ideas with perceived representations.

3. Despite the sparsity of direct evidence, there are strong a priori grounds for believing that it is the interactive combination of both perceptual experiment and mental imagery that most often leads to visual invention. Thus the third and most practical method of creating new visual objects is a compromise forced upon artists by the difficulties of manipulating prototypes and the frailty of mental imagery. A special purpose representational system is needed which can be
1. What is a Sketch?

Manipulated easily and fast enough to keep pace with visual thought. Ideally this system should allow the simultaneous combination of imagery and perception in harmonious collaboration (See Chapter 5). Such a system must be easier and faster to manipulate than actual prototypes of the objects for which it must substitute, whilst at the same time it encourages, clarifies and amplifies the users private mental imagery. Some of the attributes needed by the medium of such an representational system can be deduced from this requirement.

For instance the ideal the medium should

a. - be cheap and easily obtained
b. - be portable and robust because visual thought occurs in many places and contexts.
c. - allow the fast creation of multiple images due to the large number of visual concepts and modifications of visual concepts that may need to be considered.
d. - allow the easy erasure and alteration of images.
e. - be easy to use and to learn to use.
f. - not interfere with the user's main cognitive task and its related imagery and percepts.
g. - support the easy storage and recall of large numbers of related images
h. - support the simultaneous viewing for comparison of many related images.
i. - record the stages in the generation of a visual object or ideas for them so as to allow easy control of backtracking and correction.
j. - record verbal or descriptive modes of information as well as depictive representations and implement and convenient system of relating the two (labelling the structure depicted, see Chapter 3.5).

The traditional and still most widely used medium for such a function is the paper sketch-book with drawing instruments such as pencils or pens and combined, when colour is needed, with crayons or paint. This medium fails to meet all the proposed requirements. Indeed it seems at first sight that the requirements are in irreconcilable conflict. The paper sketch-book meets requirements a and b. It is cheap and portable (per image), provided of course that very large images are not needed. It meets requirements c. and e. provided the user is able to draw well enough to support her or his visual thought. Requirement d. is not very well catered for. Pencil and charcoal lines can be erased partially using a rubber but ink, crayon or water-colour only with great difficulty. It meets requirement e. provided again that the user can draw. I shall discuss requirement f. later. Here I
1. What is a Sketch?

...will merely note that whilst the ordinary pressure sensitive pencil is good in some respects, in use hand and pencil must interfere with part of the user's view of the drawing. Paints and palettes cause cognitive interference by dividing attention (Chapter 6). Requirement g is partly met. Sketch-books are easy to store but unless their user has implemented a system of indexes, which is unlikely, the retrieval of a partly remembered earlier sketch may be tedious. Requirement h is met quite well but at a cost to requirement g. Many pages of a sketch-book can be torn out and pinned together on a wall or spread on a table top. This is an advantage of paper and pencil over the single screen displays used by affordable machine sketching systems. The pages of a sketch-book provide a crude way of meeting requirement i, only as far as whole sketches record stages of the design or composition progress. However paper has no memory and it is obviously impossible to record in detail the within sketch sequence of decisions. Rawson (1969) argues that in drawings by skilled draughtsmen the marks themselves contain sequential information. He claims for example that he can usually tell by inspection where a drawing started. This is not equivalent to a detailed sequential record. The reasons why such a record may be desirable will be discussed later. Finally the sketch-book meets requirement j. but only crudely. As already noted artists frequently write scribbled notes on or beside their drawings. The connection between the written note and a sketch part must be indicated by a line or arrow which naturally interferes with the depictive part of the drawing.

Thus as a visualising medium, paper and pencil or the sketch-book though obviously convenient and widely used have many short-comings. They score about five out of ten for my suggested requirements. (for a and b, - 2 points, half points only for c, e, f, h, j, g.). However the points they lose on - fast manipulation and information recording - are important. Computer sketching and image processing systems have the potential at least to score much more highly on these points. Existing design specific "paint" systems (used as a medium for sketching in the early stages of design) score highly on c, d, g, but lose on h and score less well on the others. However, as I shall argue later, improved design resulting from a better understanding of sketch function could radically improve this score. As portable visual computers become cheaper and more powerful, the quality of imagery and the capacity for fast manipulation, improved image storage and detailed event recording continually improve. As a visualising medium, even in the very early stages of visual invention, the computer is surely destined to out-perform paper in almost every respect. (Admittedly requirement h is hard to meet. Cheap, very large area, thin display systems may come in the near future).
1. What is a Sketch?

Why is the humble sketch-book still in such widespread use? Why are computer systems only used by artists who specialise in computer art and by designers mainly for the later illustrative and marketing stages of design (Scrivener and Clark, S. 1992, Makirinne-Crofts et al. 1996). The reason is only partly that existing systems (in 1995) are too expensive and inconvenient to use. A more profound reason lies in the fact that the design of sketching systems in such a new medium raises questions about the cognitive functions of sketch attributes which have never been convincingly answered. How do the marks made in the traditional sketch represent their subject and how do they interact with representations in the brain of their user? Clearly, with sketching as described above, the medium is not the "message" (McLuhan 1969). If the representational system is medium independent it should be possible to transfer it to an improved medium without losing any of its crucial attributes. In other words system designers could adopt the "black box" approach to mental processes, assume that the user knows best what type of representational system to use and simply imitate the sketching attributes in a supposedly more powerful and flexible medium. This is the "idiot slave", "partitioned" approach to C.A.D. characterised by Negroponte (1977). According to this approach, all the user's thought processes are outside the province of the sketching system and its design. It is simply a medium which acts as a "tool-box" with which the user can implement his or her desired imagery as flexibly and perfectly as possible. Such an approach to the design of sketching systems in a new medium, only makes sense if the medium independent attributes of a visualising representation can be fully distinguished from those attributes which arise either accidentally or intentionally from the medium in which they are generated and manipulated. Unfortunately this is not the case with traditional media. Although it is clear that the sketching medium of paper and pencil or paint does not determine the representational system and its meaning, it constrains and enriches the functional attributes of the sketch in ways that are familiar but poorly understood.

In view of the deficiencies of paper as a medium for the fast manipulation and recording of images, it would be easy to jump to the conclusion that many of the characteristic attributes of paper sketches are undesirable consequences of an inadequate medium. Is not the incompleteness and sparsity of many sketches, attractive as it may be, a simple consequence of the difficulty of hand and eye to generate percepts on paper which can keep pace with thought? Are not hazy or multiple contour lines an undesirable consequence of the tedious difficulty of erasing the unwanted lines? Are not wobbly lines an undesirable result of a wobbly hand trying to control lines intended to be straight or perfect curves? Are not
untidy smudges and ambiguous scribbles caused by the users impatience with an imperfect medium? When water-colour on paper flows over the contour of a shape which is intended to be coloured, is this not a flaw in the medium of representation?

In short, it appears at first sight that many of the characteristic attributes of sketches are a hindrance rather than a help to sketch users who need a clear and accurate vision of the effects of their experiments on the appearance of an imagined object. What the artist or designer needs, according to this argument, is to produce as quickly and as early in the design process as possible, detailed, accurate and completely realistic images. If it is impossibly expensive, difficult and slow to manipulate life-size replicas of the object itself, then surely, the next best alternative is to manipulate ersatz ("virtual" is the term in vogue) objects. However where these are images of an imaginary object then the hypothetical ideal medium must emulate by precedent and calculation the light pattern which would fall upon the retina were such an object to be illuminated and placed in front of the user. Since photographs (including their descendants film and television) are the cultural images closest to this, the thoughtless term "photographic realism" is sometimes used for this ideal. So great is the potential of computer image handling systems for interactively manipulating the multiple appearances of three dimensional objects that the question "need artists and designers sketch at all?" may well be asked. In other words is it not possible and desirable to substitute detailed machine assisted realism for the clumsy drawings and hazy mental imagery inherited by tradition from the pre-computer age? This is not a completely vain question. So far the "photographic realism" school for the design of CAD systems has had things all their own way. A new generation of increasingly realistic, design specific visualising systems are already in common use. The discussion of such systems and their possible improvement must be deferred to Chapter 9 after a theory of sketch function has been presented.

We need first to investigate how traditional sketches represent real or imaginary objects and by what mechanisms they may assist visual thought. Sketches are only a tiny, visible component of a much richer private dialogue which artists and designers conduct in their minds. To understand the functions of the sketch it is necessary to consider what kinds of representation and mental process are necessary. It is also necessary to ask what resources are available in the brains of visual artists which enable visual "thought". These resources are enormous, but they are certainly limited. Their limitations arise, I will argue, from the recentness in biological time of the cultural need to imagine future man-made objects and the consequently "improvised" nature of the neural resources available
to the brain of an artist for this purpose (see Chapters 4, 5 and 8). What are these limits and how can they be better supported by the new media and technologies now available to our culture?

As a preliminary, it is useful to examine the historical origins of some sketch attributes and to illustrate briefly how sketches are used in practice. Then I shall attempt to use contemporary theories of cognition to develop plausible hypotheses about how "privileged" attributes of sketches serve their visualising function. This is the approach which I consider most appropriate for this study. For there is a gap in our current understanding of sketches and other visual representations which is handicapping the way we design and use new visual media. I have tried to make a bridge across this gap in understanding without losing sight of the practical objective. The unprecedented image handling capacity of computers poses a problem for the design of sketching systems for the visual arts which has not yet been solved. Before machine functions can be properly designed it is necessary to understand the relationship between the familiar attributes of sketches and their function. For these attributes are constrained and modified by the nature of the traditional media used by artists - pencils, pens, brushes, crayons, paints, inks paper, canvas etc. They are also constrained by the hurried, casual and private nature of many sketches. Despite this, when skilfully used, their untidy attributes have the power to excite the imagination in ways that other types of image cannot emulate. It is therefore not at all clear to what extent or by what mechanisms the "privileged" characteristics of sketches are essential to the visual purposes of the sketch or to what extent visual thought is handicapped by the constraints of the medium.

1.5 Sketching from the Motif and from Imagination.

Most figurative artists draw and colour continually from real scenes and objects and this activity is often referred to as "sketching". Designers may also sketch from motifs as a source of ideas. However when designing a visual object which does not yet exist, designers and artists alike can draw only upon imagination, memory and of course the sketch itself. Thus the difference between sketching from the motif and sketching from imagination appears to be fundamental. Compared to perception visual memory and imagination are vague and ephemeral yet unconstrained. A drawing from the motif can be corrected with new information for as long as desired by using the eyes. Nature provides precise constraints but also surprises of a different nature from those available when sketching from imagination. Because it is biased by expectation or remembered prototypes, drawing from memory tends to be less surprising (and thus less interesting) than drawing from nature. Artists sometimes refer to this as "false"
1. What is a Sketch?

versus "true" drawing (Ayrton, 1957). However he difference in "sketching" from the motif and from imagination is much less clear. If the purpose of drawing from nature or a motif is simply pleasure or to gather information for unspecified later use then it is not fulfilling the primary function of the sketch. Two questions should be asked before deciding whether or not such activity counts as "sketching".

The first concerns the method of information collection. If the purpose (of sketching from the motif) is purely information gathering then would not photography or video serve the function better? Will future landscape painters, for example, abandon paper sketching from nature and instead take video camera and portable computer image analysers to record at high speed the scenes of interest and at the same time manipulate the digitised images to extract and simplify the spatial and chromatic material of interest. Such systems, already in use, will soon be affordable by most artists or designers. When that happens what possible advantage could sketching by perception with paper and pencil (or stylus and tablet for recording by perception does not rule out electronic media) have for this purpose?

The question is based on false assumptions. Whilst photography and other automatic recording methods have since Impressionism had many uses to artists, they cannot compete with or replace drawing by visual examination. Even if an artist states that when sketching she or he is collecting information, there are at least two crucial hidden functions of the drawing act. The type of perception necessary for drawing must be learnt and whilst drawing artists and designers are, throughout their professional lives, continually and perhaps unconsciously developing their perceptual skills. Artists draw to amplify their perception of the motif. Perception is a very complex and largely unconscious process of extracting meaning from the mosaic of colours and luminances before the eyes. A drawing contains the carefully selected partial results of this processing. Such visually extracted information has meaning and value to the artist's purpose which is completely lacking in the unprocessed array of luminances recorded by the camera. A second, related cognitive function of sketching "from nature" may be equally important. In order to design or compose new visual objects an artist draws upon a huge remembered store of visual sources which can only be collected by visual experience. Throughout a life-time this visual knowledge becomes more sophisticated and more specific to an artist's individual style. This unconscious long term store of remembered shapes, structures, compositions and colour relationships will have many sources. However there are reasons for believing that the ability to store and recall information over long periods depends upon the meaning and depth of processing that the information has received (Craik, F. &
1. What is a Sketch?

Lockhart 1972, Craik, F. 1977, Baddeley 1990). Drawing (and colouring) during difficult perception is exactly the type of processing that might be expected to build up a necessary store of visual knowledge. If the hand and eye co-ordination needed for drawing from the motif has these two cognitive functions, then it follows that they cannot be replaced by "passive" machine image handling, however powerful. This is relevant to art education. A worrying possibility is that future machines, designed to make drawing easier or even unnecessary, might stunt the growth in students of visual capacities they need to acquire but do not understand.

The second question concerns the extent to which "sketching" from a motif is purely perceptual. It is certainly often the case with fine artists at least (e.g. Turner, Bonnard) that drawing (or colouring) from nature is as much concerned with the planning and manipulation of a future work as it is with information gathering. In this case the sketch function of invention and experiment is as important as it is when sketching is entirely "from imagination". The difference is that, although unconstrained by nature, the sketch image can also draw upon and be refreshed by the image upon the retina. Photographic evidence of the freedom with which Constable and Cezanne (amongst others) manipulated the visual structure of real landscape motifs for compositional purposes, is given by Gombrich (1962). In this study, I will treat only this manipulative and visualising way of working from the motif as true sketching.

1.6 Sketching in Context.

1.6.1 "Les Demoiselles d'Avignon": A Fine Art example.

Most of Picasso's best known works were painted in the studio. Much of the composition of his very innovative painting was developed by direct manipulation on the canvas. For example, Arnheim has documented in detail the development of his "Guernica" on and off the canvas (Arnheim, 1962). However Picasso, a brilliant draughtsman, also sketched copiously in preparation for major paintings. This example has been chosen to illustrate the context of sketching in fine art because of the historical importance of this particular painting and because of the influence that Picasso himself has had on other 20th century artists. It is also interesting because of the length of the time he devoted to sketching and planning the painting (many months) and the brevity of the time that he spent in actually executing it (a few days).

"Les Demoiselles d'Avignon" (Illustration XIII) was probably conceived by Picasso in 1906, during a brief stay in Catalonia, as a large figure painting in a new style. After months of work on drawings and studies in which he filled eight sketch books (Glimcher, 1986) he painted the large eight foot square canvas in
1. What is a Sketch?

only a few days during the spring of 1907 (Penrose, 1962). The painting, with its revolutionary way of portraying human figures, was influenced by Egyptian art, Catalanian medieval frescoes and (almost certainly) by African sculpture. It is important because it represents a turning point in the development of Western painting, anticipating the later invention of cubism by Picasso and Braque. The fact that Picasso spent so much longer thinking about and planning the painting, with many small very private sketches, than in its execution is significant. It indicates that, despite the sparse, skeletal nature of most of the sketches and the absence in them of colour, Picasso must have used the sketches to develop a clear mental picture of the painting before he lifted his paint brush. Thus the private, cognitive nature of sketching and the need for a theory of mental representation to explain them is confirmed. On one sketch-book, Picasso actually wrote in large letters "Je suis le cahier" ("I am the sketch-book") painting around the printed word "CAHIER" and writing decoratively the words "Je suis le" above. Although the meaning of this is ambiguous I think Picasso meant "My Mind is the Sketch-Book. These marks are merely reminders and entry points to my private mental imagery".

As usual with surviving examples of an artist's private sketches, the exact temporal sequence can only be guessed. However they can be arranged in a sort of tree with early sketches showing the layout of rough but conventional figures gradually being replaced by a clearer composition and figures closer to the final inventive geometry. Many of the sketches are thoughts about the global composition with signs of mind changing leading to back-tracking and a new branch on the tree.(see Robert Rosenblum in Glimcher, 1986). For example an early sketch shows a male figure entering from the left. Another one shows this being crossed out. Later sketches show further simplifications. Interestingly, as the sketches approach the final composition they become more skeletal and sparser. Does this imply that as the mental idea becomes clearer the sketch need be less detailed or that attention is becoming more specific? I will discuss later the grounds for believing that sparse segregation of specific visual attributes may assist selective thought. Other sketches show Picasso experimenting with the geometry of individual figures including some which were used in the final painting. This parallel series illustrates a point discussed in Chapter 2 that we cannot assume that all visual composition passes in simple stages from global structure to local detail the reverse process may occur as well. (See Illustrations XI & XII).

The sketches for "Les Demoiselles" illustrates a typical working practice of composing by successive refinement in stages. The early conceptual stage is followed by a long period of visual gestation, with hazy ideas about a new way of representing form becoming gradually clarified by mental trial and error with many
1. What is a Sketch?

sketches as the chief tool. The final spatial composition and the colour structure seems to have been worked out amazingly quickly on the canvas. It is of course wrong to assume that all artists work in this way. There are probably as many ways of inventing in Fine art as there are artists. In this example Picasso clearly used his early sketches to think about form, with colour decisions being made later. Turner's colour sketches on the other hand often indicate the reverse sequence with colour thoughts initially preceding and over-riding those of form (Illustration XXXV). This may relate to the individual way in which different artists construct a painting. Picasso commonly drew before using colour (see Clouzot's film "Le Mystere de Picasso"). Dufy often used the reverse process, that of outlining the structure in washes of colour and of then later drawing on top to define the form (Cogniat, 1970). However the method of at first thinking separately about different elements of a complex composition, which are only later manipulated into a unified whole, is common.

1.6.2 The "Yamazaki Serving Collection": A Design Example.

In Fine Art the somewhat haphazard and private nature of sketches reflects the fact that both the motivation and the constraints on fine art are generated privately and only represented implicitly in sketches. In design on the other hand the constraints are often explicit and likely to reflect the fact that many of the design constraints and part of the motivation are imposed externally by clients or the market. Although a designer's sketches have similar inventive functions to those of Fine artists, it is also likely they will be used to communicate design ideas to clients or management. The "Yamazaki serving Collection" has been chosen to illustrate sketching in this context. (Illustrations XIV, XV, XVI). It is a collection of stainless steel tableware designed in 1981 for the Japanese manufacturing company Yamazaki by the British designer Robert Welch (Illustration XVII, Crawford, 1983 Welch, 1986). Yamazaki who manufacture stainless steel were hoping to break into the American market, under their own name, for the first time and they needed a launching range of table-ware. Traditionally, stainless steel has only enjoyed prestige in the United States as a material for cutlery. Recently, however the American public has come to see the virtues of a material which, if highly polished in manufacture, can look like silver but which costs less and can go in the dishwasher. Hence the brief set out by Katherine Vockings, the director of Marketing at Yamazaki Tableware Inc. The designs had to be highly polished, to look like silver. They had to be distinguished in appearance so that they could take the place of traditional silverware on the American table. They also had to be modern in feeling, for the market was believed to lie between the ages of twenty
1. What is a Sketch?

and forty. During the seventeen months when the Serving Collection was on his
drawing board, Robert Welch produced hundreds of drawings. This example is
appropriate because sketches played an especially important part in it. Welch
conceived the idea early on of using spiral flutes on his designs to capture the
brilliant reflections which he knew were a valuable characteristic of highly polished
steel. However it proved impossible to make models of the pieces in shiny metal to
assist his clients to judge their effectiveness. The drawings had to do all the work.
Only the cream jug and sugar bowl were modelled fully. Most of the other pieces
were accepted on the basis of drawings and a wooden model. At this stage Welch
was in competition with four other designers. The trays, being impossible to
simulate were accepted as technical drawings. Here we see sketches being used at
several overlapping stages of the design process. There is a crucial first statement
of the design problem and Welch's conceptually clear but visually vague ideas
about its solution. There then follow rather vague early sketches, covered with
written notes, which perfectly illustrate the need to preserve in the representation
multiple alternatives. Even a very precise design concept can have a one to many
mapping to its visual representations. As Welch clarified his ideas, the drawings
gradually became less ambiguous and were sometimes sent to the clients,
accompanied by wooden models. Finally two metal prototypes were made by hand,
by Welch's assistant, and a series of much more finished drawings sent to the
clients for approval. These show skilful use of ink, chalk and wash to suggest the
surface reflections so hard to model. (Illustration XVI) They are not "realistic" in
the computer graphics sense of the word but elicit in the viewer's "mind's eye"
vivid imagery of the finished product. This exploits, at a late stage, the same
imagery processes in the client's mind that the attributes of early sketches exploit
in the designer's mind. That they achieved their purpose is shown by Katherine
Vocking's reply

"Dear Robert,

I have had these lovely renderings on the wall since they arrived. However I
think you should have them back for your permanent file. They lent inspiration and
urgency and now that the project seems on target, they should go home."

Illustration XVII shows the final result.

1.7 Summary and Conclusions.

The word "sketch" has a meaning which is evolving. It refers to a natural
class of visual representations with an indeterminate boundary. Sketches can be
defined to be any visual representation used for modelling or thinking about
something else. Alternatively they can be defined as visual representations which
possess, and are valued for so doing, attributes of indeterminacy such as
incompletion, untidyness and ambiguity. Such attributes demand a greater
cognitive contribution from the viewer than those of other types of representation.
Without claiming to define the sketch, this study will focus on two-dimensional,
viewer-centred images which belong also to two classes of visual representation,
namely representations that are used for visualising something else and which also
possess the attributes of indeterminacy outlined above. They are thus untidy or
indeterminate, views of imagined appearance, valued not for any intrinsic merit,
but for their usefulness in the creation of something else. This is regarded as a
crucially interesting subset of all the images, studies and models which are
sometimes called "sketches".

Two tasks with which we commonly associate sketching can be summarised
as

1 - Facilitating the invention and manipulation of imaginary objects
2 - Facilitating the imaginary manipulation of existing objects

It is the purpose of this study to examine how the special attributes of
sketches assist the mind to perform these overlapping functions. This is an essential
precondition for the design of improved sketching systems and media.
2. SKETCH ATTRIBUTES AND THEIR INFLUENCE.

"Confused things rouse the mind to new inventions"

Leonardo da Vinci "Treatise on Painting"

2.1 Three Kinds of Indeterminacy.

Typical "privileged" attributes of sketches have already been listed in Chapter 1. These can be summarised by saying that they are all types of visual indeterminacy. It has been argued that indeterminacy is a necessary component of visual experiment by successive refinement. However the indeterminacies are of many different kinds and we need a system for classifying them which might help to put constraints upon their meaning. However this is very difficult. The indeterminacies of interest are sketch representations which have a one to many relationship with the possible visual appearances of that object. This reflects an artist's need to "vague" or, more precisely, to preserve the right not to limit prematurely the number of visual options. Such indeterminacies are partial and their limits are hard to define because their meanings are unpredictable and dependent upon context. However their value to invention springs not only from their unpredictability but also from the constraints upon them which successively limit the range of their meanings. In skilled hands, the indeterminacy of a sketch is a controlled indeterminacy with implicit constraints well understood by their user. Pomerantz and Kubovy (1986) attempt to apply the theory of information to the perceived components of pictures. According to these authors the information carried by a component of a line drawing is inversely related to its predictability. However this fact is of little value to a theory of the sketch unless we know the dimensions of complexity of a sketch component and the number of represented meanings to which it may refer. Pomerantz and Kubovy (1986) discuss the difficulty of applying information theories of perception to a scribble. "For example, it is impossible to assess the complexity (or information load) of a scribble drawn on paper without knowing the dimensions along which the scribble was free to vary, or in other words, what alternative scribbles might have occurred instead of the one actually produced. Scribbles seem quite complex perceptually, and indeed they are hard to encode, describe, copy and remember. Scribbles are ill-defined patterns produced with very few constraints. Thus every scribble has a large number of alternatives and so carries a high information load. But unless the few constraints that do exist can be stated precisely, the exact information load (or uncertainty) of scribbles is infinite, since with no constraints an infinite number of scribbles can be generated.". The constraints on scribbles in sketches are provided,
2. Sketch Attributes and their Influence.

Black (1937) distinguishes three notions of partial indeterminacy which have been discussed recently by Herbert Toth (1994) in relation to fuzzy set theory. As a start (and it is only a start) it may be useful to see how Black's distinctions apply to the indeterminate attributes of sketches.

2.1.1 Generality

According to Black (1937) this "is constituted by the application of a symbol to a multiplicity of objects in the fields of reference." The meaning of a generality is clear even though it has many exemplars. Written notes on sketches have this kind of indeterminacy. It is a mistake to suppose that a verbal proposition about a design is necessarily vague even though it has a one to many (often infinite) mapping to its visual alternatives. Thus Millet's "green rushes" and "stones set on top of each other" are not vague because they specify definable sets (Illustration XXIII). They constrain visual form without specifying its fine detail, as is necessary in the early stages of composition by successive refinement. Generalising or categorising (I will argue) is an essential component of the visual thought that accompanies the sketch, even though it is not necessarily explicitly represented.

2.1.2 Ambiguity

This is "the association of a finite number of alternative meanings having the same (phonetic) form" (Black 1937, Toth 1994). Visual ambiguity is most useful in the early stages of composition when a limited number of alternatives are possible. Some of the Leonardo type one pentimenti are of this kind. Multiple contours elicit specific alternative images which may belong to different categories. Robert Welch's fuzzy representation of the grooves of his serving dish (Illustration XIV) were intended to represent two alternative types of profile. Here the ambiguity is clarified by the added note. A much more drastic ambiguity of category may also occur in both perception and visual invention (compare Illustrations II and XXXI).

2.1.3 Vagueness

Black (1937) distinguishes Vagueness from Generality and Ambiguity as follows. "It is characteristic of the vague symbol that there are no alternative symbols in the language, and its vagueness is a feature of the boundary of its extension, and is not constituted by the extension itself." Examples of vagueness in the sketch are missing or fuzzy contours, blurred or unspecified backgrounds,
overflowing or unspecified colour and low resolution of form. A vague sketch attribute by this definition is one in which the defining boundary of its possible representational meanings is itself indeterminate. It is not enough that it should be general or ambiguous.

2.1.4 Discussion

The Black definitions of Generality, Ambiguity and Vagueness were conceived in relation to symbols and it is not clear that they are as meaningful when applied to the depictive elements of sketches (see Chapter 3). It is clear that the three types of indeterminacy are not mutually exclusive and that they frequently occur in inseparable combination. Thus the sketch of a figure in illustration XXII has generality as it specifies a class of figures, it has some ambiguity about the possible positions of the limbs and there is much vagueness of visual detail. However the distinction could be important in choosing the right questions to ask about sketch functions. It is easy and probably common to confound the three meanings in working practice as "Vagueness". For example when a designer complains that a certain visualising system makes it "hard to be vague" (Scrivener and Edmonds, 1980, Makirinne-Crofts et al. 1992) we should ask does she/he refer to the fact that

1. Visual Generality is hard to record
2. Ambiguity is hard to represent
3. Fuzzy visual extension (Vagueness) is difficult to represent?

All three indeterminacies are important to sketching but have different implications for the design of machine data structures and processes.

2.2 The Origins of Sketch Indeterminacies.

The indeterminacy in sketches may be manifested by under-representation as in visually segregated, sparse or incomplete drawings. It may be manifested in over-representation as with multiple contours (pentimenti), incompletely erased alternative forms and densely untidy backgrounds. It may be manifested by tolerant representation as when colour is allowed to overflow the linear boundaries of a represented surface. As already suggested, such indeterminacy arises from the need to preserve varying degrees of uncertainty in the early stages of visual composition. However artists have not always tolerated the untidy indeterminacies which are now so familiar in their working drawings and colour studies (Gombrich, 1966a). Before the Renaissance, medieval European artists prided themselves on the skill and tidy accuracy of their working drawings and studies. In some respects the standards used to assess such studies were similar to those which are implicit in the design of contemporary machine CAD systems. The anecdote of
Giotto drawing a perfect circle at his first attempt when the king of Naples asked for a token of his skill, is well known (Vasari, 1878). Today a child can emulate Giotto using any cheap C.A.D. system. Whilst it is harder to draw by hand a perfect circle than a wobbly, imperfect one it is the opposite for the programmer of a CAD system. Perfect lines and geometric shapes are easily programmed because they are plotted from a mathematical formula using the screen co-ordinates. Imitating a hand drawn circle is much harder. Villard d'Honnecourt's "Swan" given by Gombrich as an example of medieval drawing, has a clean and perfect outline with no evidence of adjustment or error correction (Illustration I). Its placing on a page with another unrelated drawing suggests indeed that it had the function of a sketch, but unlike sketches by later artists there are no untidy breaks or indeterminacies in the definition of the contour. This is also a quality encouraged in computer "paint" systems as it is necessary for the application of colour using the "flood-fill" facility. With this, as all CAD users know, the slightest break in a contour, even a microscopic one, leads to the colour leaking out and flooding the background. Thus whilst the tidy application of colour is much easier by machine than by hand, it has as its prerequisite very tidy drawing.

The medieval concern for tidiness and perfection is related, Gombrich implies, to cultural conventions which valued the formula of the pattern-book and encouraged apprentices to learn by copying exactly the works of his masters (Gombrich, 1952, 1966b). Drawing served a different purpose then, when the artist was so much guided by tradition and pattern. This is further historical confirmation of the a priori point deduced in Chapter 1 when discussing the need for visual experiment. "Where invention is not expected and demanded of the artist, emphasis must be on his facility in mastering the 'simile', the formula, and fumbling will therefore be frowned upon" (Gombrich, 1952). It is remarkable how rarely untidy corrections occur in pre-fifteenth century working drawings. As a rule if a medieval artist had doubts about which formula to adopt for a composition he preferred to begin afresh.

According to Gombrich it was Leonardo da Vinci, perhaps the most influential draughtsman of all time, who first advocated the deliberate use of indeterminacy in the sketch. Leonardo gave two reasons for recommending what Gombrich argues is a revolutionary change in established workshop practice. The first is explained by comparing the painter to the poet. "Now have you never thought about how poets compose their verse ? They do not trouble to trace beautiful letters nor do they mind crossing out several lines so as to make them better." (Leonardo, 1270). Gombrich interprets Leonardo as meaning that painting like poetry is a product of the mind and that to "lay stress on tidiness of drawing
execution is just as philistine and unworthy as to judge a poet's draft by the beauty of his handwriting". However, if this is really what Leonardo meant, (it is not clear from the Treatise quotation) then it is an unconvincing point. A sceptic could point out that there is a crucial difference between the draft of a poem and a working drawing. The handwriting does not matter in the draft of a poem because poetry is made of symbols (words) the meaning of which is unaffected by the style of their execution. They need not even be visual, they can be read aloud, and although a good reading is preferred to a bad one, the poem remains the same. "Poetry concrete" excepted, no-one claims that the quality of a poem is affected by the medium in which the words are presented. With a working drawing, however, the execution is part of its form and can alter the meaning and quality of the finished work it represents. This does not imply that poems cannot also be "sketched" with many early crossings out and multiple alternatives forms. However it is the words, found or suggested, which matter in the unfinished poem not the handwriting. Nevertheless, there is a sense in which drawing resembles poetry, which is what Leonardo really meant. Both drawing and poetry can have an ambiguity or vagueness which is intrinsic to their purpose. This only becomes clear in relation to the second and much more interesting reason for Leonardo's advocacy of the indeterminate.

Leonardo's second and main point is that incompletion and indeterminacy in a sketch can stimulate the mind to invent new visual form whilst preserving multiple options for detail to be resolved later.

"For you must understand that if only you have hit off such an untidy composition in accordance with the subject, it will give all the more satisfaction when it is later clothed in the perfection appropriate to all its parts. I have even seen shapes in clouds and on patchy walls which have roused me to beautiful inventions of various things, and even though such shapes totally lack finish in any single part they were yet not devoid of perfection in their gestures or other movements." (Leonardo, 1270)

Leonardo is really making two points here, although they are interdependent. These twin functions of indeterminacy, the arousal of visual invention and the deliberate use of vagueness or incompletion to avoid premature decisions are confused in Leonardo's Treatise, but they deserve to be distinguished. The first is most useful in the early conceptual stages of visual composition and is characterised by the unexpected, almost accidental relationship between the indeterminate percept and the new meaning or visual idea it elicits. I shall argue later (Chapter 7) that unexpected sensory experience has the function in creative behaviour of shaking the visual brain out of over-learned habitual
processing patterns. The second use of indeterminacy occurs in the post-conceptual stages of visual creation. This is when the general content and structure of a composition is decided but not the details of form and colour. The sketch user at this stage frequently feels the need to represent aspects of his design or composition incompletely or with varying tolerances, preserving some options whilst determining others. The first use in sketching can be more general than the second and discoveries made through it may find unexpected applications in other works. There is a parallel here between a distinction implicit in Leonardo's working methods and a distinction made today between two methods of computer soft-ware design. In "top-down" design the objectives and global structure of an individual program are defined early and the modules of which it is composed are written later specifically to serve that initial structure. In "object-orientated" design it may sometimes be useful to proceed "bottom-up" writing important modules or routines early before global structure has been determined. If flexible and portable, such modules can save work, and be used repeatedly in future programs, the functions and structure of which is still unknown. By analogy artists and designers usually work "top-down" on the composition of a specific commission. At other times, though, they will work "bottom-up" exploring visual ideas later to be used as components of works the structure and content of which is still unknown. See for example Picasso's experimental sketches of figures used later in "The Demoiselles d'Avignon" (Illustrations XI,XII).

In the first role suggested by Leonardo, indeterminacies help their user to imagine completely new visual ideas, amplifying invention and originality. The indeterminacies in this case may be random and have a purely accidental resemblance to the invented image.

"I shall not refrain from including among these precepts a new and speculative idea, which although it may seem trivial and almost laughable, is none the less of great value in quickening the spirit of invention. It is this: that you look at certain walls stained with damp or at stones of uneven colour. If you have to invent some setting you will be able to see in these the likeness of divine landscapes, adorned with mountains, ruins, rocks, woods, great plains, hills, and valleys in great variety; and then again you will see there battles and strange figures in violent action, expression of faces, and clothes, and an infinity of things which you will be able to reduce to their complete and proper forms. In such walls the same thing happens as in the sound of bells, in whose strokes you may find every named word which you can imagine." (Leonardo 1270)

and "For confused things rouse the mind to new inventions" (ibid)
Leonardo is careful to emphasise that this use of the indeterminate is no easy trick for would be painters. It requires skill and it is necessary that the artist should possess a detailed visual knowledge of the objects and parts which the indeterminacies evoke. "But see to it that you first know all the parts of the things you want to represent, be it those of animals, of landscapes, of rocks, plants or others." It is clearly implied that the richness and quality of the visual invention provoked by his indeterminacies is proportionate to the visual erudition and experience of the user. "The mind of the painter" he wrote "should be like a mirror which is filled with as many images as there are things placed before him" (from his notebooks, MacCurdy 1938)

In the second role the indeterminacies are necessary to allow the preservation of options during the gradual refinement of an idea already conceived.

"Sketch subject picture quickly and do not give the limbs too much finish: indicate their position, which you can then work out at leisure"

and

"You will first attempt in a drawing to give to the eye an indication of the intention and invention which you first made in your imagination, then proceed to take away and add till you are satisfied, and then let draped or nude models be posed in the manner in which you have arranged the work; and see to it that they accord in measurement and scale to perspective so that there should be nothing in the work that is not in accord with reason and natural effects".

Here Leonardo is advocating the method of visual composition by gradual refinement from global composition to local detail. He criticises artists who prematurely give determinate outlines to picture parts for once such a painter has given "a beautiful and graceful finish" to these parts he will think it damaging to change them. "and these people do not deserve the slightest praise in their art".

As an example of Leonardo's use of indeterminacy to "rouse the mind to new inventions" Gombrich (1966a) gives a Study for the Battle of Anghiari, an amazing confused sketch of men fighting on horse-back (Illustration II upper). Gombrich believes this may have inspired the completely different study, for Neptune (Illustration II lower). The Neptune study is known to have originated from the time when Leonardo was working on the Battle of Anghiari.

There are many, well known examples of Leonardo's drawings where deliberate pentimenti (multiple contours, confused smudges etc) are used to allow varying positions of limbs or picture parts or to suggest movement. However Gombrich quotes one extreme example which although it is borderline, belongs to the second category of indeterminacy for refinement, rather than the first (my claim not Gombrich's). This is the "Study for the Virgin with St Anne" which is so
2. Sketch Attributes and their Influence.

dense with multiple pentimenti that the central region is completely black (Illustration III). Yet Leonardo knew his "parts" so well that he could visualise the multiple arrangements of limbs implicit in this confusion and select from them as a tracing on the reverse side of this drawing shows.

The distinction between these two uses of indeterminacy is only one of degree, for both use indeterminacy to suggest multiple alternatives, although at different stages in the creative process. Also implied but not explained in Leonardo's recommendations is the inverse notion that early, visual detail and finish may actually impede invention by making it harder to visualise the effects of alterations and changes.

Leonardo's fame and his influence over the working methods of Western artists was widespread. Botticelli and Raphael are thought to have altered their sketching style under his influence. Where the preparatory sketches of influential artists have survived in bulk, (e.g. by Rubens, Rembrandt, Goya, Delacroix and Turner's colour sketches) they show all the indeterminacies Leonardo advocated and more. Mostly the indeterminacies are of the second type used for development rather than conception. Leonardo's influence cannot usually be proved, and may in any case have been indirect. It is also possible that the humanist idea of "progress" in art (Gombrich 1966b) and the rising interest in visual experiment, led many artists to discover these attributes for themselves. Compare Rembrandt's study for the "Rape of Ganymede" with the finished painting (Illustrations XX & XXI). Leonardo would surely have approved.

There exists in addition a minor tradition in painting where the influence of Leonardo's first type of indeterminacy is also used. According to Kenneth Clark (Clark, K. 1939), Goya used the method of looking at random stains to furbish his imagination. It is easy to believe this in relation to the late paintings though I have found no further evidence. The superb sketch of a crowd scene (Illustration XXII) has all the hall marks of deliberate Leonardo confusion. One perceives in this with the "mind's eye" a number of figures in seething movement, but in exactly how many and in what positions and stances is different each time the drawing is examined.

The use of this method was made explicit in a system for composing landscapes invented between 1780 and 1785 by the English artist Alexander Cozens who published his method in a book (Cozens, 1785). Cozens belonged to the classical tradition of artificial but realistic composed landscapes. In his work and that of his son the tradition of Claude, Rosa and Gaspard Poussin was transformed into an English mode. Although Cozens is usually regarded as only a minor artist his work influenced both Constable and Turner. Constable actually
copied sketches from Cozens in his note-books (Marqusee, 1977). These artists influenced in turn the French Impressionists and hence the mainstream of European painting.

Cozens described the accidental discovery which lead to his "blotting" technique in these words -

"Reflecting one day in company with a pupil of great natural capacity, on original composition of landscape in contradistinction to copying, I lamented the want of a mechanical method sufficiently expeditious and extensive to draw forth the ideas of an ingenious mind disposed to the art of designing. At this instant happening to have a piece of soiled paper under my hand, and casting my eyes on it slightly, I sketched something like a landscape on it, with a pencil, in order to catch some hint which might be improved into a rule. The stains, though extremely faint, appeared upon revival to have influenced me, insensibly, in expressing the general appearance of a landscape".

Having found that this helped his student greatly, Cozens developed his "blotting" technique into a systematic set of rules for inventing landscapes. The starting "blots" were not completely random. In his book he recommends the reader to "Possess your mind strongly with a subject" and then with black ink and a paint-brush to "make all possible variety of shapes and marks upon your paper, combining the disposition of the whole to the general subject in mind". A piece of tracing paper is then placed over these marks and a landscape sketched out with a fine pencil using the blots to stimulate the imagination (Illustration IV). Cozens claimed that he discovered the method, before he was informed that "something of the same kind had been mentioned by Leonardo da Vinci" in his Treatise on Painting. However "it may easily be imagined how eagerly I consulted the book" which tended to "confirm my own opinions". It is clear that he had read the Leonardo passages carefully before writing up the method.

Another more recent artist who explicitly used and was influenced by the Leonardo type I indeterminacy was Max Ernst. In a way of working he called the "frottage" Ernst used artificially generated marks to suggest figurative and often fantastic imagery (Spies, 1968). He commonly started by making pencil rubbings of all manner of patterns and textures obtained at chance from the objects around him. He then developed these according to the imagery they suggested (Illustration V). Ernst’s own account of his discovery of this method is revealing -

"On August 10th 1925, I was seized with an unbearable visual need to discover the technical means whereby this theory of Leonardo’s is clearly worked out in practice. It began with a … childhood memory, in the course of which an imitation mahogany board in my bed played the part of optical provocateur in a
daydream. On a rainy evening, I found myself in a hotel on the French coast when I was gripped by an obsession that made me stare excitedly at the deeply grooved cracks in the floorboards. I decided to yield to the symbolism of this obsession. To sustain my potential for meditation and hallucination, I made a series of sketches on the floorboards by arbitrarily placing a few sheets of paper on them and then began to rub them with a black pencil. When I closely scrutinized the sketches thus made - 'the darker areas and other, delicately lit half-dark areas' - I was amazed at the sudden intensification of my visionary capabilities and the hallucinatory result of the contrasting pictures." (Spies, 1968)

The discovery that certain types of visual indeterminacy can have such a powerful visionary effect is not due to solely to the influence of Leonardo's genius. A similar technique was discovered and used earlier and independently in the East.

The Eleventh century Chinese artist Sung Ti is reported to have criticised the paintings of a colleague for their "want of natural effect" and recommended the following course: "You should choose an old tumbledown wall and throw over it a piece of white silk. Then morning and evening you should gaze at it until, at length, you can see the ruins through the silk, its prominences, its levels, its zigzags, and its cleavages, storing then up in your mind and fixing them in your eyes. Make the prominences your mountains, the lower part your water, the hollows your ravines, the cracks your streams, the lighter parts your nearest points, the darker parts your most distant points. Get all these thoroughly into you, and soon you will see men, birds, plants and trees, flying and moving among them. You may then ply your brush according to your fancy and the result will be of heaven not of men." (Quoted by Marqusee, 1977).

In an excellent analysis of the nature of drawing technique, Philip Rawson (1969) describes the "blob" (or splash) as intrinsically representing the element of hazard in drawing. Some styles he writes, exile hazard and the blob completely. Others cultivate it deliberately. The "po-mo" technique of the Chinese ink-painters of the Sung dynasty and later Chinese and Japanese periods developed the use of "spilled ink" as a deliberate stimulus to invention. The Sesshu, "spilled ink" landscape (Rawson, 1969) goes further than the Cozens blot since indeterminacies remain as the "beholder's share" in the finished drawing, whilst Cozens make its clear that his blot indeterminacies are strictly for the artist's use only.

2.3 Indeterminacies in Fine Art; The Beholder's Share.

The Sesshu illustration alluded to above introduces a third role for indeterminacy which is not limited to the sketch. The functions of indeterminacy in finished art are not identical to those in sketches. However they are relevant to the discussion of sketch practice because attributes of painting which we now take for
2. Sketch Attributes and their Influence.

granted have their origins in the sketch and have in turn influenced sketch style. Without understanding this function, the sceptical computer system designer might suggest that although like most of us, artists may see faces and figures in the stains on walls or the dying embers of a fire these images are usually too vague and limited to find applications in design. The deliberate use of such a stimulus is the activity of an unrepresentative minority of artists who have unusual powers of vivid hallucinatory, imagery. She or he might further argue that an advantage of modern computer aids to visualising over traditional media, is that it offers the potential to replace the user's own vague and often weak visualising powers with vivid and detailed source imagery, drawn perhaps with some unexpected randomness from a large pictorial data-base. Will not such systems when fully developed with "photographic realism" become a powerful alternative aid to visual invention making the ancient, splash, blot and scribble completely out of date? Perhaps, but it would be very foolish to attempt to design such invention eliciting systems without first considering the role that indeterminacy has traditionally played in figurative art and without investigating the meaningfulness of the "photographic realism" concept. For the history of figurative art, until photography at least, is also the history of the attempt to create "realistic" illusory representations of visual experience. This shows, (Gombrich 1962), that acceptable realism in man-made images relies heavily both upon cultural convention and the beholder's tacit knowledge of the visual world and the objects represented. Gradually, from the post Leonardo Renaissance onwards, forms of visual suggestion by indeterminacy became admired and absorbed into the mainstream of European painting. Vasari, commenting on the later "broken colour" style of Titian, noted how crudely daubed and separated brush-strokes were executed in such a way that "one sees nothing at close quarters, though they look perfect from a distance." (Vasari, 1878). This became an admired ability of many later painters of quality. Velasquez for example used especially long brushes to keep his distance from the canvas. His biographer, Palomino, remarked that his portraits "are unintelligible from close quarters but miraculous seen from afar" (quoted in Gombrich 1962). In the hands of such artists such controlled use of indeterminacy was an intuitive way of achieving a kind of realism which is impossible by greater attention to fine detail. It works, as was realised by the time of Reynolds (Gombrich, 1962), because the viewer unconsciously superimposes memory images of the subject over the brush-marks provided he is far enough away not to see their boundaries clearly. Thus portraits painted in this manner would seem to have an illusory memory based likeness to relatives of the sitter and probably looked slightly different to each viewer.
However it was not until Impressionism that the dividing line between the privileged indeterminacies of working sketches and finished paintings was truly blurred. The vigorous landscapes painted before the motif by Constable and now much admired for their own sake, were because of their rough untidiness, intended only as coloured sketches to be finished in full detail in the studio. They were not intended for sale and were valued at only a few pounds after his death. (Compare the sketch for the "Hay Wain" now in the Victoria and Albert Museum with the finished painting). It appears to have been Manet who initiated the Impressionist doctrine that "a picture must be frankly stamped with the spontaneous character of a sketch" (Wilenski, 1944). The Impressionist insistence on painting directly from the motif and the high value they placed on an effect of accident in the composition also helps to blur the distinction between the sketch and the finished painting. It is due mainly to the Impressionists that, as noted in Chapter 1, the meaning of the word "sketch" is no longer clearly definable in terms of either its function or its attributes. According to Gombrich (1962) the beholder's share was increased in Impressionism because the viewer must from a distance perceive the suggested form and light "across" and despite the brush-marks whereas the brush-marks of the late Titian or Velasquez assist the viewers imagery by their direction and shape. Since Impressionism, this type of tacit dependence on the viewer has been taken for granted in much figurative and semi-figurative painting. A painting by Frank Auerbach (Circa 1960) in the Tate Gallery viewed from a few feet is simply a meaningless abstraction of deeply furrowed white, grey and pinkish brown paint piled 1 - 2 inches thick. Seen from a distance it becomes a reclining nude with strange solidity and volume. A supreme example of such controlled exploitation of the beholder's imagery is the Monet "Nymphéas" paintings in the Orangerie, Paris. If viewed close to one sees in these thin washes of graduated colour superimposed with rough and tangled or widely separated brush-marks of more saturated colour. From further away one experiences strong illusions (hallucinations?) of water rippling and of light being reflected on water whilst at the same time of looking through water into another world.

The Impressionists sought total objectivity and the innocent eye in their attempts to capture fleeting patterns of light upon the retina. This claim, remarks Gombrich, is at variance with the subjective visual knowledge on which their paintings rely. But surely there is no paradox here. In searching for a true description of visual experience, the Impressionists discovered that it was only by combining a selective objectivity with the indeterminacy of nature, that the artist can evoke the viewer's unconscious memory images. It is the mysterious combination of the beholder's subjective visual memory and a skillfully constructed
objective percept, which creates the "impression" of changing, reality in "Impressionism".

520 years after the publication of Leonardo's "Treatise" we still do not have a detailed or convincing theory to explain the evocative and "hallucinatory" effects of visual indeterminacy which Leonardo recommended as an aid to invention and which has been used to such effect in art. In his "Art and Illusion" Gombrich (1962) refers to the theory of "projection" which was applied by psychologists when they used, for example, the Rorschach ink-blot test as an aid to psychoanalysis (Rorschach, 1942). In this diagnostic tool, patients are asked to inspect a standard set of bilaterally symmetrical abstract blots and to say what they suggest or resemble. It is hoped that their responses will reveal to the analyst unconsciously repressed sources of their anxiety or neurosis. However the projection "theory" with its implication that a visual memory is somehow projected onto the perceived blot explains little that Leonardo could not have surmised himself. It is obvious of course that the "projection" cannot be taken literally, since conscious percepts are the subjective experience of neural activity in the visual cortex of the brain. Crucial questions remain about how this supposed combined percept and projected memory is represented in the mind and what processes generate and use it. For example

Why and how does a confused indeterminacy elicit such "projected" imagery more strongly than a blank sheet of paper for example or closed eyes?

How is projected imagery used in the generation and refining of visual ideas?

Does its use require special and rare talent or can it be developed by anyone?

Should we be educating our students in its use, and if so how?

Lastly, should not the users labile and often vague imagery be advantageously replaced by the "virtual reality" of future machine systems?

It is clear that these and other questions central to the design of improved visualising aids are questions about human cognitive capacity and the possible processes of visual thought available to artists and designers. They will be addressed later.

2.4 Photography, Realism and the Sketch.

The invention of photography by Niepce and Daguerre in France and by Talbot in Britain, during the 1820s was officially announced at the French Academy of Sciences in 1839. Its influence on the history of art has been well documented by Aaron Scharf (1968). At the time that its possibilities first became familiar to artists one of the most fashionable styles of painting resembled
photography in some respects. In France between 1850 and 1859 a school of "Realism" advocated extreme pictorial objectivity of the kind to which the camera seems ideally suited. Some artists saw the camera, with its capacity to capture quickly fine nuances of tonality and minute details of surface texture, as a dangerous competitor. Others, who saw their livelihoods threatened including many of the miniature portrait painters turned to photography themselves. The more far-sighted however saw photography as a liberating tool which could be used discreetly (and often secretly) to broaden the information gathering function of the sketch. Courbet is known to have used photography extensively though he very rarely copied directly from them. His "Burial at Ornans" scandalized visitors to the 1850-51 Salon who saw in it the objective "ugliness" of the daguerreotype. The Fashionable "realism" was the false realism of Classical Idealism where behind the illusion of life-like realism lies the artist’s careful filtering out of all unwanted aspects of natural motifs.

So does the comparison with very objective painting confirm the idea that the photograph represents the pinnacle of representational realism? This certainly seems to be the implication behind the contemporary use of the term "photographic realism" as a desirable goal for computer graphics. The answer to this question depends upon whether the term realism means true to the image as a pattern of light focused upon the retina or true to the image as it is consciously experienced in perception. However as was understood before the turn of the century, by science (Hering, 1964 (original 1874), Chevreul, 1839, Helmholtz, 1910 (original 1870)) and by perceptive artists (Scharf, 1968) that for both form and colour these two are quite different. Thus we are not normally conscious of the extreme effects of perspective on the retina (a halving of size with a doubling of distance). Before we are aware of "seeing", the brain compensates for the resulting distortion of size, shape and parallelism, by inferring the causative changes of distance and orientation and constructing for conscious experience a scene in which objects preserve their constant shape and colour (e.g. Gregory 1970, Zusne 1970, Beck 1972, Rock 1983 see also this text 4.2.1). A photograph records the transitory and unprocessed mosaic of luminances as they are focused on a human retina. However unlike the image on the retina, most photographs are not viewed with the eye from the single unique camera position which is the only position from which the perspective is optically "correct" (i.e resembles that on the retina). This can result in the most "unreal" distortion. The problem of the dependence of perspective on exact eye position was understood in the Renaissance When Brunelleschi first demonstrated the new perspective, he arranged for the picture to be viewed reflected in a mirror, as seen from behind the surface it was painted on, through a
2. Sketch Attributes and their Influence.

hole drilled in the vanishing point, thus ensuring that the eye position was correct and could not move (Vasari, 1878). Computer "Virtual reality" systems update the Brunelleschi demonstration with simulated movement and stereoscopy. In 1892 after half a century of debate on the causes of photographic aberrations of perspective, Streintz (Scharf, 1968) published three photographs taken of the same street scene and landscape from the same position using three different lenses of 6, 25 and 73 degrees angle of vision. The three perspective views (as Alberti and Brunelleschi could have predicted four centuries earlier) are completely different. The closer the effective foreground of the motif to the lens the greater will be the exaggeration of perspective scale in the photograph. This type of exaggeration, impossible with the eye, can be manipulated for visual effect in the cinema. Both Italian Renaissance painting (White, 1987) and modern Virtual Reality (Gray, 1992) present an artificial perspective by calculation. The photograph records perspective directly from the physical world. This is a distinction too easily forgotten in machine assisted imagery.

Even "objective" painting from the motif preserves, in subtle ways, the "conclusions" of tacit interpretation and recognition processes within the brain of the user. Delacroix, who used photographs extensively as source material, understood this intuitively. Scharf demonstrates that when one compares the posed photograph of a nude which Delacroix used for his "Odalisque" of 1857 the painter deliberately elongated the fore-shortened thighs and other aspects of the photographic pose which would have been unbelievable in the painting (Scharf, 1974. illustrations on page 124) Delacroix wrote in his essay of 1850 "The daguerreotype is only a reflection of the real, only a copy, in some ways false just because it is so exact... The eye corrects" (Scharf, 1974 page 120)

There is a sense in which the photograph is more real than the painting, simply because its viewer knows or presumes that it has been generated by direct uninterpreted contact with the light reflected from its motif (see Chapter 3). The subtleties of tonality and surface texture in a good photograph are often valued, in a good photograph because they are, like nature, so unpredictable and unimaginable. This type of objective realism, "optical" realism, was close to some of the stated objectives of the Impressionists. Because the light pattern of the photograph is so detailed a record of the light reflect from real objects, it appears at first sight to have attributes completely incompatible with the sparse or confused and subjective indeterminacies of the sketch. It is therefore rather surprising to find that indeterminacy and incompleteness in early photographs were important influences on Impressionism and hence indirectly on modern painting. These attributes were doubtless regarded as short-comings of early photography.
Nevertheless, the history of their influence shows that photography can have cognitive functions in the visual arts which has nothing to do with objective realism, but which are complementary to those of the sketch.

The first and most obvious indeterminacy in early photographs is that like drawings and pen and wash sketches they are monochrome. I will argue later that separation of tonality from chromaticity make its easier to attend to, to perceive and to appreciate tonal relationships. In fact, until the invention of the panchromatic film at the beginning of the twentieth century, the relative tonality of different colours was badly distorted. Colours of the longer wavelengths (red and yellow) were rendered too dark in relation to blue and violet. Contrasts of tonality in the finished print, often had to be artificially manipulated during development (another source of unreality). Despite these known indeterminacies, the subtle tonality which is so much easier to perceive in a monochrome photograph had a direct influence on the method of describing form in painting by directly applied tonal relationships. Manet, a pioneer of Impressionism used this method impressively and Scharf (1974, pages 62-75) has documented in detail how this technique was influenced by photographs. Without photography Manet's famous "Execution of the Emperor Maximilian" would have been impossible. Yet Manet rarely copied directly from photographs. Rather he used them as source material just as pre-photography history painters used their sketches from models. Thus the use of photography for tonal information is using the camera not for its realism, but for its image processing qualities as an aid to visual awareness and tonal decision taking, a point that I shall develop later. Wilenski (1945) has made a vitriolic attack on 19th century "photographic" tonal painters including, amongst others John Singer Sargent, whom he accused of mindless duplication of many of the tonal faults of photography. He justified this in part by his emphasis of the need in figurative painting for perceptual intelligence in the description of form as opposed to the mechanical transcription of tone values as if by an eye without a brain. Wilenski singled out for example the merging of two visual objects into one which sometimes occurs by the "blind" translation of photographic tone into paint. His criticisms seem absurd if applied to Manet's discriminating use of photography as a source of otherwise unavailable visual information. What would he have thought however, of the "blind" simulation of surface reflection and shadow used by contemporary machine solid modelling systems?

Blurring was another type of indeterminacy caused by technical deficiency in the early photograph which gave it an affinity with the sketch in stimulating invention in painting. There were two sources of blurring in the early daguerreotypes and calotypes. One, known as "halation" occurred on glass plates
when light penetrating the emulsion strikes the uncoated side of the glass and is reflected back through the emulsion from behind. The other is caused by movement which was especially obvious in the early days before the development of fast "instantaneous" film. Corot's famous blurred trees seem to have originated in this way (Illustration XXXIII). He is known to have been friends with a photographer, Adalbert Cuvelier. Scharf (1974, page 93) illustrates a blurred landscape signed "A.C. 1852" which is attributed to Cuvelier and has a striking similarity to Corot's painting of that time. At his death over 300 photographic plates were found in Corot's studio. Though these are now lost, Scharf provides convincing evidence of the role that photography played in the development of his characteristic style which anticipates Impressionism. Photographic blurring of objects caused by movement generated other forms of visual invention. Early photographs of street scenes are strangely empty, the figures do not register. Later they appear as strange blurs or smudges (Scharf, 1974). The many pedestrian figures and their movement which Impressionist street scenes conjured into the mind's eye are represented similarly with skilful and sparse blurred marks. A good example is Pissarro's "Place du Theatre Francais in the Rain" painted in 1898 and now in the Minneapolis Institute of Arts (Rewald, 1991 page 121). Such figurative suggestion is so familiar in painting today that it is easy to forget that it was a new discovery which Pissarro's contemporaries found hard to accept.

The influence of photography on 20th century painting has been widespread and profound. It is natural that a painter such as Richard Hamilton, who is so interested in the popular imagery of our culture and whose work seems to comment on the nature of representation should remark that "I sketch with photographs" (BBC 2 "Painting with Light" 1986). He now works extensively with Quantel's electronic image handling system, "Paint Box" collaging together popular icons in unexpected ways. His work and the work of many other painters in the history of figurative art demonstrates the value of photography to complement, in often unexpected ways, the indeterminacies of traditional sketches as aids to visual thought. There is no doubt that the power and convenience of photography and its derivatives, film and television, have caused it to have a pervading influence on the way we perceive and think visually. Despite this, history suggests that photographic attributes may not provide good criteria for assessing the "realism" of a representation. Faith in the veracity of a photograph depends upon the fact that it is contingent upon the light reflected by the objects it represents. If this is its claim to "realism", then the term "photographic realism" applied to machine processed images which lack such contingency is a contradiction. Whether it is justified to regard real photographs as sketches depends upon their unique attributes and how
they are used. It is the rather unexpected use of their indeterminacies rather than their "realism" which gives them a functional affinity to sketches as defined here.

2.5 Sketches and Written Notes.

Sketches by both fine artists and designer frequently contain written notes. At the early stages of visual composition there are certain to be many ideas about the projected work which are either too vague or too general to be given visual form. Sometimes the notes are scribbled on a note-book page opposite the sketch, sometimes they are written on the sketch itself as labels or attached to a sketch part with an arrow.

An obvious use of the note is to provide information which cannot be represented in the sketch medium. Painters sketch drawings are often accompanied with notes about colour. Many of Turner's notebooks in the British Museum consist of tiny spidery outline drawings accompanied with such colour notes (Wilkinson, 1975). The drawing by Millet "A Heath" (Illustration XXIII), obviously a study for a projected painting, contains the following notes written on the parts to which they refer:-

"cleared patches, tillage, rye on the heights, roughly cleared land, most of the stones set on top of each other, broom, green rushes" and below :- "cleared and ploughed land having a thoroughly Celtic look".

But such notes are not simply signs of the limitations of the medium. They are often firm propositions which limit without fixing the scope of multiple options. Colour notes such as "pale green" refer to many hundreds of possible colours. To represent the colour at that stage with paint might force an unwanted premature decision. It can also be assumed that phrases such as "Celtic look" conjured multiple images in Millet's mind which he could project upon the sketch.

Not surprisingly written notes are commoner in sketches by designers than by fine artists. Designers usually have many verbally represented constraints and they must often communicate their own pre-visual concepts to clients and colleagues. Thus sketches by Fashion designers typically contain notes about fabrics and about detail for pattern cutters. Illustration IX shows Karl Lagerfeld thinking to himself both verbally and visually at the same time. Early sketches by Robert Welch for stainless steel table ware, (Illustration XIV) contain notes about specific alternative structures (such as the profile of surface grooves) and also annotated diagrams about manufacture.

It must be admitted that artists sketches often contain no writing at all. However it seems likely that in the early stages of composition there are always, non-visually implemented thoughts about the projected work which are associated in memory with the sketch and which have a similar function to written notes.
Thus Vincent Van Gogh describing an unfinished painting to Theo wrote "Another canvas of size 30. This time its just simply my bedroom, only here colour is to do everything, and giving by its simplification a grander style to things, is to be suggestive here of rest or of sleep in general." (Roskill, 1963). There is no need to record such private descriptions on the sketch as they are easily remembered concepts, instantly aroused in the artist's mind by the sketches they generate.

Discussion of the function of written notes is continued in Chapter 3.

2.6 On Being the Right Size.

One of the most familiar forms of indeterminacy in the sketch is uncertainty of visual structure due to the fact that the sketch is usually much smaller than the object that it represents. Even when the size of the represented object is specified (as in a Mondrian sketch with labelled measurements, see Stoichita, 1979 page 17), generality and vagueness are introduced by the fact that the difference of scale between representation and represented introduces the "problem of the model". The intimate dependence of functional structure upon size is well illustrated in biology where the relation between size and structure is a fundamental principle of evolution (Haldane 1928).. If an organism changes its size but not its form, its volume increases with the cube of its height whilst its surface area only increases with the square. As larger species evolved, drastic increases in complexity were necessary to increase the ratio of surface area to volume in all organs to preserve their function. An analogous but visual relationship exists in art and design. In three dimensions, if a designed object doubles in size the ratio between its volume and its surface also doubles, with visible consequences. It looks heavier and less tactile. In two dimensions the length of lines and associated edges of shapes increase in proportion to the width of a drawing, whilst the area of the shapes and their associated colour increases with the square of this width. Increasing the area of colour or tone in relation to the edges where two or more colours make contact must be expected to have dramatic effects upon colour appearance. The importance of a reproduced drawing being the correct size in order to perceive and understand it is emphasised by Rawson (1969).

Understanding this fact intuitively, artists sometimes insist on sketching to the scale of the represented work. Thus mural painters commonly sketch to scale on the wall surface itself, in a labile medium which is painted over later. Matisse prepared full scale sketches for his decorations of the chapel at Vence (Diehl, 1958). He drew outlines with charcoal on the end of a pole to preserve a correct view point. However in the early stages, artists and designers are mostly content to make sketches smaller than the works they represent. For visualising large objects the requirement of fast and easy manipulation is incompatible with size equivalence
for both existing traditional and computer media and is presumably more important. It can be noted however that small sketches, produced early in the compositional process, usually differ in important ways from later, more detailed visualising aids. The very indeterminacies and omissions of detail already noted may have size indeterminacy functions. The artist may intuitively recognise that, for example, edge to area relationships can be adjusted by successively adding more colour edges/detail as sketch size approaches reality. Some sketches have an implicit built in scale based on a well known and familiar scale of the object represented. However this is not the case for unfamiliar or non-figurative design. It may even be the case that the scale of the finished work is itself indeterminate. The indeterminate represented size of a sketch frame may facilitate the use of "the mind's eye" to experimentally imagine scale changes, or to predict the consequences of scale correction. Fine artists can often make later real size related surface to volume or edge to area adjustments by direct manipulation on canvas or clay. Where this is not possible, as in the design of manufactured objects, full scale models may be used. The manipulation of "toile" garments by Fashion designers and the use of full scale wax models by car body designers are well known examples. The fast manipulation of detailed, full scale, late stage representations is a possible objective for "Virtual Reality" systems. Detailed understanding of the visual consequences of size changes as an idea progresses from sketch to model and to full scale prototype will depend upon the medium and type of design.

The whole question of size in relation to sketch function is given further significance in Chapter 6.9 in relation to the way a sketch may bias the way we perceive and attend to object structure.

2.7 Discussion and Conclusions.

This chapter is an attempt to review what is already known or can be deduced from what is already known about the privileged attributes of sketches. I have introduced Black's distinction between three kinds of indeterminacy - generality, ambiguity and vagueness in the hope that it will prove applicable to sketches and sketching systems.

The historical origins of sketch indeterminacies in Western Art can be traced partly at least to Leonardo's "Treatise on Painting" and his famous comments on the value of visual "confusion". However the fact that a similar parallel tradition exists in Eastern art of also using perceptual indeterminacy to stimulate invention, is evidence that the use of such deliberate untidyness is not only an invented tradition of style but also a repeatable discovery made independently by several artists. This is consistent with Cozens' claim that he made the "discovery" of his blotting method before he had read Leonardo.
I have claimed that Leonardo's sketch indeterminacies are of two kinds, those used for "rousing the mind to new inventions" (type a) and those used for refining an existing idea (type b). Applying Black's distinctions, the Leonardo type a indeterminacies are clearly ambiguous since they can evoke a radically new idea. The type b indeterminacies (in the position of a limb for example (Illustration XXVIII)) are perhaps better classified as indeterminacies due to generality. A degree of vagueness must be assumed to exist in both types of indeterminacy. How much vagueness a sketch component possesses however, would be expected to vary greatly with its context and its meaning to the sketch user. It remains for future researchers to determine to what extent these three aspects of indeterminacy in the sketch are separable in practice.

In order to better understand the nature of such indeterminacies it is useful to explore the way in which they have, since impressionism become accepted and admired in Fine Art. Attributes originally used in preparatory studies to enhance the artist's own invention were progressively transferred to finished works to enhance the " beholder's share" by similar mechanisms. Impressionist and post-impressionist figurative painting is a useful source of information for sketch theory. The role that deficiencies in early photography may have played in this process proves that photography was not used only for enhancing figurative realism.

In documenting briefly the frequent occurrence of written notes on sketches, I am preparing the ground for the translation function of the sketch which will be discussed in Chapter 3. Clearly written notes are frequently used to record an indeterminacy of the generality type. Such notes are not necessarily vague but they may simply stand for many visual options which the sketch user wishes to preserve.

One further type of indeterminacy of sketch representation - that of size - is rarely discussed. Here I have drawn a parallel with the influence of size upon biological structure which I have used in teaching and which I am convinced from personal experience is relevant to the visual arts. However the physical effects of size (such as the ratio of shape area to shape edge) are complicated by the way a sketch is perceived and attended to. This problem is addressed again in Chapter 6.

In conclusion, the history of working practice in the arts shows that many, but not all, types of indeterminate sketch attribute have had cognitive functions valued, but not well understood, by their artist users. The theory of mental "projection" quoted by Gombrich (1962) is incomplete and now dated. It is now time to explore the possibility of using contemporary theories of cognition and
All the five hypothetical functions of the sketch are ways of amplifying visual thought and invention. However they are interdependent in other ways that are not yet clear.
neuroscience to explain in more detail how sketch attributes might assist visual thought.

Diagram I outlines the functions of the sketch as a hierarchy, in which the primary function, "Amplifying Visual Thought and Invention" will be treated as comprising five component functions presented as hypotheses in the next five chapters. How these five hypothetical functions may be inter-connected is discussed in Chapter 8. The secondary functions of the sketch, "Recording Ideas" and "Communicating Ideas" although mentioned in Chapter 1 are outside the scope of this thesis.
3. SKETCHES AND REPRESENTATIONAL TYPE.
"Good representation then is a release from intellectual bondage." Jerome Bruner (1979)

3.1 Introduction

Hypothesis One states that "Sketches have attributes which make them intermediate between depictive (picture-like) and descriptive (language-like) representations. Such attributes are used to facilitate the bidirectional translation between descriptive and depictive modes of thought."

In order to think about complex or imaginary objects and events, we need an appropriate way of representing to ourselves the objects we are interested in and their constituent parts. Bruner, (1979) believes that successful learning and problem solving depend upon the quality and appropriateness of the representational systems used. It has also been shown that theories about the way a problem of reasoning is represented mentally, can account for the type of errors subjects make when attempting to solve it (Johnson-Laird, 1983). Representation is equally important to visual invention. There is a analogy between sketching and learning mathematics.

"I have been struck repeatedly by the economical significance of a good mode of representing things to oneself... The crutch provided by a matrix that gets all the combinations out of the head on to paper or the blackboard makes it possible to look at the group structure as a whole, to go beyond it to the task of seeing whether it has interesting properties and familiar isomorphs." (Bruner 1990, page 26)

3.2 Description versus Depiction: Defining the attributes.

The distinction between images, which are recognisable because they are "imitations", and symbols such as words which are based on imposed conventions, has been discussed in semiotics ever since Plato (Eco 1976, Gombrich 1982a). But these most familiar types of representation are only examples of a more general distinction which has been discussed by psychologists and philosophers as the distinction between "analog" and "propositional" (or "Fregian") forms of representation.

The term "analog" is often used to mean representation by some analogy or similarity between the represented and the representing worlds. Unfortunately the term is a bad one. It can be confused with the common use in engineering measurement to mean continuously variable as opposed to discreetly variable as in "digital" (as is assumed by Baddeley, 1990, page 98). In the context of
representational type "analog" does not mean this. Therefore I prefer the almost synonymous word "depictive" Kosslyn (1980, 1983).

As Paivio (1986) has pointed out, propositional theories of cognitive representation do not always make it clear whether the postulated propositions are the representational primitives used by the mind itself or whether they are simply being used as descriptions of such postulated representations. If the former then it seems clear that there are symbolic ways of representing visual objects in which relations and hence truth values are not explicit (for example sign systems in diagrams). The term description is used here in a broader sense than "proposition" which is simply one type of descriptive statement.

The most familiar examples of descriptive systems are verbal and of depictive systems are pictorial. However any attempt to show that a particular example of a representational system is purely depictive usually fails (see for example Goodman 1968, Wollheim, 1977). Similarly descriptive systems, such as writing, can be shown to have had a partly depictive ancestry, traces of which remain. (For example the letter "A" originally depicted a bull's head). As I argue here, the distinction can only be understood in terms of continuously varying clusters of descriptive or depictive attributes which representational systems may possess. Despite this the descriptive/depictive distinction is usually discussed as if it were a clear dichotomy. For example Kosslyn (1980, 1983) in connection with his work on mental imagery compares a descriptive (Propositional) representation of a ball on a box (ON(BALL, BOX)) to an outline drawing of a ball on a box. Kosslyn then lists seven differences to characterise the description/depiction distinction.

1. The proposition uses discrete symbols (for the relation and its objects) which cannot be subdivided meaningfully. On the other hand, in the drawing the "symbols" can be arbitrarily broken (half a ball, one corner of the box etc).

2. There are different types of symbol in the proposition (e.g a relation symbol is different in kind from an object symbol). On the other hand, (Kosslyn claims) there are no distinct types of symbol in the drawing. For example there is no explicit representation of the relation "on" which is simply implicit from the position of the ball and box in the drawing.

3. In a descriptive proposition the symbols are arranged according to a set of rules (a grammar) which is not the case in the drawing.

4. The description is abstract. That is it can refer to categories of objects and relations, and it is amodal (not tied to one mode of perception only). In contrast the drawing is concrete. Different attributes of an object such as size,
shape and position cannot be represented separately. It is also modal, i.e. associated only with one mode of perception (vision).

5. The descriptive proposition is unambiguous (a crucial difference from natural language) whilst the drawing is ambiguous and open to differing interpretation.

Finally Kosslyn offers two "crucial" properties which he terms privileged because they are exclusive to depictions.

6. Depictions can only exist in a medium which "functions as a space". This space may be physical such as a piece of paper or it may be functional as is a computer bit-map corresponding to a raster display.

7. The other privileged property amounts to a claim of spatial correspondence between the represented and the representing worlds. "Every part of a depiction corresponds to a part of the object represented".

Although this account draws attention to interesting attributes of representational type, it is not convincing as a definition of the underlying nature of the descriptive-depictive distinction. For example a study of drawings from different cultures and of the differences between adult and child drawing suggests that drawings do indeed use symbols (Arnheim, 1974, Deregowski, 1973, 1977, Willats, 1977, 1984, 1985). It is probably true that symbols are more discrete in the proposition than in the drawing and that the rules governing their use are more formal and restrictive but the distinction is relative rather than absolute. This does not mean that Kosslyn's suggested differences are not helpful in clarifying the nature of visual representation but it does suggest that the examples chosen do not belong to disjoint sets.

Other authors have attempted to isolate a single distinguishing characteristic from which the other differences are derived. Shepard (1981, 1982, 1984) has emphasised the idea of "isomorphism" (similarity of structure) as the distinguishing characteristic of analog (depictive) representation. He distinguishes Primary isomorphism such as the spatial and chromatic similarity between a model and the object it represents from Secondary isomorphism. He claims that whilst it is obviously impossible for the brain to use Primary Isomorphism (a green object cannot be represented by something green in the brain) Secondary Isomorphisms which preserve functional characteristics of time and space in the physical world, are quite feasible. Palmer (1975b, 1978) finds this distinction unclear and unconvincing. He himself has emphasised the distinction between Intrinsic and Extrinsic representation. Propositional (descriptive) representation represents by making external references which must be known. "whatever structure there is in a propositional representation exists solely by virtue of the
extrinsic constraints placed upon it by the truth-preserving informational correspondence with the represented world". Most of Kosslyn's first five attributes can be regarded as corollaries of this. In contrast, the representing information of analog (depictive) representation is intrinsic because it is contained within the structure of the representation itself. Moreover "Whatever structure is present in an analog representation exists by virtue of the inherent constraints within the representing world itself, without reference to the represented world".

Perhaps the clearest attempt to summarise the distinction is Sloman's description of correspondence or non-correspondence between the represented and the representing (Sloman 1975). Sloman uses the term "Fregean" to refer to representation by propositions or descriptive symbols. "Roughly, in a complex Fregean symbol the structure of the symbol corresponds not to the structure of the thing denoted but to the structure of the procedure by which the thing is identified or computed". In the analogical (depictive) case however "There must be at least some correspondence between the structure of the thing represented and the structure of the representation." (Sloman, 1975). This overlaps with Palmer's extrinsic/intrinsic distinction but is not exactly equivalent to it. Thus the structure of a sentence depends upon the history of the language and the way it is intended to be read, but not upon what it represents. But it also depends upon the reader having access to an external vocabulary of word meanings. The structure of a picture represents by correspondences which exists between it and the objects represented. Information contained within the picture may be extracted from such correspondences without external reference by the viewer's visual system.

A final distinction which is also a corollary of the different ways in which information is extracted from depictive and descriptive representations is the difference between implicit and explicit information (Finke, 1989). The information in a depictive representation is only implied by the isomorphism or correspondences and must be extracted by visual processing. The information in descriptive representations (such as propositions or mathematical equations) is made explicit by their external reference. As far as visual cognition is concerned the distinction is between the raw optical data projected as a pattern of luminances on the retina and the explicit description derived by the brain. This is extracted by complex analysis and is enriched with stored knowledge of the categories of objects and scenes depicted (the extrinsic component).

To the probably endless debate about a distinction I believe to be central to the nature of sketches, I wish to make the following contributions.

1. Most of the authors quoted above treat the depictive/descriptive distinction as if was a simple dichotomy. However the examination of cultural
representations shows that purely depictive or descriptive types are rare (if they exist at all) and that most representational systems fall along a rough continuum from very depictive to very descriptive systems with many intermediate types. Language has a small but significant depictive component, photographs a significant degree of external reference (a descriptive component).

2. Depiction and description are best characterised as distinguishing variable attributes of a particular representational system rather than the representation as a whole.

3. Although most commonly used systems fit a linear continuum in which depictiveness decreases as descriptiveness increases, they are not necessarily mutually antagonistic in this way. Using the terminology suggested by J.J.Gibson (1937) in relation to dimensions of perception, descriptiveness and depictiveness are best treated as two semi-independent "intensive" dimensions rather than a single "oppositional" dimension. Intensive dimensions can vary in quantity independently of each other. Oppositional dimensions on the other hand counter-act each other so that as one increases the other must decrease. According to Hering's theory of colour, redness and greenness are oppositional but whiteness and blackness are intensive (Hering, 1964). It is possible for example to label a depictive illustration in such a way that its descriptiveness increases without decreasing its depictiveness. I suggest that it is due more to cultural convenience than to theoretical necessity, that most familiar systems of representations fall along a continuum in which as descriptiveness increases so depictiveness decreases.

It may be useful to term systems with a high combined total of descriptive depictiveness as Dense Representations as distinguished from Sparse Representations with a low combined total of descriptive and depictive content. Egyptian tomb painting and 3-D medical models would be classed as "dense"; algebraic equations and unlabelled drawings as "sparse" using these terms. I believe that the distinction has a special relevance to theories of cognitive representation. These must be much richer or "denser" than the relatively sparse cultural representations, such as writing and drawing which are used to support them.

3.3 Indices, Icons and Symbols: The Sagacity of C.S. Peirce.

A hundred years ago C.S. Peirce was interested in the nature of meaning which he showed depends not only upon a sign and a referent but also upon an "interpreant" by which he meant a representation of the world in terms of which the "sign-referent" relationship is mediated (Peirce 1960, Bruner 1990). Peirce distinguished three representational types, the "Index", the "Icon" and the "Symbol" by the way their constituents were related to what they represented. The
relation between the Symbol and its referent depends upon the existence of a rule-governed system or "language" that determines its position within a system of signs. "Description" defined here is equivalent to systems of Peirce's Symbols. The Icon bears a "resemblance" relationship to its referent and so is roughly equivalent to our "Depictive" representation. It has to be mentioned however that more recent theorists have shown the concept of resemblance or "similarity" in cognition is difficult to pin down (Palmer 1978, Treisman 1986, Tversky, 1977), hence Sloman's substitution of "correspondence".

The Index which Peirce treats as a separate type, represents by being "contingent" upon its referent as "smoke is to fire". The signs within the representing world of an Index system are related to the represented world by cause and effect. This is really another way of saying that the Index is non-semantically (non-cognitively) generated. The importance of distinguishing Indices from Icons, or rather index like attributes from Icon like attributes in the design of sketching systems lies in the impartiality of the information provided. Indices provide evidence (and this includes useful surprises), about the physical reality of the objects they represent. The example of photography has already been discussed in relation to its historical influence on sketch attributes (Chapter 2). It is now possible with computer systems to generate Icons which can successfully masquerade as Indices. This raises questions about the value of "realism" in visual images and how we distinguish the genuine index from its imitative relative.

At first glance Peirce's tripartite division, if accepted, seems to discredit the whole idea of a simple continuum of types. We seem to be dealing with a triangular conceptual space (or a prism if we include my suggested dense \ sparse dimension). However I believe a simple continuum from very depictive to very descriptive remains a useful if approximate ordering system for those representational types which can be used for visualising purposes. My reasons are as follows :-

1. It is difficult even to imagine a representational system which is purely contingent (i.e an Index). As a near example we might consider how an X-ray diffraction pattern could be used to represent the structure of a molecule. This certainly does not resemble the molecule and its crystal structure on which it is contingent. However by Sloman's criterion it has a correspondence (indirect and without resemblance), with its referent and hence has a type of depictiveness. Indices and Icons are not disjoint sets.

2. Visual recall is indirectly contingent upon the object perceived, encoded and stored. A degree of contingency is arguably present in most and perhaps all types of depictive representation including the most imaginatively generated ones.
Drawing from the motif is an act partly contingent upon the light pattern falling on the retina and partly upon non-contingent cognitive processes. Sketching from imagination may sometimes seem to be without any contingency from the imagined objects represented. However imagination depends heavily upon visual memory. As Thomas Hobbes (1651) argued, we can create images of novel objects because we have seen the component parts at one time or another, "As when from the sight of a man at one time and a horse at another, we conceive in our mind a Centaure" (Hobbes, 1651). However memories are faulty and continually altered upon recall by our expectations and stored schema (Bartlett 1932). Thus there is an unknown and indirect contingent (index) component even in representations constructed from imagination.

There is also a semantically generated (non-contingent) content in all useful contingent representations. The meaning and information content of the most impartial photograph is dependant upon its context and the way its view point is selected by the photographer. New technology has created a whole range of contingent (index) representational types for medical and scientific purposes. (For example - X-Ray photography, electron and ultra-violet microscopy, infra-red and radio astronomy, nuclear magnetic resonance imaging of body tissues etc).

3.4 Visual Representation: A Continuum of Type.

It is worth considering briefly what alternatives are available within the continuum of type, in order to decide where sketches belong in this untidy continuum of intermediate cultural types. A list of examples from the very depictive to the very descriptive are described briefly below (paragraphs 3.4.1 to 3.4.12). The continuum is complicated the added contingent/imaginary and sparse/dense dimensions. I have based position in the continuum upon depictiveness, with descriptiveness usually but not always increasing as depictiveness decreases. The sum of depictive and descriptive attributes rather than their ratio determines the sparsity/density attribute.

3.4.1 Replicas and Examples.

The most depictive type of representation is an exact replica. A replica can qualify as a representational system if it used as a prototype or example. A replica is an unique object which represents a class of future objects, typically preceding manufacture by mass production. Exemplification (or partial replication) might include swatch-books of materials, samples or component-parts. According to Goodman (1968), exemplification is a mode of symbolization, but as he later points out this depends upon the extent to which "reference" is intended.
3.4.2 Working Models and Simulations.

The next level contains a number of parallel representational types. A detailed working scale model of a car is a theoretical type rather than a practical possibility. As soon as size correspondence is sacrificed, functional isomorphism becomes impossible. This is the "problem of scale" discussed in Chapter 2. If the model included labels, notes, handbooks etc then it is a mixed type which descends the continuum. "Virtual Reality" systems aspire to model appearance at this level.

3.4.3 Static Three-D Models

Scale models and wax-works have no temporal depictive correspondences. They cannot simulate events. They are designed to duplicate as closely as possible the unprocessed light pattern presented by the surfaces of real objects. In the late stages of design there is sometimes a conflict between the need to manipulate and the desire for detailed depiction. Labelling and written specifications can make this representational system denser. Car designers still build life-size clay models of new designs. Here it is worth noting that isomorphism of appearance only ("resemblance") is classified high in the depictive sequence as it is not combined with other non-isomorphic references.

3.4.4 Holograms, Television and Photography.

High in contingency, this class of cultural representation exhibits correspondence of optical (light) structure. Cylindrical holograms have the highest degree of light correspondence because they preserve in the form of wave interference patterns information about the light reflected from objects in many directions. The holographic plate is a good example of the distinction between correspondence and "resemblance". Although it contains much more detailed spatial information than a photograph its abstract pattern of lines has no resemblance to the object it represents until the light pattern it encodes is re-created by special viewing conditions (usually with a laser). Television and photography (moving or static) encode only a tiny sample of the luminous information, that which happens to be reflected in the direction of the camera lens. This sacrifices the third dimension and destroys spatial correspondence. The correspondence of the photograph to the scene represented is the correspondence of projective geometry, where for example, real relative size and apparent relative size, due to differences of distance, are confused under the umbrella of visual angle.

3.4.5 Figurative Art and Animated Graphics.

Lower in the continuum are all representations in which cognitive processes are explicit in both generative and interpretive processes. Within this class there is
however a wide range of degrees of depictive correspondence between what is represented and what is used to represent. Modern portrait sculpture, representing form without colour, clearly uses suggestion as does most portrait painting. Most theorists now agree that cultural convention plays an important role in the way that painting represents. (Gombrich, 1962).

3.4.6 Line Drawings and Illustrations.

As has been repeatedly pointed out, the lines used in drawing have no existence in the retinal images projected by real objects. Linear biological illustrations or engineering drawings of automobile structure preserve different types of spatial isomorphism with the represented objects but include non-isomorphic sign systems. Artists' drawings on the other hand are typically much freer in the variety of meanings which may be assigned to lines.

It has been shown that very young children who have never been shown line drawings before, are able without difficulty to recognise the objects and scenes they represent (Hochberg and Brooks, 1962). One explanation of this otherwise strange finding is that line drawings access an internal cognitive representation created by the brain during the early edge extraction stages of visual processing (Frisby 1979, Marr and Hildreth, 1980). The "self-occluding" contour lines used in drawings make a good example of a genuinely intermediate depictive/descriptive type. Their shape has a direct correspondence ("natural isomorphism" Palmer 1978) with the shape of the object represented but line thickness has a meaning which is descriptive and does necessarily correspond to any part of the object represented. Another example of a genuinely intermediate component might be the line representing the combined head and body in children's "tadpole" drawings (Willats, 1985).

As the result of an interesting study of the drawing methods of children, untrained adults and art students, Pratt (1984) has concluded that the viewer-centred "accuracy" (i.e representational isomorphism) of drawing from nature is related to the level of perceptual analysis that is available to the drawing maker. The lower the level of perceptual analysis the more depictive and "view-specific" the drawing. Further, Pratt claims that the description first available to a would be draughtsman is at the highest descriptive level and that access to lower levels (needed for view specific representation) involves learning to descend through the intermediate levels. This provides a further clue to the difference between depiction and description.

Besides the optional use of labels and codes referred to above, illustrations in books have explicit reference to the accompanying text, which in turn can refer
back to the illustrations. Since the text is itself a descriptive system dependent upon external reference to meaning, it may be necessary here to distinguish direct and indirect types of external reference. Gombrich (1982a) draws attention to the intermediate status of posters and comics. He suggests that the relative amount of learning necessary to use a representational system may be a criterion for assessing the advantages of intermediate depictive / descriptive types. The success of comics is due to the advantages they offer to the learner. The comic strip provides a series of related images which support each other. Such drawings are depictive but supported not only by the words in "balloons" but also by easily learned sign systems such as "streaks" indicating movement, stars in front of the eyes and question marks indicating pain or puzzlement.

3.4.7 Sketches

Tentatively I place artists' sketches somewhere inbetween linear illustrations and maps or diagrams. Their indeterminacies and incompletions ensure that they are less depictive than drawings or linear illustrations, but that they are more depictive than maps and most diagrams, because they are viewer centred and correspond more closely in this sense with appearances. When they contain added verbal notes and labels, they are (like maps) mixed depictive and descriptive systems. Sketches which are drawings also possess intermediate depictive / descriptive components, such as lines used to represent figure-ground contours. Colour can be either descriptive (ie coded) or depictive depending on how it is used. It can also be, (as in post-impressionist painting), depictive in one dimension (such as value) and symbolic in another (such as hue). Many of the characteristic incompletions and indeterminacies of sketches are intermediate representational components which facilitate descriptive to depictive translation. Thus pentimenti can depict a number of related edges corresponding to different examples of a possible object silhouette viewpoint. Sketches viewed in isolation are sparse representations and small alterations in the ratio of depictive to descriptive attributes will make them move up or down in the continuum. Thus their exact position is variable.

3.4.8 Maps and Diagrams

In the conventional map, two spatial dimensions parallel to the ground are represented depictively. Distances and directions are implicit and must be extracted by visual scanning or measurement. Heights represented by contour lines are equivalent to a number of maps superimposed. Topographically placed symbols provide descriptive encoding. A map of interest as an analogy, is the "War chart" with moveable flags and markers placed upon it o represent the changing states of
a battle. Here the chart has the function of a depictive spatial medium to assist the visualising of spatial relationships of mobile descriptive information. Each flag makes descriptive reference to fire power, division strength, velocity etc. In principle such a system of representation can be hierarchic to any depth. The descriptive flags may point to secondary depictive maps which in turn can possess topographically specified descriptive pointers. The hierarchical labelled map is a plausible candidate for the type of mental representation used by visual working memory which is specifically supported by the sketch (see Chapter 4).

A spatial isomorphism can be either discrete or continuous. An example of a discrete isomorphism is a pixellated raster display. Once the correspondence between spatial location in the represented and the representing worlds is limited to discrete fields of view then the mapping may be between spatial fields with different size ratios of represented to representing fields. Thus a street map may include a "blow-up" of a densely crowded town centre. Street widths are usually grossly exaggerated to allow printed names etc.

Neither is it necessary to preserve isomorphism of inter-locational direction and distance, unless the purpose of the map is to allow these attributes to be extracted by spatial measurements. Electronic-circuit diagrams and the famous London Underground map assist the interpretation of connectivity by sacrificing isomorphism of direction and distance. Perhaps it is a general principle of cultural visual representations that they facilitate cognitive processes (in low capacity short-term memory) by eliminating or simplifying the irrelevant (see Chapter 6).

Diagrams usually preserve a deliberately limited and specified isomorphism. Characteristically the preserved isomorphisms are those of relationships, such as "connected to", "part of" etc. Those components of vision which are least relevant to the content of a diagram are used symbolically. For example colour can be used symbolically in a diagram depicting spatial structure, whilst spatial arrangement can be symbolic in a diagram depicting colour (as used in charts of colour classification systems). There is evidence that the symbolic use of space to represent non-spatial quantitative dimensions is used unconsciously in solving mental problems (Moyer & Bayer 1976).

3.4.9 Pictorial Notation Systems, Graphic Signs.

There is, perhaps, a gap in the proposed visual continuum somewhere between diagrams and natural language. Labelled diagrams, for example, are mixed representational systems where words and depictive spatial or connectivity systems co-exist in an uncomfortable compromise. Evidence for genuine continuity between pictorial and visual verbal systems can be found in the pictographic
origins of writing and non-viewer centred modes of illustration as used, for example, in ancient Egypt of the ninth to eighteenth dynasties (Gregory, 1970). Modern types at this level abound in science and technology. An example is the molecular structure diagram with topologically depictive bond representation combined with symbolic atoms.

3.4.10 Natural Language

Language is the classic descriptive system with extrinsic rules of formation and learned rules of meaning. Hockett (1960) has proposed a number of such defining criteria. Human infants and animals species without language can recognise and categorise objects in ways which suggest that much stored visual information must be represented descriptively in the brain (Sutherland, 1968, see also Chapter 4). Although there are good reasons for believing that descriptive information encoded for recognition must be non-verbal, it may be that language is necessary to make such visual description accessible to consciousness. To explain the facts of language comprehension and production despite its many ambiguities and context dependencies Chomsky (1957, 1965, 1980) has proposed a **transformational** grammar in which comprehension involves transforming the **surface structure** of language to a cognitively represented **deep structure** in which the ambiguities have been removed and from which the meaning is extracted.

The relevance of theories of language representation to the sketch is shown by the written notes which often accompany artist's sketches and by the sketch function of translating verbally (propositionally) expressed ideas into depictive form. It is interesting that natural language, in distinction to more formal notation systems, shares many of the characterising attributes of sketches.

For example:

1. The meaning of both verbal phrases and sketch line segments is often context dependant.

2. Both natural language (e.g poetry) and sketch components may be enriched by ambiguity.

3. Both sketch and natural language comprehension may require mental completion from memory.

4. Both representational systems are strongly dependant for their meaning upon poorly understood knowledge systems (semantic memory).

Because of these shared characteristics, language is an appropriate descriptive partner to the depictive components of sketches. It will be argued in Chapter 4 that some sketch attributes depend for their meaning upon a one to many
mapping from their depictive ambiguities to an analogous descriptive "deep" representation of object structure.

3.4.11 Pure Notation Systems.

According to Goodman (1968), neither pictures nor natural language qualify as pure notation systems. He has proposed five stringent requirements which are most nearly met by musical notation.

1. Firstly, all examples of characters in a notation system must belong to disjoint sets.
2. The second syntactic requirement according to Goodman is that the characters must be "finitely differentiated".
   The remaining three requirements proposed by Goodman are semantic.
3. No character in a notation system can be ambiguous.
4. Different characters must have "semantic disjointness". That is the disjoint sets representing characters must refer to equally disjoint sets of elements in the represented world ("compliants").
5. The final requirement for a notation system is "semantic finite differentiation". "For any two characters $K$ and $K'$ such that their compliance classes are not identical, and every object $h$ that does not comply with both, determination that either $h$ does not comply with $K$ or that $h$ does not comply with $K'$ must be theoretically possible".

Goodman argues that pictures are descriptive symbol systems (although of an impure kind) so he would not agree with their position here towards the depictive end of a depictive to descriptive continuum. His definition of a pure notation system is a conceptual ideal which is rarely if ever met in practice. Certainly sketches and most diagrams would fail on both the syntactic and semantic requirements. Natural language fails on the ambiguity requirement but even if it did not it would fail on semantic disjointness.

3.4.12 Mathematical Symbol Systems.

Mathematics is linked to language by propositional calculus (Russell, 1903) thus reinforcing the idea of a cultural continuum of representation with poorly defined boundaries between types. It is also the most abstract and the most capable of extension and generalisation. It is probably true that all descriptive theories of representation aspire to the condition of mathematics because mathematical description is so firmly based on explicit premises and explicit demonstrable deductive argument. Russell (1903) defined pure mathematics as the "Class of all propositions of the form $p$ implies $q$ where $p$ and $q$ are propositions....". For instance, the history of geometry suggests that there may be analogies to be found
between the processes used by the biological visual system for extracting abstract and categorical descriptive information from retinal images and the historical evolution of mathematical systems for proving and describing spatial relationships (Bronowski 1973, Kline 1954).

3.5 Intermediate types as Translation Systems.

Because the most familiar representations of our culture tend to be polarised, as pictures or language, at the two ends of the continuum (Arnheim 1970), it is natural to conclude that depiction and description are independent modes of representation. I have attempted to characterise the difference and to show that there is a continuum of intermediate types in which the distinction is less obvious. I now wish to propose that depiction and depiction are inter-dependant attributes, both of which are necessary components of any useful representational system. All the representational types in the above continuum depend for their interpretation and usefulness upon cognition. Words, the most prototypical of descriptive systems, depend for their meaning as much upon the mental images they evoke as upon other words (Paivio 1971b, 1986). If mental images are depictive (as argued in Chapter 5) then writing when used to assist visual thought has a greater depictive component by association than it appears to have on paper. There is also evidence that logical inference may depend upon the use of depictive prototypes and mental models (Johnson-Laird 1983). Description used to represent categories often depends upon the ability to mentally generate typical exemplars of the depictive possibilities which are implied. Depictive possibilities must be scanned and analysed to extract new and perhaps unexpected, explicit descriptive information. Such descriptions can in turn be used to modify or to generate anew further depictive information.

In considering in principle how descriptive information about objects, scenes and events can be mapped back into depictive form, it is important to distinguish different levels of description. Jackendoff (1983) has argued persuasively that many types of word meaning depend upon mentally stored representations of spatial and three-D structure used for visual recognition. He proposes that the translation of representational type which allows us to "talk about pictures" or to visualise the meanings of words occurs by shared visual and linguistic access to a high level "conceptual" information structure (Jackendoff, 1983, 1992).

In machines used for image manipulation, an analogous hierarchy of descriptive levels exists. At the lowest level, (I will call it the "array level") that is used by the computer "bit-map" or raster display, we can divide an image up into small elements, and assign each one propositional symbols representing colour and
position, or (as in maps) address pointers to separately stored object descriptions. At other levels, visual structure may be described incompletely using, for example, a coordinate system normalised for size, position and orientation or describing some structural relationships but not others (as in G.K.S graphics systems, Hopgood et al. 1983). It is possible to depict such descriptions (i.e. translate) using conventions to sample the multiple alternatives they define. When such descriptions define or modify explicitly visual or spatial structure, I will term this the **figural** level (following Goldschmidt 1992). Shape grammars (e.g. Leeuwenberg 1982) fit a low version of this level. At higher levels, description may be used to categorise the objects to be depicted or make statements about them in ways which imply visual structure, but only indirectly. Here a mapping (from scratch) to depiction is impossible without adding information by indeterminate invention or from outside sources. However some constraints at this level may be used interactively to re-depict or re-describe an existing representation. I will term this the **conceptual** level. Applied to sketching the three levels constitute an untidy classification with many intermediate types. But then untidyness is a principal characteristic of sketching!

Sketches resemble maps of visual appearance in some respects but are unlike them in others. They resemble maps in their intermediate status containing spatially depictive representations combined with descriptive labels, codes and sign systems (Fish and Scrivener, 1990). They differ from them in their untidy indeterminacies and their fluidity. This is a consequence of their function of assisting visual invention. Because maps are shared representations used for communicating shared information they must be clear, tidy and conform to agreed conventions. In contrast a sketch is usually unique and personal. Every sketch differs from every other sketch because its function as a "cognitive catalyst" is unique and specific to a projected invention. In art and design the "reactants" (starting constraints) may be known but the end products are not and the nature of the cognitive "reaction" to be catalysed is obscure. The untidiness of the sketch, as already suggested, is necessary to its function. But perhaps the most significant difference is the extent to which sketches in contrast to maps rely upon and interact with invisible cognitive representations within the brain of the user. Their sparsity and disordered untidiness only make sense if we assume that they are only the representational "tip of an iceberg" (see chapters 4 and 5). Sketch-books are simply collections of fragmentary traces (cognitive "footprints"), providing evidence of a much wider and richer cognitive dialogue conducted within the mind of their user. The mental images and propositions which form the conscious component of this internal dialogue are known only in memory by the artist. The nature of the
representations available to the brain for visual invention can only be inferred from
the experimental data of cognitive psychology and neuroscience.

3.6 Cognitive Representation: Is there also a Continuum of Type?

To what extent does the cultural continuum of type mirror an analogous
continuum of type within the brain? What can cognitive science infer about the
cognitive representations which form such an important component of the sketch?

3.6.1 The Neural Medium.

At the lowest level of analysis all forms of representation within the brain
are neural. Until recently, discussions of representational type in cognitive
psychology would omit completely any attempt to map theories of representation
onto known neurophysiology (eg Palmer 1977, Rumelhart & Norman 1985,
Pylyshyn 1984). This is due to the influence of functionalism. It is often assumed
that issues of representation are most fruitfully discussed independently of the "low
level" questions of medium or physical implementation. However there are now
signs of a thaw in this attitude. Useful cross-fertilisation between neuroscience and
cognitive psychology is occurring (LeDoux & Hirst 1986, Kosslyn and Koenig,
1992, Kosslyn 1994, Crick 1994). Psychologists discussing mental representation
are now beginning to quote findings in neuroscience as evidence (Ullman 1989,
Tarr & Pinker 1989, Kosslyn 1994). This may be the only approach which can
resolve some long standing controversies. Anderson, J. (1978) for example has
argued that it may only be possible to distinguish depictive from descriptive
theories of mental imagery at a physiological level. Even if this is not correct
(Hayes-Roth 1979), it is becoming increasingly necessary to ask how a theory of
representation could in principle be consistent with known properties of the visual
pathway within the brain.

This is particularly true in considering what it can mean to say that some
forms of mental representation are depictive (analog). It is much easier to see how
the brain can represent descriptively. The elements of any representation within the
brain must be constituted from the electrical signals by which neurons
communicate with each other, by their state or readiness to send such impulses and
by the way they are connected. It is through their synapses that neurons
communicate with each other. These are finely branching microscopic junctions
with other neurons which communicate by passing chemical transmitter substances
across minute gaps between terminal "boutons" from the sending cell and the
membrane of the receiving cell. The pattern of connectivity of the neurons is
determined not only by the detailed neuroanatomy of the brain but also by the
varying and adaptive physiology of the synapses. There are about a 100,000 million neurons in the brain each of which may make up to 20,000 synapses with other neurons. When a neuron receives synaptic signals from other neurons, a graded potential develops which must reach a threshold value before it in turn fires a chain of impulses to other neurons. Some synapses are excitatory increasing the graded potential, others are inhibitory, reducing it. Thus the output from an individual neuron is the result of a summing effect of positive and negative signals from other neurons. Recordings from single cells in the neocortex of the primate brain have shown that the processing of visual information from the eyes is subdivided into specialised parallel pathways within which much sequential processing takes place. Within the cortex there are now believed to be twenty four or more "computational maps" of the visual field (Kosslyn and Koenig, 1992). The visual cortex is divided into at least six distinct layers of neurons. At right angles to these, organised columns of cells are found, which each respond to a different part of the visual field, its receptive field. In order to function as the medium for a depictive representation however, such maps would have to be associated with the appropriate information extraction and scanning processes. To date the physiology of such processes remains obscure.

In area V1 of the primate brain, the first cortical area to receive signals from the eyes and the best studied, the cortical columns map the visual field of both eyes with uniform columns of cells which respond to colour and to lines or edges at different orientations (Hubel and Wiesel 1977 1979, Hubel et Al. 1978). Although the columns are uniform in size and structure, the size of their receptive fields varies, being smallest where they map the central foveal region of the retina and larger for the periphery thus reflecting the variation of resolution of the visual field when the eyes are fixated. It is tempting to suggest that these columns of cells "represent" edge orientation. However Hubel and Wiesel are careful not to make this claim remarking only that they demonstrate early stages in the analysis of the retinal image. This early stage cortical map has also been shown to encode in parallel up to nine different spatial frequencies (Maffei & Fiorentini 1977). (The significance of this for theories of the sketch is discussed in chapters 5 & 6). Maps receiving information later in the visual pathway are specialised for processing different visual attributes. Thus V5 is specialised for movement analysis and V4 for colour and perhaps form. The processing of spatial position, movement and attentional orientation appear to be separated from that of colour and object form or identity (Hubel & Livingstone 1987, Livingstone 1988). The reason why there are so many spatially organised maps of the visual field in the cortex has been discussed by Cowey (1981). He suggests that the maps reflect an efficient
processing division of labour for each point within the visual field. The maps are organised in such a way as to minimise the number of long connections (axons) between neurons since these are inefficient and biologically "expensive" in resources. If there were too much processing to be done within any one receptive field the neural columns within a single map would become overcrowded and require longer connections to communicate with their neighbouring columns. Thus the adjacent columns within a cortical map contain those neurons which need to communicate with each other most frequently.

An influential idea to explain learning and memory is the cell assembly, a closed loop of cells which continue to signal after the initial stimulus has ceased and in which repeated stimulation causes permanent changes in the excitatory capacity of synapses (Hebb 1949, Jusczyk & Klein 1980). In Parallel Distributed Processing (PDP) theories, representation is supposed to be distributed throughout a net in such a way that the output is a matrix sum of the inputs. PDP models of recognition and learning have excited interest because they demonstrate properties which are hard to model in other ways, such as graceful degradation and the capacity for completion when elements are missing. However it is still not clear whether the fundamental unit of visual representation is the neuron itself (or the neural column) or the synapses that it makes. The evidence from neuroscience is very complex and deserves a more detailed treatment than it can have here. However the point is that what is now known is about the brain's physiology is consistent with the possibility of a continuum of representational types from the partially depictive labelled map (spatial location represented depictively with object attributes descriptively) to abstract descriptive representations of concepts and categories by either neuron clusters or localised PDP networks. The significance to the sketch of the modularity in the visual pathway will be discussed in Chapter 6.

3.6.2 Different Approaches.

Most of the debate about mental representation has been at the functional level. This is the level at which it is thought to be meaningful, for example, to implement computer models of cognitive representations. At this level, as Paivio (1983) has pointed out, there are at least three different sub levels of meaning to "mental representation" which are often confused. At one level we may be simply referring to the conscious percepts and memories of cultural representations. In this sense there must be, by definition, a mental continuum corresponding to our perception of the cultural one. This is almost the level of meaning of Paivio's own "Dual Code" theory. This proposes that language and images are independent parallel "symbolic" encoding systems used for memory and thought (Paivio,
1971b, 1983). Although they can operate independently the two systems are also interconnected and can influence each other in many ways that Paivio has researched. He has shown for example that concrete nouns are more strongly associated with mental images than abstract nouns in memory experiments.

At another level, theories of mental representation can be theories about how best to represent the hypothetical mental representations for modelling purposes. It is not always clear in this context what propositional theories of cognition mean. Are they simply used as notational devices for describing mental representations or do they refer to the form of the representations themselves. To me, only the former meaning makes sense, when theorists refer for example to Structural Descriptions as nodes connected by "non-verbal" propositions (Palmer, 1977)

It is the intermediate level of representation, how images and language are themselves represented and whether there is a continuum analogous to the cultural continuum which concerns us here. Paivio himself has adopted Morton's "Logogen" theory of language representation as an underlying representational theory of language and proposes that an analogous symbolic unit the "imagen" may underlie mental images. This does not help resolve the question of how depictively images are represented. Nor does Paivio discuss the possibility that intermediate representational types may be used for translation purposes. The two systems are simply "interconnected".

3.6.3 Cognitive Depiction.

Some theorists believe that all representation within the brain is best described as propositional (ie descriptively) and that there cannot be depictive or analog representations. The debate which is now twenty five years old has centred around the possible depictive nature of mental images. The evidence and arguments can only be briefly summarised here. Clearly at one level, - the response pattern of the neural medium - the propositionalists (Pylyshyn 1973, 1975, 1981, 1984, Anderson, J. 1983, Simon 1972, Hinton 1979) must be right. Cultural depictive representation may reflect light patterns which resemble the light patterns reflected by real objects. However within the brain colour and luminance can only be represented by patterns of neural activity which is surely not depictive in Shepard's first order sense (Shepard, 1982). Spatial correspondences are preserved in the cortical maps mentioned above. Unfortunately these cortical maps are not in themselves evidence for the existence of the functional depictive representation of space. It does not matter that there is no physical isomorphism between, for example, the size and shape of the retinal image and the size and shape of the
corresponding cortical maps. However to show that such maps function as
depictive representation of space it would be necessary to show that there are
processes (perhaps some neural equivalent of visual scanning) which use the
connectivity relationships between the cortical columns, in say area V1, to extract
and make explicit new descriptive information, which is only implicit in the
response pattern encoded in a cortical map. Until such processes are discovered,
the physiological status of cortical maps as depictive representations is purely
speculative. However the known physiology certainly does not preclude the
existence of neurally encoded representations of visual structure, which function as
spatial media, from which descriptive information can be extracted. The theoretical
grounds for believing that both perception and mental imagery depend upon
representation within such spatial arrays is very persuasive (see Chapter 5).

Evidence that percepts and images share representational resources is
provided in Chapter 5.

3.6.4 Cognitive Description.

There is little controversy that the brain stores and manipulates information
represented symbolically. Functionalists who believe this is the only form of
cognitive representation have already been referred to. Johnson-Laird (1983) cites
Craik’s three cognitive processes of human reasoning in symbols.

"1. A 'translation' of some external process into an internal representation
in terms of words, numbers or other symbols.

2. The derivation of other symbols from them by some sort of inferential
process

3. A 'retranslation' of these symbols into actions, or at least a recognition
of the correspondence between these symbols and external events, as in realizing
that a prediction is fulfilled". (Craik, K. 1943).

However it is not clear that, as Pylyshyn did later, Craik, K. used the word
"symbol" to rule out spatial arrays which represent by correspondence between the
represented and the representing. Here I use "descriptive" to mean representation
without such correspondence. (Some authors use "symbol" in a sense which could
include spatial arrays).

We use language (and other cultural symbols) privately as well as socially
to represent and process visual concepts and categories which must be represented
descriptively. The huge literature on this subject is beyond the scope of this study.
I have mentioned, only briefly, Chomsky’s theories of how language structure is
represented, Morton’s "logogen" theory of language generation and access (Morton
1969 1979) and Paivio’s "dual code" theory (Paivio 1971a 1986). (See also Schank
& Nash-Webber (1975)). It is interesting to speculate that language is the means by which humans make unconscious descriptive information available for voluntary manipulation. Information, that is, which is represented and processed automatically in non-verbal animals. Although language is used for descriptive representation, it does not follow that its representation in working memory is descriptive. We use spatially depictive images of words and symbols when we think and the "inner voice" - auditorily depictive images of verbal sounds - when we reason (Baddeley 1986).

### 3.6.5 Intermediate or Mixed Types and Cognitive Type translation.

Most theories of perceptual processing are (implicitly) theories of depictive to descriptive translation. How do the early pre-conscious stages of vision capture and then represent, from momentary retinal information, the perceptual invariants of size, shape etc? In his influential computational study of early visual processing Marr asks the question when does vision "go symbolic" (Marr 1982). His widely quoted representational structures were

1. The "Primal Sketch" which represented the early two-dimensional results of extracting edges, blobs and bars.

2. The "2½ D Sketch" which represents the orientation and depth of surfaces from one view-point, a Viewer-centred representation. This is the representation of a scene which is implicitly available to conscious attention.

3. A Three-D object centred representation.

Marr's theoretical representational systems form an interesting series, postulated as stages in the visual processing of objects. The primal sketch and the 2½D sketch can both be considered as being at about the level of labelled maps in the continuum discussed above. The crucial transition occurring, according to Marr, before a recognition match, is from a viewer-centred coordinate system to an object centred coordinate system. In so far as this is a change away from correspondence with the image on the retina but towards an abstract description of the three dimensional object, it is a depictive to descriptive shift. However the transition in perception is not necessarily simply one of depictive to descriptive translation, although it is often described in these terms. As will be discussed Chapter 4, there are grounds for believing that conscious percepts are late constructions in which descriptive information in semantic memory is used to enhance rather than replace their depictiveness. This reverse translation process is explicit in the Kosslyn model of mental image generation. The possible role of a matching reverse translation of type in perception has been little considered. Marr's object centred theory of recognition (Marr & Nishihara 1978) was based on
a multi-level hierarchy of generalised cones. Although this tackled the theoretical problems, evidence since been provided that the brain uses viewer-centred representations as well as object centred for recognition. (Palmer et Al. 1981, Jolicoeur & Kosslyn 1983, Humphreys & Quinlan 1987, Pinker 1984a). Ullman (1989) has proposed an interesting intermediate type for object shape recognition, the "Pictorial Description". This overcomes some of the objections that are usually made to template matching (ie spatially depictive) theories of recognition. Ullman envisages spatial structure of objects represented in a depictive map but normalised for position, size and orientation. He then considers the computational problems of aligning such a match. This leads to problems only if purely depictive representations are considered. However Ullman considers the possibility that abstract descriptions (obtained by part decomposition) could be added at spatial points. The resulting scheme is attractive and is consistent with the finding that spatial location and object detail are processed in different regions of the visual cortex.(Van Essen & Maunsell 1983, Mishkin et Al. 1983, Ungerleider 1985). Tarr and Pinker (1989) suggest two possible representational alternatives for object recognition to the Marr and Nishihara object centred proposal. Objects could be represented in a single canonical orientation, so that mental rotation operations are necessary to transform an input shape before input and memory are compared. Alternatively shapes could be stored in a set of representations each corresponding to a different orientation. They present experimental results which suggest a hybrid hypothesis in which there are stored multiple views, but that transformation of the input shape is necessary to find a match either with the nearest stored orientation or with one at a canonical orientation. Palmer et Al. (1981) have presented evidence that such "canonical" perspective views of objects are views that encode the most information and produce the fastest recognition response times.

We can now speculate on two ways in which the special attributes of sketches might facilitate the early and involuntary stages of visual perception recognition.

1. Sketch contour lines and other marks may provide fast, pre-processed access to early representational stages of perception, which extract edges and figure / ground relationships etc. (as suggested above see Frisby 1979).

2. Sketch components may exploit stored prototypical views of objects and contain descriptive labels which facilitate alignment and matching processes with stored "pictorial descriptions" in long term memory.

Thus unlike photographs and other realistic images sketches despite their sparsity preserve the results of earlier cognitive processing which can be exploited
by the visual system. The unproven assumption here is that this facilitation releases some cognitive resource for visual invention.

Probably more important to sketch function are the consciously controlled translations of type which occur when sketching. One example of depictive to descriptive translation, curve tracing, has been investigated by Jolicoeur et Al. (1986). These workers investigated the time taken by subjects to determine whether or not a small marker, such as a cross, lay on or off one of two irregularly folded, bending lines. They found that the response times were linearly related to the lengths of the lines even when exposure times were too brief for eye-movements to be possible. The results are consistent with an internal "covert" scanning movement of attention on a neurally represented spatial medium, necessary to convert implicit depictive information to a descriptively explicit form. Ullman (1984) has described a number of other such operations which might plausibly be used to extract descriptive spatial relations from early depictive visual representations. Examples are "inside / outside" relations, "abstract shape qualities" and "open-endedness". He further suggests that these operations might be assembled into "visual routines" which operate on a base representation to generate a sequence of temporary "Incremental representations". Here we see the idea of representation as a dynamic structure. His later idea of Pictorial Descriptions, might be structured this way, with a 2½D spatial map acting as a base representation for increasingly developed descriptive labels connected to locations at different degrees of resolution. It is easy to see that many higher level spatial descriptions of interest to artists and designers are also extracted as they study their sketches. Those who accept the depictiveness of mental images point to exactly analogous operations believed to operate internally (Finke & Shepard 1986, Finke 1989). An example, quoted by Paivio (1971b), is how we answer the question "How many corners are there on a block letter E?". Most people report that they have to generate an image of the letter and then sequentially count the corners. The "mind's eye" at work again.

Nothing demonstrates the interdependence of depiction and description better than the evidence that people use (and need) depictive representations to solve problems of propositional reasoning. In an extensive program of research Johnson-Laird (1982 1983), has deduced that people use Mental Models when they perform logical reasoning tasks. He distinguishes mental models from propositions and mental images. Mental models according to Johnson-Laird are incomplete "structural analogues" of part of the world. They are clearly depictive in our sense. Mental images Johnson-Laird suggests are perceptual views of mental models. The assumption that is most relevant to the intermediate nature of sketches is that "the
semantics of the mental language maps propositional representations into mental models of real or imaginary worlds." In other words propositional representations are interpreted with respect to mental models. In effect he has also concluded, though by a completely different path to mine, that description and depiction are interdependent. It is not clear how depictive Johnson-Laird's mental models are assumed to be. Logically it would seem most efficient to preserve in the model symbolic representations for non-critical elements. In this case they would be intermediate types with similar descriptive-depictive catalytic properties to those I have outlined. It would be interesting to know whether he would agree that versions of Ullman's "Pictorial Descriptions" are ideal components of his mental models. Here I suggest that sequences of sketches related to individual visual concepts can be thought of as visual reasoning analogues of Johnson-Laird's mental models.

3.7 Evidence for Sketch Mediation of Type Translation.

Most design or fine art projects start with propositional constraints, vague descriptive ideas and no doubt some internal imagery of varying degrees of vividness. Indirect evidence that sketches mediate the translation of descriptively represented concepts to depictive form and back from such tentative depictions to new description, is provided by the written notes often found on sketches themselves. However as we have already seen this does not mean that the descriptive concepts implicit in sketching are necessarily verbal, only that words are the dominant means available to make non-verbal figural description available to others. A good design example is Robert Welch's fuzzy representation of the projected grooves on his serving vessel, with the bifurcating arrow pointing to two possible groove profiles. The often fleeting and unrecorded process of descriptive to depictive translation has here been captured by the designer's note (Illustration XIV). In Fashion design as well, early concepts are usually descriptively represented, and early sketches often densely larded with notes (Illustration IX). The one to many descriptive to depictive mapping is well illustrated by a sketch for an Opera Cape by the designer Charles James (Illustration VII). The notes indicate that James, C. saw the garment as either single or double breasted, with or without a collar and either 5/8ths or 7/8ths length. The fashion designer Jean Muir (1984) saw the translation of "concept" to "shape" as the most important part of the design process.

"I happen to find the working out of these shapes more interesting than thinking of the ideas in the first place"

"During the making of a collection, I constantly scribble the shapes at home in the evening to see what I'm getting, or where they can lead. In the end I have
masses of rough drawings. This is my way of hopefully making certain that I have explored all the possibilities of shape in terms of ideas - suitability of cloth - seasonal needs - enough diverse shapes to suit different figures etc."

There is every reason to believe that Fine artists are also involved in descriptive to depictive translation whilst sketching. Often however, the mentally represented descriptive notes are only implicit and must be inferred from published conversation and letters (Roskill 1963, Goldwater and Treeves 1976). Ingres drawing described by the Ingres museum as "Nude with four left arms" (Illustration XXVIII) shows literally the one to many mapping implicit in the need to translate a general proposition into depictive alternatives for comparison.

A famous exception was Delacroix who once had ambitions as a writer (Pach 1937). His notebooks from North Africa contain masterly drawings and colour studies densely surrounded by written descriptions of the scenes he is recording. The depiction seems to amplify the description which is in turn developed by reference to the depiction.

There are only a few systematic studies of sketching in practice. Goldschmidt (1991) has made an interesting study of sketching by architects. She refers to sketching as "interactive imagery" and distinguishes two forms of reasoning whilst sketching as "seeing as" and "seeing that". "Seeing as" refers to pictorial or figural reasoning and "seeing that" refers to nonfigural reasoning or arguments about the entity being designed. In a later paper (1992) Goldschmidt prefers the terms "figural" and "conceptual". She reports protocol studies (see Newell and Simon 1972), in which seven experienced architects and one architecture student participated in a specified library design problem whilst their "thinking aloud" was recorded. The protocols (recorded thought and sketching acts), were parsed into design "moves" defined as coherent, detectable design operations. "Moves" were themselves analysed further into "Arguments" which are reasoning acts leading to a design move. Her results show frequent alternation between "seeing that" and "seeing as" whilst sketching or contemplating the sketch. Goldschmidt agrees that her use of "figural" ("seeing as") and "conceptual" ("seeing that") overlaps with the terms "depictive" and "descriptive" as used here and to which she referred (Fish and Scrivener 1990). However her distinction is not the same as ours, in that "figural" does not necessarily exclude mental descriptions of visual structure and it is not clear that mentally depictive reasoning (Johnson-Laird referred to above) is excluded from her "conceptual". It must also be remembered that protocol analysis has been criticised on the grounds that the act of introspection and verbal report may change the nature of the internal processes to which they refer (Evans, J. 1980). Despite these provisos, Goldschmidt's
valuable studies are consistent with the descriptive to depictive translation hypothesis.

Further suggestive evidence is provided by Scrivener and Clark, S. (1992) in which they discuss the results obtained from a program of research intended primarily to investigate the nature of collaborative design at a distance, using remote terminals with a shared work space and with visual and auditory contact. They discovered that "sketching acts" (distinguished on the basis of timed pauses and recorded automatically), reflected meaningful components of the sketch structure. Sketch acts could be classified into those which were depictive, (eg object drawings), those which were descriptive (eg alphanumeric notes and labels) and those which they termed "split acts" in which depictive and descriptive information is realised in a "single burst of activity". "The results suggest that descriptive acts normally accompany drawing acts in sketching". The authors conclude that "a drawing act in sketching is not an attempt to represent a solution as such, rather it is a notational device that helps its creator to reason with complex and labile mental structures." Further inferences about the nature of such structures will be made in the next chapter.

3.8 Conclusions concerning Representational Type.

1. Artists' sketches belong to a continuum of depictive to descriptive representational types which vary from those in which highly concrete and specific information is conveyed implicitly by an isomorphism or correspondence with what is represented, to very abstract and general information represented by explicitly defined symbol systems and rules of interpretation.

2. Sketches are intermediate between the two extremes, being roughly more descriptive than pure drawing and more depictive than scientific diagrams or maps. Reference to a complete range of representational type is necessary to distinguish sketch components which show attributes genuinely intermediate in the depictive-descriptive continuum and global sketch attributes which arise from the mixing of components belonging to other systems, either higher or lower in the continuum. Examples of mixed depictive descriptive components are illustrations, in which written notes or numerical measurements are attached by marker lines to spatially depictive positions.

3. Sketching shares with drawing, painting, and creative writing the property of semantic symmetry (my term) since cognitive processes of equal importance are used (in a feed-back loop) both for their generation and interpretation. (This property contrasts with the semantic asymmetry of machine processed photographs and other representations which are less semantically generated than interpreted.)
4. It is proposed that an important function of the class of intermediate and mixed representational types to which sketches belong is depictive-descriptive and descriptive-depictive translation. These processes are used in a feed-back loop to support visual invention. However the two processes can be asymmetric with respect to the sources (internal or external) and levels (concrete/local or abstract/global) of the information they alter, remove or add.

5. Many of the characteristic properties of the "sketch class" of intermediate representational types, such as "context dependence", "external reference but without explicit rules", and the presence of indeterminacies leading to multiple interpretation or cognitive completion, show that they must be regarded as incomplete representational systems. Any understanding of their cultural functions will require a theory of the mental (cognitive-neural) representational systems of which they are only a small, though more visible component. The attributes of sketches are only understandable as a supplementary component of a huge system of mental representation, both conscious and unconscious with which their visible components interact. Chapters 4 to 7 attempt to unravel some of these processes further.

6. Theories of cognitive representation which may underlie the visible characteristics of sketches and which have been proposed at different levels of analysis, from neural medium to cognitive function are, on the whole, consistent with the "type translation within a continuum" theory of sketch function. I will claim that the advantages posited by Ullman for *Pictorial Descriptions*, (clearly an intermediate type) for his "matching by alignment" theory of recognition, seem to have equally tangible advantages for theories of visual thought and invention.

The debate about whether or not mental images can be depictive is filled with too many misunderstandings to be any longer fruitful. Once the physiological plausibility of neural structures acting as functional depictive systems is established the argument collapses. just as depictive theories cannot function without descriptive interpreters, so descriptive theories ultimately require depictive representation to be meaningful. Even Pylyshyn, probably the most consistent "proposition only" theorist, admits that word meanings (and by implication other descriptive representations) must at some point invoke perceptual (depictive) input to avoid circularity (Paivio 1983 Pylyshyn 1984).

The problem of finding a non-humunculus theory of the mind's eye still casts its shadow over any simplistic theory of mental depiction. It is less often pointed out that a similar paradox must be resolved with propositional (descriptive) theories. How can visual structure be represented by propositions if there is no inner mind to interpret the propositions? This problem is addressed in Chapter 7.
7. The visible attributes of most published sketches and the few available experimental studies of sketching, suggest that sketches function at the interface between continuously interacting descriptive and depictive thought processes. The realisation that depiction and description are interdependent suggests that sketching is used to turn vague descriptions and sparse depictive imagery into continuously richer ("denser") combined descriptively labelled depictions and depictively interpreted descriptions.

3.9 Summary

These conclusions are summarised in the first hypothesis about sketch function.

**Hypothesis One**

Sketches have attributes which make them intermediate between depictive (picture-like) and descriptive (language-like) representations. Such attributes are used to facilitate the bidirectional translation between descriptive and depictive modes of thought.
4. SKETCH PERCEPTION AND RECOGNITION.
"Heard melodies are sweet but those unheard are sweeter"
John Keats "Ode on a Grecian Urn" 1820

4.1 Introduction

Hypothesis Two states that "Human perception and object recognition processes enable accidentally impoverished stimuli fast automatic access to long term memory for perceptual completion. Sketch attributes exploit such unconscious processes providing memory search and retrieval cues that improve the availability of tacit visual knowledge for invention."

Chapters 4 and 5 address the issue of how sparse, untidy or confused sketch attributes amplify invention in the two ways suggested by Leonardo. It is suggested that biological processes which evolved for analysing, recognising and completing naturally impoverished stimuli are exploited by sketches for a different cultural purpose. The cognitive mechanisms likely to be involved are here conceived as divided into two stages. In the first phase automatic perceptual processes group sketch components into candidate objects which are then used with other automatic processes as "look up" keys to long term visual memory. If the keys are incomplete (e.g fragments of contour) or ambiguous and confused then it is likely that the search and match processes used for recognition will unearth more "buried treasures" from long term memory than do tidier, less ambiguous or more complete representations. This chapter examines the extent to which existing theories of picture perception and object recognition support this idea.

In the next chapter a second phase of this process will be proposed. It will be suggested that candidate recognition matches lead to transitory percept completion by spatially superimposed mental imagery. These two phases provide two distinct but related functions - perceptual retrieval and perceptual completion. The first phase is based primarily upon access to stored descriptive representations. The second phase involves the mental generation of spatially depictive representations. The mental representation of the sketch percept is viewed as a descriptively labelled 2½D spatial array (3.4.7 & 3.6.5).

4.2 Sketching and Theories of Perception.

4.2.1 The Classical Problems of Perception

Before discussing a perceptual function peculiar to sketches, it is worth considering how theories of picture perception in general can be applied to sketches. Such theories are themselves heavily influenced by which of several schools of thought about the nature of perception their authors espouse.
The problem for theories of perceptual organisation is that of explaining how the perceptual experience of a meaningful three-dimensional world of segregated objects and events is constructed from the apparently inadequate information provided by the flow of images presented to the retina. The retinal image is a rather poorly formed two-dimensional array of luminances and chromaticities. Because it is a flat projection from a three-dimensional world, information available in the retinal image about an object, confuses real size with projected size (which changes with distance from the eye), and confuses real shape with projected shape which changes with orientation. (Gregory 1969, Kaufman, 1974, Wilding, 1982). Moreover, useful colour information about an object (its surface spectral reflectance) is not given explicitly in the retinal image, since information about local luminance and chromaticity does not distinguish surface reflectance from the colour and intensity of the illumination (Beck 1972).

Despite this, the visual system is able, in natural scenes at least, to distinguish object size and shape from object distance and orientation and object surface reflectance from surface illumination, providing remarkably valid perceptual experience. Thus visual experience is the end result of cerebral processes which extract the "invariances" of three-dimensional objects from fleeting stimulus information (For theoretical reviews see Bruce & Green 1986, Gordon 1989). The many "accidental" variances in the image caused by vagaries of view-point and illumination, although of interest to artists, are mostly irrelevant to the biological functions of recognising and interacting with three-dimensional objects. So in representing and manipulating stationary views of imaginary objects, the brains of artists must exploit visual mechanisms which have evolved for responding to real objects of survival value in the natural environment of our hominid ancestors. These mechanisms, which extract the "constancies" of the three-dimensional world, discount or under-represent in visual consciousness many of the view dependent features of vision that are of interest to artists.

Only a very small central region of the retina, the fovea, receiving light from a visual angle of about 2° is capable of high resolution colour vision. Away from the fovea the retina is progressively less able to distinguish detail and colour although it remains sensitive at low resolution to movement and to low contrast stimuli. For this reason the normal viewing of a sketch consists of brief fixations of small fields of interest lasting 200-300 milliseconds followed by sudden unconsciously computed jumps to new fixations points ("saccades") (Yarbus, 1967). There are two separated classes of signal which can be detected arriving from the eye, a transient, fast signal sensitive to low contrast and a slower signal sensitive to detail and colour. There is evidence that during a saccade the transient
low resolution signal is suppressed, probably because it would generate erroneous signals about movement (Burr et al. 1994). It is likely that the transient system is used to select regions of interest and to drive the eye to new locations for more detailed processing (Schiller 1985). Despite this we perceive sketches like real scenes as integrated wholes. Thus any theory of sketch perception must explain how the brain is able to integrate a scattered sequence of saccadic foveal "snapshots" and interpolate briefly remembered detail for parts not fixated. This point has been emphasised repeatedly by Hochberg (1968, 1980, 1984) and in used in his theory of picture perception referred to below.

Rock (1983) has classified the variety of theoretical approaches to perception into "stimulus theories" and "constructive theories". The latter he subdivides again into "interaction theory" and "cognitive theory". These three classes are roughly synonymous with the "Gibsonian", the "Gestalt" and the "Helmholtzian" schools of thought.

4.2.2 The Gibsonian Theory of "Direct Perception" and its Influence.

Stimulus theories attempt to explain perceptual performance in terms of a detailed analysis of the information provided to the senses (the "stimulus"). The most influential and complete of such theories is J.J. Gibson's theory of "Direct Perception" (Gibson 1950, 1966, 1971, 1979). Gibson and his followers assume that the nervous system responds to this information directly in ways which explain the structure of perceptual experience. Gibson's achievement has been to show the richness of the optic information available in natural, "ecologically valid" conditions. In his theory perception is an active process of exploration in which the nervous system responds directly to "higher order" stimulus relationships which are sufficient to define the "invariants" of real objects and scenes. The Gibson school denies the importance of either memory or stored mental representations in perception. Gibson placed great emphasis on the fact that in normal perception we receive a dynamic family of changing views of objects and scenes in which certain relationships remain constant. The information "picked up" by the visual system is not that contained in static retinal images but in those relationships which can be shown to remain invariant in the changing pattern of light as we move our heads and bodies and as objects come in and out of view. As Marr (1982) has stated, Gibson asked the "critically important question, How does one obtain constant perception in everyday life on the basis of continually changing perceptions?". According to Marr the fatal shortcomings of Gibson's analysis was that he failed to realise that the detection of physical invariants is an information processing task and that he "underrated the sheer difficulty of such detection".
Gibson's experimental work drawing attention to the importance of such factors as perspective, texture gradients, the horizon line, movement parallax etc in explaining the perception of three dimensional surfaces (and no doubt some of the attributes of sketches) deserves respect. However, as a contribution to the understanding of sketch attributes, there is a serious problem with the theory of direct perception. Many of Gibson's claims are incompatible with the theory presented here which relies unashamedly upon a "stored representation" account of perception. Gibson dismissed such "mind's eye" theories of perception as old fashioned "mentalism" and the theory that we "project" memory images onto indeterminate stimuli even as "mischiefous nonsense" (Gibson, 1979, page 282). It is possible to answer Gibson's claim that only "direct perception" has ecological validity by pointing to the ecological necessity for early man to make fast perceptual decisions in conditions of transitory, poorly illuminated, camouflaged, or partly concealed object stimuli and the evidence that attention limits the capacity to process all the information available to the eyes. But an even stronger refutation of this negative aspect of "Gibsonism" emerges when efforts are made to suggest how such higher order "pick-up" might be implemented by the brain (Marr 1982). Shepard in a series of papers (1975, 1981, 1982, 1984) has tried to reconcile his theory of internally represented "secondary isomorphisms" with the Gibsonian invariants and "affordances". "Although I agree with Gibson that the brain has evolved to extract invariants under favourable conditions, I also presume that it has evolved to serve the organism under less favourable conditions of night time, obstructed and spatially or temporally limited viewing and even of structural damage to the brain itself" (Shepard, 1984). Unlike Gibson, Shepard believes that the nature of internal representations was the crucial issue for the understanding of perception.

The theory of "Direct" perception (Gibson 1979) and the denial of any role for attention, memory and internal representation is so different from the theory of sketches proposed here that it would be tempting to skip any discussion of the Gibsonian school. However this is impossible due to the influence (and potential future influence) of his theories on the design of visualising systems (Smets, 1992). Gibson's theory has also been influential in theories of picture perception (Hagen 1981 1982 1986). Pictures, especially sparse indeterminate ones present a problem for direct perception theory. For pictures, even very realistic ones like photographs do not contain any of the movement based cues to object invariances which Gibson maintained were important. Moreover, the information provided by gradients of texture, shadow and linear perspective and by exact relationships to the ground plane and the horizon may be present in photographs, but are missing in most
sparse or untidy sketches. Gibson's own theory of picture perception, on which he wrote repeatedly, has changed during his life (Gibson 1971 1979, Hagen 1981). Originally he proposed that pictures, "re-presented" the array of light presented to the retina by real objects. "A picture is a surface so treated that a delimited optic array to a point of observation is made available which contains the same kind of information that is found in the ambient optic arrays of an ordinary environment" (Gibson 1971). In other words they succeeded to the extent that they mimicked photographs. This theory hardly accounts for the sketch attributes discussed in Chapter 2. Later he modified this view and claimed that they capture aspects of the "Invariant information" available in the ordinary environment (Gibson 1979). This later theory was never explained in detail and it has been left to one of his disciples - Margaret Hagen, to develop from Gibson's theory of stimulus invariances what she has called a "Generative theory" of picture perception. Briefly, Hagen suggests that although a particular perspective view in a picture may be ambiguous in theory, it belongs to a "generative set" of views which are known to, and can be generated by, the viewer. "The static information in a perspective picture may in this light be regarded as the frozen presentation of a single aspect, one instance generated by a rule, and therefore perceivable by any possessor of generative percepts." (Hagen 1981). The psychological availability of such generative rules is shown, Hagen claims, by the facility with which novel views of objects and scenes are perceived in pictures even by small children. In framing her Generative theory Hagen claims to be closing the gap between Gibsonian theories of pictures and other theories which assume a crucial role for memory and "unconscious inference". However there is no admission that novel views of objects might be generated, not only from perspective rules, but also from images created internally from representations of remembered objects. "All visual experience will be in accord with the rules of ecological geometry. An object of a particular form looked at from a particular place will have a particular appearance, and it does not matter that this particular appearance was never beheld before. It is generated by the projective rules which govern the appearance of things. The particular application of these rules with regard to specific objects and events does not require the laborious association of experiences from the past." This aspect of Hagen's theory is in contrast with the theory to be proposed here.

4.2.3 The Gestalt School and its Influence.

The Gestalt school of psychology which flourished in the 1920 and 30's was pioneered mainly by Max Wertheimer, Kurt Koffka and Wolfgang Kohler. Their writings and those of their disciples have left a lasting influence on theories
of perception, more by the phenomena they described (such the "Laws of Perceptual Organization" and the principle of "Pragnantz") than by their theories about how this might be implemented in the brain (see Koffka 1938, Kohler 1947, Ellis, W. 1938). The latter have been overtaken by contemporary neuroscience and need not be discussed here. Gestalt psychologists emphasised the organisation of perceptual experience into "wholes" with "emerging" features which could not be understood, they believed, by an analysis of the component parts. Wertheimer’s paper demonstrating the ways by which perceived elements are involuntarily organised according to the rules of proximity, similarity, good continuation, closure and common fate is directly relevant to one of the functions of the sketch, that of manipulating disparate units of form and colour into a single visual whole. Their documentation of the properties of figure ground separation is also important to the visual arts. The Gestalt psychologists proposed laws of perceptual organisation such as the principle of "Pragnantz" or simplicity which they believed were innate. The Gestalt psychologist often used simple line drawings to illustrate their proposed "laws" and thus their work is incidentally relevant to sketches. Koffka’s classic (1938) chapter on Figure and Ground relationships could almost be taken as part of a description of sketch attributes. For example he pointed out that when an outlined shape is seen as figure, the contour line is always perceived as belonging to the figure. The figure he found is more complex than the ground, the texture clearer, the colour more saturated than the ground. Whilst more recent research has confirmed many of the Gestalt school's findings, theories about their significance have changed. It now appears likely that the "laws of organisation" reflect the conscious products of early automatic, and "pre-attentive" stages of visual processing, used to extract edges, surfaces and candidate objects for further processing (see Chapter 6). This, at least, is the interpretation favoured here. Gestalt theory however fails to account for the fact that in addition to perceiving "emergent features" and "integrated wholes" we must be able to perceive and attend to segregated image components. Because they never developed a plausible theory of object recognition Gestalt theory fails to explain how inventive imagery is elicited by sketch indeterminacies.

Rudolf Arnheim (1966, 1970, 1974) has made a comprehensive statement about the relationships between Gestalt psychology and art. He has studied the connection between Gestalt laws and visual composition and form perception in painting and drawing (1974). He has gone rather beyond the Gestalt theories in his notion of the "visual concept" which he uses to explain the way pictures represent. Arnheim’s "visual concept" of an object "conceives of the object as being three-dimensional, of constant shape, and not limited to any particular projective aspect"
(1974). In adults and artists the visual concept "embraces the multiplicity of its appearances, the foreshortening, the slants, the symmetries and asymmetries, the partial concealments and the deployments, the head on flatness and the pronounced volumes" (Arnheim 1970). In Arnheim's theory of pictorial representation the image is considered to be an expressive equivalent of an object not a "snap shot of it". Further, objects must be represented from (one of) their most characteristic viewpoints in order to successfully accomplish this equivalence. According to Arnheim, certain projective views of any object contain what he called "Renvois" that is references which point beyond a given view to adjacent ones. The difficulty of interpreting uncharacteristic views is shown in the well known joke drawings, "a giraffe's neck passing a window" or "a Mexican riding a bicycle". Whilst it is true that an artist or designer may occasionally have cause to sketch an imagined object from an uncharacteristic view, very few such views occur in published sketches (e.g. Picasso (Glimcher and Glimcher) 1986, Welch 1986, St Laurent 1987). Experimental support for Arnheim's theory is discussed later (4.3.5).

4.2.4 Helmholtz's Theory of "Unconscious Conclusions".

It is an implicit (though perhaps unacknowledged) fact that both the Gibsonian and the Gestalt approaches to perception assign to neural physiology very complex but unexplained information processing tasks. In the Gibsonian school this task is the "picking up" of higher order invariances from the changing flow of the "ambient optic array". Their metaphor of "resonance" in which the nervous system is compared to a tuned radio circuit appears to absolve the psychologist from any need to consider how visual information might be represented internally. In the Gestalt approach the task of an innate and largely automatic, physiology is not so much to pick up structure already present in the optical environment but to add structure to the unstructured stimulus before it reaches conscious awareness. Thus the Gestalt school is more concerned in describing the results of innate visual processing of stimuli. Both schools emphasise the determinate aspects of perception. In contrast the Helmholtzian approach emphasises the indeterminate probabilistic aspects. Helmholtz, the father of this approach, likened conscious percepts as the "conclusions" of unconscious inferences. In a much quoted chapter of his Treatise of Physiological Optics he refers to how an astronomer may compute the distances and positions of the stars using a knowledge of optics. "In the ordinary acts of vision this knowledge of optics is missing. Still it may be permissible to speak of the psychic acts of ordinary perception as unconscious conclusions, thereby making a distinction of some sort between them and the common so-called conscious conclusions. And
while it is true that there has been and always will be, a measure of doubt as to the similarity between the results of the psychic activity in the two cases, there can be no doubt as to the similarity between the results of such unconscious conclusions and those of conscious conclusions. " Thus what we experience when we see is not given to us by the stimulus (as the Gibsonians claim) but has meaning and structure added to it by some inferential process. Richard Gregory one of the strongest supporters of the Helmholtz approach has developed the idea by making a detailed comparison between perceptual inference and the inferences of science (Gregory 1970, 1973, 1980). Irwin Rock (1983) has also made a strong case for visual processes which resemble reasoning. He distinguishes, "form construction" which he says occurs by a process of description from "problem solving" which is used to resolve ambiguous stimuli and "unconscious inference" proper where perception reflects a rational solution to an indeterminate stimulus based upon "assumptions" about the real world. A moderate Helmholtzian view is also taken by Julian Hochberg (1968, 1981, 1984). Hochberg has emphasised the dependence of perception upon stored representations. His studies, showing how object percepts can be mentally completed from incomplete line fragments viewed sequentially through holes, is relevant to the problem of how very sparse sketch components are recognised. The problem with unconscious inference theories is that the nature of consciousness and therefore of the difference between unconscious and conscious inference is left unspecified. Helmholtz's disciples would recognise that "unconscious inference" is distinguished from ordinary inference in that it occurs without voluntary control and that it is not subject to much conscious knowledge about the object perceived. Thus illusions and other non-veridical perceptions persist in their appearance even when the subject has full knowledge of their non-veridical nature. In contrast conscious inference must surely take all the known facts into account before framing a hypothesis. It is important to note Helmholtz's own caution in the quoted sequence above - "there will always be a measure of doubt as to the similarity between the results of the psychic activity in the two cases". Should this doubt not cause one to insert the words "as if" before unconscious inference accounts of perception? Another possible criticism is that unconscious inference accounts say nothing about how perceptual "conclusions" are represented. The term "inference" suggests a propositional or descriptive reasoning process. However if, as I have argued in Chapter 3 percepts have a strongly depictive component then how is that component derived by inference? An unconscious inference theory might explain in the "Hidden Man" demonstration (Illustration XXXII) why the flat black shapes can be descriptively categorised as a head but a further theory is needed (in my
opinion) to explain the depictive change that is experienced when a pattern of flat shapes suddenly becomes a three-dimensional head and shoulders. Despite these reservations, the Helmholtzian approach has the most to offer of the three schools for the understanding of sketch attributes. Because Helmholtz's disciples believe that perceptual experience is under-specified by normal stimuli they are naturally interested in illusions and in the way we perceive ambiguous and incomplete images. They have a vested interest in being able to explain how impoverished stimuli such as pictures can be so rich perceptually.

As Gregory remarked (1970 page 32) "Pictures have a double reality. Drawings, paintings and photographs are objects in their own right - patterns on a flat sheet - and at the same time entirely different objects to the eye. We see both a pattern of marks on paper, with shading, brush-strokes or photographic grain, and at the same time we see that these compose a face, a house or a ship on a stormy sea. Pictures are unique among objects; for they are seen both as themselves and as some other thing, entirely different from the paper or canvas of the picture. Pictures are paradoxes." If this is true of pictures then indeterminate sketches are even greater paradoxes for in them we see not one stable "other thing" but many changing things. Two paragraphs further on Gregory makes a comment which can be developed in terms of the depictive - descriptive translation function of sketches (Chapter 3). "Perhaps man's ability to respond to absent imaginary situations in pictures represents an essential step towards the development of abstract thought. Pictures are perhaps the first step away from immediate reality, and without this reality cannot be deeply understood." (Gregory 1974). Later in the same work Gregory discusses how writing may have developed from pictures and pictographs, which perhaps indicates the direction this earlier thought was taking..

Julian Hochberg' "Mental Structure" theory of picture perception is also Helmholtzian in tone and perhaps comes closest to this thesis in explaining the nature of sketches, although it is tantalisingly incomplete. In agreement with Gregory, Hochberg believes that "The fact that pictures appear to the viewer to be so similar to the things they represent, although they are so different physically, is a major fact that any theory of perception must confront" (Hochberg 1980). He then goes own to argue with numerous examples, that pictures display the operation of mental structures -

"that is of visual knowledge, not given in the stimulus display, about the physical properties of objects and of their spatial relationships. If that is true, where do the structures come from? Why should we have these abilities, if direct response to information in the visual stimulus array that is presented by the world normally suffices for veridical perception ?

"
The answer to that puzzle, I believe, is that pictures are not unique in being ambiguous and incomplete: The objects of the real world, as they are glimpsed within each momentary glance, are usually partially hidden from sight, and are ambiguous and incomplete as far as usable stimulus information is concerned... I believe that mental structures of sensory expectation are developed to bridge the successive glances at the world; that the difference between realistic pictures and the scenes they represent may be negligible to the momentary glance; and that pictures therefore draw upon mental structures normally developed in the service of seeing the real world”.

4.2.5 The Ideas of Ernst Gombrich.

Ernst Gombrich’s theories of picture perception have already been referred to. In his erudite description of the historical relationship between style and culture in figurative art, he presents a theory which is broadly Helmholzian. Thus what we see in a painting or sculpture depends both upon cultural convention and upon the unconscious hypotheses we frame about what is represented. These in turn cause the viewer to unconsciously "project" his or her hypotheses upon the indeterminate components of the works we are looking at (Gombrich 1962, Chapter 4). How this is possible is explained by the theory of "projection" used by Rorschach when he invented his ink blot aid to psychoanalysis. Rather paradoxically Gombrich also gives credence to Gibson's theory of direct perception (Gombrich 1982b). Gombrich is the only one of the above authors who specifically addresses the problem posed by the deliberate use of indeterminacy to stimulate invention either as a private working method or to increase the beholder's share of the finished work.

4.2.6 Perception as a Multi-Stage process: Towards a Synthesis

I have attempted to summarise the most widely adopted stances on perception together with their consequences for picture theories for two reasons.

Firstly I want to show that theories of the design of new sketching systems can be influenced, in crucial ways, by which of the rival theories of picture perception is adopted. For example, the Gibsonians believe that the information necessary to explain perception is available and "picked-up" from the flow of light presented by the environment with no necessary involvement of knowledge or memory. It follows from this that generating realistic imagery which mimics the changing light patterns presented to the eye by real objects, is the natural way to assist artists to visualise imaginary objects. It is therefore not surprising for example, that Professor Smets should entitle the presentation of her laboratory's interesting work on very realistic imaging systems for designers "The Theory of
Direct Perception and CAD in Virtual Reality" (Smets, 1992). Unfortunately, there was no time at the symposium at which she delivered the above paper, for me to question her on this point ("Virtual Representations for Design and Manufacture" Coventry University, 1992). However it is clear from the title and contents of her paper that her approach towards visualising systems for designers is strongly influenced by, or even based on, Gibson's theory. The adoption of a theory of unconscious inference or mental structure would support a contrasting approach. As Hochberg and other theorists believe, perception is forced by the limited capacity of saccadic integration, foveal attention, etc to interpolate by inference or expectation the unnoticed components of a scene. It follows that very detailed over-complex imagery may distract automatic attentional processes from what is relevant or important (see Chapter 6) and hinder inventive processes of mental completion.

Secondly, none of the above theories of picture perception is sufficient to explain the "privileged" attributes of sketches. All the existing theories leave unanswered questions about the nature of the representations or physiological processes which the brain uses to interpret sketch attributes and to interactively generate from them new visual ideas.

The constructive or positive aspects of the three schools of thought about perception are not mutually exclusive. More recently attempts to model perception and cognition as information processing (e.g. Marr 1982, Watt 1988, Rumelhart et Al. 1988, Humphreys & Bruce 1989) and recent ideas in visual neuroscience (Hubel & Livingstone 1987, Kosslyn & Koenig 1992, Zeki 1993, Milner & Goodale, 1995) suggest the possibility of a new synthesis using parts of all the above theories. Thus Gibson's study of the higher order variables in natural stimuli may indeed help us to understand how the brain acquires conceptual knowledge from novel stimuli. However this does not mean that all the ecological information available to the visual system, is actually used by the visual system (as Gibson claimed), as the eye quickly samples a visual scene with a few fragmented glances. It is surely true that even granted the richness of the information available in natural scenes, knowledge and experience normally provide important enhancements of perceptual experience. The gestalt laws of perception may prove to be revealing consequences of important but early, pre-recognition, "bottom-up" processes which derive from the retinal images a structured representation containing candidate objects. This temporary representation is then subjected to a second phase of "top-down" processing in which an automatic search occurs for a match between the partly structured early representation and stored representations in the same format in memory. Recognition is fast. It is at this second stage that it
is plausible to postulate an inference-like filter to constrain what would otherwise be a laborious memory search. There are two ways in which object, scene or event probability might result in perceptual experience which resembles inference. Firstly, innate visual mechanisms are likely, by natural selection, to reflect unchanging expectations in the natural environment, such as the effects of gravity on objects and movement and the most frequent distributions of illumination and texture on the ground plane. Secondly, visual experience since infancy must have caused the construction of organised priority lists in memory of the most likely objects and events to be anticipated in different contexts. If this priority is constructed in a sufficiently adaptive manner then perceptual "conclusions" as if an unconscious inference has occurred might result.

Thus the stance taken here is that conscious percepts are a synthesis from many specialised visual processes and that sketch components trigger such processes, many of which are automatic and unconscious. The evidence that perceptual experience refers to a late construction in a multi-component component process of visual analysis, with many parallel and sequential stages, is persuasive. The studies by Marcel (1983a 1983b) on backward masking seem to prove, at least for word perception, that both memory access to the shape of a word and memory access to the meaning of a word, occurs before we are conscious of seeing the word at all. Marcel displayed a pattern mask at varying intervals after a briefly displayed word. Having found with each subject what interval (termed the SOA, "Stimulus Onset Asynchrony") allowed a 60% correct identification of the word, he then reduced the SOA to the point where his subjects told him they could not see the word at all. However when asked to guess which of a list of words was similar to (a) the shape of the target word or in another experiment to (b) the meaning of the target word, they consistently guessed (statistically) correctly much more often than chance. Thus information about shape and meaning appeared to be accessed before perceptual awareness. A fascinating point is that as the SOA becomes shorter and masking is greater, first visual awareness is lost, then word shape perception and finally word meaning. This order is the reverse to that which would be expected by "bottom-up" processing theories, since earlier masking would be expected to interfere with earlier stages of processing which surely must progress in the order - "analyse form, look up meaning, construct conscious percept". Marcel's explanation is that visual information only becomes conscious (and hence a "Percept") when all the different separately processed components can be successfully integrated. This reverse order of masking then perhaps makes sense if masking interferes with the construction of the late percept. In this process word meaning could be accessed and held in a temporary store before word shape is re-
used for the final integration. This interpretation may be wrong, but it is consistent with the many other experiments which have shown that briefly displayed, incomplete or ambiguous percepts are modified in subjects' reports to agree with their semantic expectations (Carmichael et al. 1932, Vernon M.D. 1970). Marcel's late construction theory is important to picture perception because of the distinction I have already drawn between semantically and non-semantically (ie "contingent") generated representations. Photographs and computed photographic simulations present or fabricate uninterpreted information similar to frozen retinal images. Sketches like drawings and paintings present information which has been selected from the structured and semantically labelled internal representation of conscious percepts or visual memories. Thus contours, figure ground relationships, scene structure and implicit meaning are presented as already given to a visual system, designed by evolution to extract these features from much more complex natural stimuli. However the "late construction" theory does not necessarily imply that the proximal (projected) sizes and shapes of objects are not available to artists when they sketch. It is true that the results of object perception and recognition usually means that we see familiar objects as preserving their size as their distance changes, but as Kaufman (1974) has pointed out there appear to be two modes of perception depending on how we attend to the information provided by our eyes. Thus normally when I am driving I see the car in front of mine as its normal size and am very aware of its distance. However by switching attention it is also easy to see that the height of the image of the car is smaller than that of my wind-screen wipers. This dual mode of perceiving is perhaps especially developed in artists and designers. It may be part of the answer to Gregory's "paradox" of picture perception where we perceive the plane of the picture and the proximal sizes of sketched objects, whilst simultaneously perceiving the represented distances and the intended object constancies (although these are more dominant in real views).

Perhaps the best parts of these rival theories of perception can be reconciled by assigning to each a different phase of this multistage process. None of the above theories of picture perception offers a convincing theory of sketch attributes. However there are elements of insight in each, to which the present theory owes a debt. Especially useful is Hochberg's comment that picture perception does not rely on special cognitive processes, but that it uses "Mental structures" which evolved not for pictures but to enable the biologically adaptive response to ambiguous and incomplete stimuli in the real world. In order to develop this idea it is necessary to ask the following questions.

1. How are sketch components represented for perception?
2. What type of memory representation might they interact with?
3. What sort of memory matching process is used?

4. What type of process might enable the mental completion of sketches? How is it distinguished from a recognition match?

4.3 Sketches and Object Recognition.

4.3.1 How Sketches Exploit Recognition Mechanisms.

In so far as sketches are private representations primarily of value to their creator, it might appear that there is no recognition problem. In generating lines, hatching, colour washes etc the artist usually knows what these marks are intended to represent before they are made. So she or he can hardly fail to recognise what they signify. This misses the point. As we have seen, the sketch marks are there to assist their user to imagine new visual structure and this implies perceiving unexpected relationships and meanings in the marks made. According to Hypothesis 2 sketch attributes exploit recognition mechanisms which evolved for the perception of accidentally, impoverished stimuli in the real world. If some of those recognition processes are involuntary and automatic then sketch components may have besides their intended meaning unintended, unexpected and sometimes useful perceptual consequences. It is the automaticity of recognition that causes the untidy pentimenti in Leonardo’s sketch for the Battle of Anghihari to suggest, unexpectedly, Neptune rising with his horses from the sea (Illustration II). As I shall argue later (7.4) novel retrieval permutations from unconscious memory may be an important component of visual inventiveness.

The term recognition refers to a cluster of functions which have in common the implicit association of stimulus derived visual information with information in memory. At the lowest level of sketch structure the artist recognises simple shapes, curves and colours associating them, perhaps unconsciously with remembered contexts. These shapes and colours are normally organised by the visual system into particular two dimensional patterns representing possible objects separated from their backgrounds. Visual recognition evolved to recognise three dimensional objects in three-dimensional space and may discount view-specific position, scale, orientation or colour (Sutherland, 1968). For this reason, I suggest it is difficult for painters and designers to make marks on two dimensional surfaces which do not have three-dimensional connotations. Theories of recognition attempt to account for the fact that we can recognise the same object from any view point, even though many of these view points will produce novel two-dimensional projections on the retina. However in sketches many of the stimulus cues, believed to be used by the visual system to infer three-dimensional structure in real scenes, are always missing (e.g. movement parallax, retinal disparity) and others are often
missing (e.g. gradients of shadow, gradients of texture and perspective projected from the correct view point). But in addition, recognition is the ability to categorise a represented object in many different ways. This type of recognition surely demands a hierarchy of semantic associations in long term memory. These three visual achievements, pattern recognition, object view recognition and object categorisation are usually but not necessarily linked. Thus one may recognise the shapes made by a tree without realising that it is a particular tree one has seen before. One may recognise a tree as a particular one without knowing what kind of tree it is. One may recognise a tree as an oak tree, or just a tree, or just as a strange plant, even though much of its shape and many of its features are unfamiliar. Finally there is the ability to give verbal labels to the tree and its parts which in turn is used to access linguistically represented knowledge. Humphreys et Al. (1988) suggest a multi-component stage model for picture recognition supported by neuropsychology and the results of "priming" experiments. They suggest that three processes of visual recognition occur in cascade (i.e. the second stage starts before the first has finished). Early bottom up visual processing derives a view-point dependant object description (such as Marr's 2½D sketch). The first stage of recognition, "Perceptual Classification" compares this early representation with stored descriptions of object appearances. The second stage, "Semantic Classification" concerns recognising the function and categorical properties of an object. The third stage "Naming" associates the object and its parts with verbal labels. This is an interesting proposal which seems to provide scope for top down semantic influences upon ambiguities encountered at the perceptual stage. Thus tentative semantic "hypotheses" may be available in cascade before the earlier perceptual classification is complete. As these authors have also pointed out there are, almost certainly, many different routes to object recognition depending upon the recognition task, its context, the degree of object familiarity and the type of information which characterises an object category.

The following theories of recognition are sometimes presented as competitors. In my view this is a mistake. It is more likely that each theory captures different aspects of the complex processes available for the perceptual recognition of sketch components.

4.3.2 Feature theories

Feature theories of recognition assume that objects have local features which can be detected directly by the nervous system. Candidate features are usually assumed to be simple geometric properties such as vertical and horizontal lines, curves, angles, corners etc. (Selfridge et Al. 1960). In its simplest form a
recognition match consists of comparing a weighted list of object features in long term memory with the relative activation strengths of the feature detectors. Some approaches assume that objects have certain invariant features which are common to all their views (Pitts et al. 1947). Others point out that some objects may possess a unique "landmark feature" which allows identification of an object at any orientation (Humphreys and Quinlan, 1987). The trouble with feature theories is that if features are assumed to be simple, then many different objects possessing the same features are not distinguished. If the features become complex, then the problem is deferred until we know how the features themselves are recognised. Thus in sketches of figures an eye might be considered to be a landmark feature for a face, but such a degree of specificity would multiply the number of feature detectors required to astronomic numbers. It makes sense to assume that feature detection is only one component of more complex recognition processes.

4.3.3 Structural Descriptions

Structural Descriptions are theories of representation based on the premise that the common properties of object views and object categories consists of the structural relationships between the parts. At least two kinds of Structural description have been proposed.

The first is purely symbolic, consisting of nodes representing object features linked by non-verbal propositions expressing the relationships between these nodes. This attempts to solve the problem as to how we recognise so easily novel exemplars of objects which occur in a multiplicity of shapes and colours. Thus the letter "T" is represented by an approximate vertical line and an horizonatal line linked by propositions stating how they are connected. If such descriptions were coded in hierarchical nets with the "part" nodes at one level being linked to further subordinate descriptions the system becomes potentially very powerful. Since only defining relationships need be represented, tolerances can be provided and default values given for variables provided by object and scene schemata (Minksy 1975). Palmer (1975a, 1977) has developed this idea, suggesting structural descriptions for simple shapes such as cubes and abstract linear figures. He provides converging evidence to support this framework in a series of experiments using simple stick figures, studying how people perceive and rate as "natural" object parts, and on the time taken for a mental synthesis of spatially separated parts into unitary figures. The difficulty with such purely descriptive theories is that they would have to be immensely complex to describe natural objects such as human figures (see Palmer's descriptive cube 1975).
Another form of "Structural Description" assumes that the spatial structure of an object is represented in a coordinate system after the extraction of an axis or frame of reference. The term "coordinate system" is here used loosely to mean any system of representation in which object spatial relationships are referred to by relative distances and directions within a single frame of reference. In terms of the depictive descriptive continuum discussed in Chapter 3, coordinate systems are intermediate types in which the descriptive elements include the rules for translation to depictive elements by a mapping into a real or imaginary spatial medium. A question which has been debated is whether such long term representations for recognition use an object centred reference frame or a viewer-centred one (Jolicoeur & Kosslyn 1983). Marr and Nishihara (1978) proposed an object centred representation, using an hierarchically structured coordinate system of generalised cones around longitudinal axes. Thus a human figure is represented globally around a single vertical axis, and in increasing detail, by subordinate axes for trunk, head, arms and legs. An arm can be represented by jointed axes for upper arm and fore-arm and hand. A hand has separate sub-ordinate axes for fingers etc. The Marr and Nishihara model is elegant and carefully thought out. Its hierarchical structure addresses the problems of the need to represent the structure common to high level and lower level shape categories allowing for successive refinement. The authors claim that it can be derived from Marr's viewer-centred "2½D sketch" (Marr 1992), using natural constraints, by a mechanism they term the "image space processor". They rejected viewer-centred representations primarily because of the large number of views which they believed would have to be stored for each object.

4.3.4 Theories of Recognition by Components

An interesting class of theories attempt to explain how an object's structural description might be derived from invariances cued by typical object views. These suggestions are relevant to sketch theory, because they have been applied to object recognition from contour profiles and to incomplete line drawings (see Illustration XXX). They are also useful because they address the problem of how we recognise objects which present incomplete contours and parts because they are partly obscured by other objects, as is often the case. I will suggest that this capacity for recognition from incomplete information is exploited in sketches for a secondary cultural purpose.

Hoffman, D. and Richards, M. (1984) suggest that the visual system decomposes object shapes into parts at regions of concavity along the visible object contour. Contour lines in sketches fit Marr's classification into: (1) contours that
define an object’s silhouette or surface discontinuity ("self occluding contours"), (2) contours that mark a change in orientation of a surface and (3) contours that mark the edges of a shadow or change of reflectance within a surface. (In sketches lines are of course also used for non-contour purposes such as representing shadow and texture). It is the first of these contours which are considered by Hoffman, D. and Richards, M.. According to their theory part boundaries are defined by a rule which exploits a uniformity of nature termed "transversality". This regularity occurs because "when two arbitrarily shaped surfaces are made to interpenetrate they always meet in a contour of concave discontinuity of their tangent planes". Thus the profile of a face divides naturally into nose, lips forehead etc according to this rule. The authors conclude that the visual system exploits transversality to categorise objects parts as one of the regularities of nature which underlies the inference of parts from images. These parts can then be used for building an object description for recognition. The theory has the merit that it can help to explain how we are able to recognise object silhouettes when part of the silhouette is hidden behind another object.

Starting from similar premises Biederman (1987) has developed a theory in which the input image is segmented at regions of concavity into 36 kinds of simple volumetric components such as blocks, cylinders, wedges and cones. These are then used as primitives to derive structural descriptions of objects which are invariant for position, size and orientation. Hummel and Biederman (1992) have implemented a computer neural net program to illustrate the theory. This uses line drawings of simple imaginary objects to derive representations which can be used to recognise the same objects from any view point, size or position. However it has yet to be shown that Biederman’s model will work with drawings of natural objects such as human figures which present more complex problems. In support of this theory Biederman and Cooper, E. (1991b) have investigated the recognition of line drawings with partly deleted contours. Such incomplete or missing contours are often found in sketches (Illustration X). They found that when contours are deleted in such a way as to leave regions of concavity (corners and edge terminations) intact object recognition is still easily achieved. However when an identical degree of contour deletion includes the concavities posited to define part boundaries then recognition is difficult or impossible. They showed further that contour deleted images of the first (completable) type primed recognition for the complementary image. (consisting of the missing contour segments). They conclude from this and other priming experiments that contour deleted images provide evidence for representations intermediate between the input contour view and a structural description of the relationships between the posited "geons".
4. Sketch Perception and Recognition

4.3.5 Object Position, Size and orientation.

A key issue for theories of sketch recognition is how the stored representations used for recognition achieve independence from the three irrelevant metric variables of viewed objects - position, size and orientation. It would be grossly inefficient for the brain to store multiple representations of every familiar object in every position, size or orientation in which we are able to recognise it. The possibility favoured by proponents of structural descriptions is that the stored representation used for recognition matching is in a form which is invariant for position, size, orientation and view-point. In a series of papers Biederman and Cooper, E. present evidence based upon the effect of priming (prior exposure of the object stimulus) that the representation used is independent of position in the visual field, independent of a reflection about the vertical axis, independent of size and independent of orientation (Biederman and Cooper, E., 1991a, 1992 Biederman & Gerhardstein 1991).

Whilst there is little reason to doubt the fact that object recognition is position invariant there is some apparently contradictory evidence suggesting (in some kinds of recognition at least) that prototypical views and normalised sizes may require a matching process which must be preceded by preliminary size scaling or mental rotation of one of the two representations.

The clearest evidence for size scaling and mental rotation is provided when the task is to match two simultaneously presented shapes. Thus the time to specify whether two random polygons are the same is monotonically related to the difference in size, suggesting that the match must be preceded by a size equalising process (Bundesen & Larsen 1975, Larsen & Bundesen 1978, Bundesen et Al. 1981). The famous results on mental rotation (Shepard & Metzler 1971, Shepard & Cooper, L. 1982) demonstrated under many conditions a monotonic relationship between the time to make same or different responses to simultaneously presented views of three-dimensional shapes and the angular orientation difference. The fact that the inferred rotation occurred equally fast in depth as in the picture plane suggests that subjects must have been able to form a three-dimensional representation of the abstract shapes used. The tasks in these experiments differ from normal recognition in that the matching does not implicate long term memory. Unlike the contrasting Biederman and Cooper, E. results where the task was the categorical recognition of very familiar objects, these matching experiments used either unfamiliar shapes or attributes such as handedness, which it has been argued, require rotation because they are not represented in orientation independent form (Shepard & Cooper, L., 1982). Unlike categorical recognition, such transformations for simultaneous comparison are usually assumed to be
related to conscious image manipulation in short term memory. In a later study Larsen and Bundesen (1978) produced evidence for a different process of size scaling, for pattern and character recognition, which is faster and, according to the authors, is a scaling operation used for matching stimulus patterns against representations in long term memory. Jolicoeur & Landau (1984) found strong systematic effects of orientation in the identification of letters and numbers. However the effects are much smaller than those found in the mental rotation experiments, suggesting again a different process. In another series of studies Jolicoeur (1985) found systematic effects of orientation in the time to name line drawings of natural objects. However with practice the effect of orientation reduced considerably, presumably because unusual orientations were being learned. Another result indicates that novel depictions of a known class of objects may be identified by a process of mental rotation. Corballis (1988) has suggested that an initial frame independent description of an object's shape is sufficient to access a stored representation of that shape in long term memory. Once it has been categorised, its axis and orientation can be determined and mental rotation then used to determine other features such as its handedness.

Thus the evidence suggests that different types of matching and representation may be used for different kinds of object recognition. Neuropsychology provides clear evidence for the distinction between object recognition based on a direct match with long term memory and object matching which uses transformation processes by mental imagery. Using previously published cases Farah (1984) collected eight cases of patients with brain damage who could perceive and recognise objects whilst having a deficit in image generation. They could not describe or draw objects from long term memory and denied having visual imagery. A further 13 patients showed the reverse dissociation. They could detect draw and describe objects but they could not recognise objects presented.

There is evidence that the long term representations used for visual recognition include viewer-centred representations despite the apparent disadvantages. Thus Jolicoeur and Kosslyn (1983) provide converging evidence that both viewer-centred and object-centred representations are used for learning and recognising novel three-d. stick figures. Rock and DiVita (1987) found that after viewing novel three-d. wire objects their subjects showed substantial view dependence in their subsequent ability to recognise these objects. This was according to the authors the result of a failure to achieve an object-centred representation.
If as the evidence suggests, object recognition frequently depends upon a match with stored view dependent representations, then a key question is how are objects recognised from so many different viewpoints? Arnheim's theory that pictures exploit characteristic views or "renvois" has already been mentioned (4.2.3). Support for the idea that certain views contain more information than others and may act as recognition prototypes is provided by Palmer, Rosch and Chase (1981). These authors used three converging measures to rate what they called the "Canonicalness" of small models of familiar objects. These were (1) subjective ratings of 'goodness' of different perspective views, (2) the perspective from which people first imagine a common object and (3) the perspective from which they choose to take the best photograph. They then demonstrated that highly Canonical views (measured as above) are identified more quickly than others with response latencies being a monotonically decreasing function of canonicalness. Finally, the same authors showed that the rated canonicalness of a perspective view could be predicted from (1) the visibility of information about the object in a given perspective view and (2) the subjective importance of that information for the identity of the object. Thus their results support the idea that there are certain perspective views which make more information for perceiving and recognising an object than others. If only a few prototypical views of an object are represented in memory then it is presumed that objects seen from other views would have to be mentally rotated until they matched one of the stored views. The alternative theory is that so many different views of every object are represented in memory that little transformation of orientation need precede the matching process. Tarr and Pinker (1989) have attempted to distinguish between these two options. They examined the effects of practice on learning to recognise a number of novel stick figures at different orientations. Their results were consistent with a hybrid between the mental rotation and the multiple view point hypotheses. They suggest that viewed shapes are transformed to match the nearest or the most canonical of multiple stored views. The authors also suggest that for very familiar shapes mental rotation may be much faster or unnecessary because many views have been learned, as is perhaps the case with disoriented letters and numbers (Corballis et Al. 1978, Corballis 1988). Evidence is also provided from mental imagery. There is neuropsychological evidence that the long term representations used for recognition are the same as those used for generating a mental image (an important finding which will be referred to again). Thus Martha Farah described four cases of brain lesions resulting in content specific impairments in the ability to recognize objects (Farah, 1988b). In all cases there was a parallel matching content specific impairment in imagery ability. Pinker et Al. (1984) found that subjects could not
visualise in one step the appearance of a three dimensional object from an arbitrary viewing angle. Instead they visualise it in some "canonical" orientation then mentally rotate it into the target orientation. These two results combined, provide further support for the existence of stored view dependant representations. Lastly it should be mentioned that Perrett et Al. (1985) have found a range of separate clusters of neurones in infero-temporal cortex of the macaque monkey which are sensitive to specific views of heads and faces, profiles, full views, half views and even rear views. Perrett et Al. (1991, 1994) have shown that although there are cells tuned to a range of different perspective views certain views such as the full face and full profile were much more strongly represented than others. Interestingly the neurophysiological evidence for the importance of particular head views parallels the importance of different views in behavioural studies. Thus there is converging evidence from different sources that Arnheim's intuition that artists and designers exploit "renvois" to assist pictorial recognition is well based. This does not however rule out the co-existence of object centred descriptive systems though the nature of these is still disputed. Besides the view specific cells in monkey inferotemporal cortex, there are also some cells which are view point invariant to individual faces. Perrett et Al. (1994) suggest that such view independent cells are formed by combining the outputs of view-specific cells. They also suggest that the coexistence of different types of stored representation reflects the requirements of different types of visual task. Thus when the viewer must reach out to grasp an object, a viewer centred frame of reference which specifies orientation is appropriate. In other tasks categorical recognition which is independent of view point is more important. In most sketches the contour of objects and the spatial shapes made by specific views are important and I would expect painter's sketches especially to depend upon view-specific mental representations of objects and scenes. Sketches of objects for three-dimensional design however would be more likely to prime stored object centred representations of the Marr\Biederman type. readers are invited to survey the sketches illustrated to decide for themselves the truth of the hypothesis that rough sketches exploit prototypical views of objects or object parts to stimulate the imagination.(e.g. Illustrations IX, X, XI, XIV, XX, XXIX, XXXIV).

4.4 Does Sketch Recognition include a Depictive Match?

"Template" theories of object recognition postulate the spatial alignment of a perceived object shape with a stored candidate shape, followed by the computation of the degree of overlap between the two. Unlike structural description theories, recognition is assumed to include a matching process which depends upon correspondences of spatial position between the parts of two
representations. Such theories imply partly depictive representations which can be mapped into a spatial medium such as a neural array. Template theories of recognition have usually been considered implausible on the grounds of inefficiency (Neisser, 1967 Pinker 1984). However criticisms of alignment matching theories have often been "tilting at a straw man". A template match at an early stage of visual processing, in which an image is represented as an array of luminances and colours, would require such an impossibly large number of stored shapes to account for the recognition of multiple views and exemplars of object categories, as to be unworkable. Such a process has never been seriously proposed. Clearly alignment matches would only be applied to higher level spatial maps in which many candidate feature such as contours and figures have already been extracted. As many authors have pointed out it is difficult to see how such a matching process could account for the ability to recognise previously unseen examples of object categories such as trees, with infinitely variable projected shapes. Also as Pinker (1984) has pointed out it has problems with three-dimensional shapes and with an unstructured part of an image containing more than one object. However template theories are more plausible when they are assumed to apply to pre-processed representations, in which candidate objects have already been segmented into parts and normalised for position, size and orientation. The evidence presented above that multiple view dependent representations can be learned and used for recognition, combined with the evidence that transformation operations are sometimes used in shape recognition, makes a role for template matching more likely.

Ullman (1989) has described a model for shape recognition by alignment using semi-depictive "pictorial descriptions". He has shown mathematically that spatial alignment of two-dimensional views is possible provided three corresponding alignment points can be identified. He suggests that these might be prominent, easily identified features which most objects possess. In this approach the recognition process is divided into two stages. The first determines the transformation in space that is necessary to bring the viewed object into alignment with possible object models. The second determines the model which best matches the viewed object. Like Biederman's contrasting model Ullman's has been implemented by computer (Huttenlocher & Ullman 1987). Thus it can claim to increase the feasibility of template matching without, of course, demonstrating that such a method is actually used by the brain. One point that Ullman's model makes clear, is that alignment matching can match spatial positions in a map-like representations where spatial locations are used not to represent unprocessed primitives such as luminance and chromaticity (as in an real image), but labels to
pre-processed part descriptions. This is one of the advantages of combining depictive maps in a spatial array with descriptive parts. It is possible to imagine such an intermediate representation to be hierarchically structured, with spatial locations indexing both categorical structures and coordinate descriptions of parts, which could themselves be expanded by further mapping onto the array to provide extra detail as needed (my interpretation of an earlier suggestion by Palmer (1978)). The perceptual components of sparse sketches could then be acting as a spatial memory "look-up" system near the "trunk" of such a depictive-descriptive tree. Thus lines identified as trees in Millet's "Curtain of Trees" (Illustration XXIV) provide spatial pointers to descriptions in memory of missing branches and twigs. The snowman's face with stones for eyes and a stick for a mouth looks like a face because it provides spatial markers to detailed mental descriptions of familiar facial features (see also "fruit faces" in Palmer, 1975b). (In Chapter 9 I suggest that it would be rewarding to include such intermediate information structures in an improved machine sketching system).

Ullman does not mention another component of visual processing which might be used to facilitate recognition matching by spatial alignment. The visual system uses, in parallel, nine or more band-pass spatial frequency channels in the early stages of image analysis. Their physiological responses have been studied in area V1 of the visual cortex. Here organised columns of neurons responsible for processing a limited region of the visual field are divided into layers which respond to different spatial frequencies and which have in consequence overlapping receptive fields of different sizes (Georgeson 1979, De Valois et Al. 1982). If such spatially organised frequency channels are also available to those other cortical regions responsible for high level vision then template matching might be facilitated by matching corresponding positions in sequence starting with low and proceeding to higher frequencies. Such a process would allow varying degrees of tolerance in an alignment match and an early rejection of bad matches. Watt (1988) has suggested a similar sequence of spatial frequency matches for position finding and grouping processes in low level vision. The existence of alignment matching is implicit in the hybrid-image theory of the sketch discussed in 5.3.7.

4.4.1 Categories versus Exemplars.

As mentioned earlier, recognising sketch components may involve the assignment of a candidate object to a category and also the identification and elicited retrieval from memory of a specific example of that category. There are grounds for supposing that the visual system uses different types of representational system for these two kinds of recognition For example patients with a cerebral
deficit known as **prosopagnosia** can often recognise a familiar person's face as a face but are unable to identify whose face it is.

**Prototype theories** suggest that we use idealised patterns based upon an average structure of many exemplars to recognise categories. Experiments have shown that for recognising families of abstract patterns (Franks and Bransford 1971) or diagrammatic faces (Reed, 1972) people find it easier to recognise a variant of the family which is closer to the prototype than a less typical example even though they have never been shown the prototype.

Relevant to sketches is a theory of face recognition which suggests that, once a face encoding has achieved a match with a stored prototypical face average, an individual's face is identified by how far it varies from this prototype in a multidimensional space of facial parameters (Valentine & Bruce 1986, Bruce, 1988). Brennan (1985) has used a similar model to explore the nature of caricatures in a computer model. She has encoded a number of faces from frontal photographs as 186 corresponding co-ordinates scaled so that they have matching x, y values for the pupils of each eye. Her "caricature generator" works by exaggerating the differences between encoded points in an individual face and in a prototype formed by calculating the average co-ordinates of all the faces. It has been shown that caricatures generated in this way as line drawings are recognised more quickly than veridical line drawings of the same faces (Rhodes, Brennan & Carey, 1988). Although such "super-recognition" attributes are most obvious in political cartoons they are used in more subtle ways in artists' sketches to emphasise important features. Note for example the exaggerated body length and necklines in Fashion sketches (Illustration VIII).

Theories of object category representation are important to sketch function because in the early stages of visual thought a "natural" perceptual device is needed to assist the mind to imagine many possible exemplars of a tacitly represented object class.

### 4.4.2 Schemata: The Importance of Context.

The deliberate vagueness of typical sketch attributes means that their context must play an even greater role in biasing recognition than it does in perceiving less ambiguous stimuli. The speed of sketch recognition and the mental ability to resolve sketch ambiguities are dependent upon a prior knowledge of the sketch subject matter. The term "**schemata**" is frequently used to refer to stored models of scenes and events, which can be used to anticipate the spatial positions and semantic relationships between candidate objects and their parts (Bartlett 1932, Palmer 1975a, Minsky 1975).
4. Sketch Perception and Recognition

In a series of studies, Biederman (1972, 1981 Biederman et al. 1973) has shown that the speed and error rate of identifying and finding familiar objects in pictures, is greatly affected by their pictorial context. When pictures of familiar scenes were jumbled, or an object was placed in an unlikely position, detection speed was greatly increased compared to when the scene was coherent. A rather startling finding was that selected objects in coherent scenes could be recognised in too short a time to allow more than one saccadic glance. Thus scene schemata appear to be activated with very briefly presented information. Biederman lists five types of object relationships that might be recorded in scene schemata - support, interposition, probability, position and size. (See Illustration XXX).

In a complementary study, Palmer (1975a) has demonstrated the influence of the prior presentation of visual scenes on the identification of briefly presented drawings of familiar objects. The number of correct responses and of confusions with visually similar objects were related to both the context and the particular target object presented. Palmer quotes Morton's (1969) "logogen" model of word recognition as providing a good "quantitative fit" to his data. One of Morton's suggestions, which has since been adapted to visual imagery (Paivio's "imagen" theory, Paivio 1986), is that internal recognition units exist which have a threshold for activation which is lowered by semantic contexts with which they are consistent. Applied by analogy to sketches rather than words, ambiguous object contour lines in the Leonardo Anghiari sketch discussed above (Illustration II) are more likely to activate "Neptune rising" recognition units or "men fighting" units depending upon the sketch schema unit currently co-active in Leonardo's mind.

Palmer (1975c) sees scene schemata as a natural development of Structural Descriptions with characteristic objects linked by propositions expressing their relationships. The difference between a scene schema and an object structural description is one of degree, with a scene schema, of necessity, being more tolerant and general. When a sketch centres around a single object then a schema for an object category might be used to limit the number of more specific structural descriptions.

Whether such scene schemata are as influential in untidy sketches as they are in the neat drawings used by Biederman and Palmer deserves to be tested. Illustration XXV shows Rousseau's "Twisted Tree" with a part misplaced and Illustration XXVI shows the same drawing corrected.

4.4.3 Memory Retrieval

When applied to sparse or untidy sketches, the theories of object recognition described above obviously depend upon poorly understood memory
retrieval mechanisms. Sketch attributes, however impoverished, act as cues to the retrieval of stored object information. There is a huge parallel literature on "memory retrieval" (reviewed Baddeley, 1990, chapter 11), which has been largely avoided here, because it is almost entirely concerned with accessing the meanings and associations of words. However there is a fascinating analogy between the theory of recognition by parts, combined with a later image completion process as presented here, and Tulving's theory of "Synergistic Ecphory" (Tulving, 1983). (According to Baddeley (1990) the term "ecphory" is based on a Greek word meaning "to be known". Tulving himself does not give the origins of his neologism). It is used by Tulving to refer to a theory of episodic memory retrieval (recalling memories of events) in which a retrieval cue is combined with the unconsciously stored representation of an event to provide a hybrid "conscious memory of certain aspects of the original event". Learning is assumed to implement the storage in Long Term Memory of a "Memory trace" or "engram". Recollection is postulated (in Tulving's model GAPs, short for General Abstract Processing System) as a symbiotic process of interaction and constructive synthesis ("Synergy") between the retrieval cue and the stored engram. Tulving's theory of episodic memory ecphory is relevant to Chapter 7 which discusses how artists may consciously use and recall immediately past sequences of their private cognitive acts. Here an interesting and possibly significant parallel between theories of visuospatial memory retrieval and event recollection is simply noted. (see also chapters 6 & 7).

4.4.4 Towards a Model of Sketch Recognition.

It is possible to conceive that the various and apparently contradictory representational systems proposed for object recognition are elicited collaboratively at different stages of the recognition process and for different purposes. Thus

- Sketch recognition starts by the automatic activation of a schema elicited by the sketch title (if it has one) and knowledge of the subject matter. The schema biases the selection of candidate object types and object parts and their anticipated spatial relationships.

- Incomplete object components and candidate features, mainly lines seen as fragments of contour are used to elicit, to construct and to match stored Structural descriptions of object categories.

- Object prototypes are used to elicit more precise spatial relationships for object type matches (e.g. a face, a chair etc). It is possible that at this stage the representations are low resolution coordinate systems. For discussion and evidence concerning the distinction between categorical spatial relations (left hemisphere)
and coordinate spatial relations (right hemisphere) see (Kosslyn et Al. 1989, 1992, Kosslyn and Shin, 1994). Matching here might be accomplished by comparing the nodes of descriptive trees attached to recently processed and to permanently stored object attributes.

- Object exemplars are recognised or searched for. At this stage matches will be tested between contour components and prototypical object views (renvois Arnheim) as suggested by Palmer et Al. (1981) and Tarr and Pinker (1989). The matching process proposed here is more complex, using an intermediate depictive-descriptive type - the descriptively labelled spatial map. This may require preliminary orientation and size transformations to match the spatial components. There is now evidence (perhaps inconclusive) that stored structural prototypes are used for recognising categories (left hemisphere) and coordinate systems for object exemplars (right hemisphere) and that they are represented separately in different parts of the brain (Kosslyn & Shin 1994).

Where such a matching process fails, either because an atypical view is not represented in memory, or because certain attributes of interest (such as handedness) are not represented, then further mental transformations of size, orientation and part permutation may be necessary to facilitate a match with memory.(Kubovy & Podgorny 1981). It is not clear whether such transformations (if they occur) use the same resources and processes used for voluntary mental imagery. I suggest they do. The fact that such posited transformations occur much faster than when imagery is used to match two simultaneously presented shapes (Jolicoeur 1985, Corballis 1988) may perhaps be explained by different degrees of automaticity used in the two tasks. Sequential matching between a remembered and a displayed shape is faster than simultaneous matching of two displayed shapes.

The evidence is consistent with the view that confused and untidy sketch attributes stimulate creativity because they elicit the automatic recognition matching processes in a set of nested search loops, which activates many more components of long term visual memory than would either a complete and unambiguous image or a very sparse one.

Although such a "top-down" sequence, from schema to exemplar to specific view, seems the most likely direction for sketch recognition, the reverse "bottom-up" sequence can play an important role in invention. Ambiguous or incompatible possible object contours can cause the search for a new sketch schema. It is posited that, during the act of sketching, unconscious processes which evolved for recognising real world objects are being pressed into service by problematic stimuli to support invention. Incomplete or ambiguous sketch parts, whether they are coded as Hoffman, D. "part boundaries" or Biederman's "geons" will often prime
4. Sketch Perception and Recognition

a prototype or object cluster with many exemplars in memory. Once the recognition system has "zeroed in" on the appropriate region of long term store I suggest that a phase of mental image generation may be facilitated with repeated automatic and sometimes unsuccessful attempts to resolve ambiguities by alignment matching (Illustrations II, III, XXII).

4.5 Discussion: Three ways of Going Beyond the Information Given.

Sketch attributes exploit for creative purposes cognitive routines which evolved in our hominid ancestors for completing impoverished real world stimuli. It is argued by at least one school of thought that the perception of normal real world objects must also constantly interpolate or "fill in the gaps" left by inadequate visual information. Rival theories about how such cognitive interpolation is implemented are linked to the question of at what stage or stages in the visual process stored knowledge and "inference like" processes are used to construct a percept. For this reason visual perception and recognition are advisedly discussed together, as inter-dependent aspects of a common process. To a "neo-Helmholtzian", perceptual completion is part of a single continuous inference-like process comparable (apart from its automaticity) to conscious thought. The phrase "beyond the information given" in the section title is a quotation from Bruner to whom perception is an act of categorisation (Bruner 1957) and is implicitly continuous with cognition.

Widely discussed examples relevant to the sketch, include the so called "cognitive contours", better referred to as an "illusory surfaces", investigated in detail by Kanizsa (Kanizsa 1976, 1979, Kanizsa et Al. 1982). Kanizsa has argued persuasively that such interpolation processes are not easily explained by an inferential or recognition process (Kanizsa 1979). Like the gestalt Laws of Perceptual Organisation "amodal" completion is part of the pre-processing which provides the "data" given by vision to further processes which include recognising and also perhaps unconscious reasoning. Kanizsa argues that there are two ways of "going beyond the information given" in perception. The first, the "primary process" extracts contours, surfaces and candidate objects and interpolates (without reasoning) object parts as if they were partly occluded. The "secondary process" which is another way of "going beyond the information given" includes recognition and reasoning upon the products of the primary process. We experience the missing contours in Matisse's nude (Illustration X) clearly but in a less vivid way and without the contrast enhancement of the Kanizsa cognitive contours. Thus it is imagined contours rather than Kanizsa's "subjective" contours which are the more common device in sketches for stimulating inventive shape perception. Although these two processes may be concurrent and interdependent, Kanizsa argues that the
4. Sketch Perception and Recognition

distinction between "seeing" and "thinking" needs to be preserved. In sketching the interdependence between two forms of perceptual completion is clear, since in sketching the artist is repeatedly "thinking" descriptively about a spatially imagined object and then imagining novel spatial and chromatic consequences of such newly derived descriptions, in a recursive loop (Chapter 3).

Although the Kanizsa distinction between completion by "seeing" and by "thinking" is useful, it does not go far enough. I argue here that, in order to account for the perception of impoverished stimuli found in sketches, we must distinguish at least three ways of "going beyond the information given". This arises from the distinction between depictive and descriptive modes of representation which may be implicit in Kanizsa's writing but is never clarified.

4.5.1 Pre-Conscious Visual Interpolation.

A set of Primary processes organise and structure the visual input in ways which includes the partial and sometimes faulty completion of candidate objects and surfaces. In pictures especially the end results of these processes are often available to consciousness in a depictive form, even if no familiar objects are recognised and even if such interpolations conflict with the inferences from recognition (Kanizsa and Gerbino 1982) These processes also result in the construction of unconscious descriptive (symbolic) representations of the image (as proposed by Marr). This stage, sometimes called "low level vision" uses the innate processes which have been acquired by evolution in adaptation to the properties of three-dimensional space, with the constraints of gravity (giving selective preferences for vertical and horizontal axes) and the ground plane.

A crucial processing requirement, possibly occurring in cascade with phase two, is the distinction of spatial perception from the recognition and categorisation of objects. The layout in egocentric space of the positions, sizes and orientations of candidate objects is needed to comprehend movement, to guide reaching and grasping and for visual search. In the sketch such spatial perception is used, I suggest, to imagine such movements. As discussed above the detailed recognition of objects is complementary to this. It demands the perception of those attributes which are independent of such spatial metrics. The distinction between spatial perception and object recognition is reflected in the dorsal and ventral visual processing routes in the neocortex (see Chapter 5).

4.5.2 Descriptive Enrichment by Knowledge and Inference.

A set of secondary processes access long term memory, resulting in various degrees and kinds of recognition (Section 4.3). This provides a much richer representation of the sketch components by linking objects parts and features to
stored semantic knowledge of scenes and objects in descriptive form. Although this process will have greatly enlarged the descriptive components of the percept it cannot by itself account for the spatially depictive enrichment of the percept which occurs after recognition (e.g when we recognise the face in the well known "Hidden man" demonstration (Illustration XXXII, Porter 1954)

4.5.3 Spatial Superimposition of Generated Imagery.

In order to synthesise a spatially complete percept, after recognition and descriptive inferential processes have occurred, I suggest a third type of enrichment of the perceptual representation occurs. This is distinguished from both Kanizsa's "primary interpolation" and from the "secondary" processes of memory retrieval for descriptive labelling and "unconscious inference". To assist spatially oriented responses, depictive images are generated from stored memory and spatially superimposed upon the final percept. Evidence that imagery is used in this way, to assist the perception of partly occluded objects, is presented by Cave & Kosslyn (1989). How such a third way of "going beyond the information given" might have become adapted to sketching is the subject of the next chapter.

These three stages of "going beyond the information given" can be related to the depictive descriptive continuum, discussed in the Chapter 3, and considered as perceptual equivalents of the representational translation processes I claimed were important to sketching. In the first stage unstructured depictive retinal information is structured into tentative objects, without reference to memory. Ambiguities such as those caused by occlusion, may be resolved by innate "wired in knowledge" of the physical environment and of visual probabilities acquired by Darwinian natural selection. At this stage the processing amounts to depictive to descriptive translation, which in turn can have descriptive to depictive consequences (Kanizsa contours, perceptual groups, figure and ground etc). The second stage is depictive to descriptive translation by access to schemata and structural descriptions in memory. The third stage, in which I suggest mental imagery may be used, is a translation of descriptive information acquired by recognition which is then used to complete the spatial structure of the perceptual construction of conscious awareness. An implication of this late construction theory, which will be referred to in the next chapter, is that conscious percepts must be represented in a temporary storage medium. There is evidence (see Chapter 5) that incoming perceptual information has access to long term memory before it is held in working memory (Logie, 1995). If, as I shall argue, conscious percepts, are dependent upon and perhaps coincident with representation in such a short term memory then theories of visuospatial working memory and of image
4. Sketch Perception and Recognition

generation are integral components of "late construction but early memory retrieval" theories of perception. The perceptual representation of the sketch is, I suggest, a temporarily stored representation of the sparse but richly encoded marks on paper, combined with a fluctuating hierarchy of depictive imagery, spatially labelled with descriptive structures retrieved from long term memory. The former provides a skeletal context and support for the latter.

The evidence discussed in this chapter supports a second hypothesis about sketch function

4.6 Conclusion: Hypothesis Two.

Human perception and object recognition processes enable accidentally impoverished stimuli fast automatic access to long term memory for perceptual completion. Sketch attributes exploit such unconscious processes providing memory search and retrieval cues that improve the availability of tacit visual knowledge for invention.
5. HOW SKETCHES COMBINE PERCEPTION WITH IMAGERY.
"I prefer stories told on the Radio to those on Television because the pictures are better"
Small girl interviewed for the B.B.C. circa 1976

5.1 Introduction

Hypothesis Three states that "Sketches are percept-memory hybrids. The incomplete, physical attributes of sketches act as a stimulus for percepts which invite completion from memory and imagination. This results in the generation of transient mental images which after a matching process are transformed and spatially superimposed upon an internal representation of the sketch."

It is argued that the processes used by the brain to create and transform mental images have had part of their evolutionary origins in the need by our hominid ancestors to complete from memory the perception of naturally occurring stimuli, impoverished by poor illumination, brief duration or partial concealment. Sketch attributes are viewed here as an adaptation, by cultural evolution, of such innate mechanisms for an inverse function. Artificially impoverished stimuli are deliberately or unconsciously created to elicit imagery which assists the planning and invention of imaginary or future objects and events. Due to its origins, inventive mental imagery needs perceptual support.

In this chapter the evidence that sketches are percept-memory hybrid images which support the generation and mental manipulation of mental images for invention is examined.

It will be argued that sketches support mental imagery in two ways. One of these is appropriate to the later stages of composition, when imagined parts are used to experiment with those undecided details necessary to refine the more certain outline of an idea. The other, in which the imagery is less constrained, is more applicable to the early inventive stages of visual design (see Chapter 2). In both functions it is posited that a variable region of this combined map of spatial location is available to focal attention. To this is ascribed the function of connecting location to stored component descriptions used for generating new depictive image parts. In turn the modified hybrid map may be used for encoding and re-encoding new structural descriptions and so completing an iterative processing loop. Thus, skilfully used, ambiguous or incomplete sketch attributes assist creativity by stimulating the visual system to generate a stream of spatially super-imposed mental imagery. This proposal is supported by evidence about the nature of mental imagery and the way that it interacts with visual perception.
5.2 Mental Imagery and Sketch function.

5.2.1 Imagery as a Medium for Visual Experiment.

In Chapter 1 preliminary reasons were given for believing that artists use mental imagery as a method of experimenting with visual appearance. In the last twenty years there has been extensive research and debate about the nature of mental imagery. Here I will briefly review the evidence for those imagery properties most relevant to the visualising functions of the sketch. For reviews see (Kosslyn 1980, 1981, 1983, 1987, 1994, Finke 1980, 1985, 1986, 1989, Pinker 1984a).

By definition mental images are experiences which introspectively resemble percepts but are generated from memory rather than information from the eyes. Historically mental images have been compared to faint percepts resulting from the memory of past visual experience (Kosslyn 1980 chapter 11). However unlike the perception of objects (or pictures of objects) physically present mental image parts can be generated, transformed, manipulated, and combined in new arrangements at speed. These properties are exactly what a visualising medium for art or design requires. The similarity to perception is important because visual invention is concerned with predicting and manipulating the appearance of hypothetical objects from particular view-points. The need to manipulate the structure of such appearances at speed without the penalties of having to use a physical medium to do so is a consequence of the astronomically large number of permutational possibilities which visual experiment implies (Chapter 1). As Kosslyn remarks "Because of the kinds of transformation we can bring to bear and the kind of representation an image is, we can use our imagery as a simulation of possible (and perhaps of impossible) transformations in the world. Thus imagery as an aid to thinking about the consequences of given actions, is a crutch to help us devise a plan for reaching some desired state of affairs." He goes on to remark that "in addition to serving some role in concept learning, reasoning and pattern recognition imagery may well serve to make unconscious thoughts and desires manifest in consciousness, as Freud and others have maintained." (Kosslyn, 1980 page 456). A point to which I will return later (Chapter 7).

Unfortunately a third class of imagery properties greatly reduces many of the advantages of the first two. Besides being functionally similar to percepts and flexibly transformable in conscious memory, visual images are often vague, difficult to form and hard to maintain, fading quickly and needing constant refreshment (Kosslyn et Al. 1983, Kosslyn and Shin 1994). Moreover the vividness with which subjects are able to form and inspect their images varies
widely (Galton 1883, Marks 1972). For these reasons mental imagery requires perceptual support.

5.2.2 The "Spatial Array" Theory of Imagery.

Historically a wide range of theories have been used to explain the functions and mental representation of imagery (for critical review see Pinker & Kosslyn 1983). However the most detailed and best supported by evidence is Kosslyn's "array" theory (Kosslyn 1980). This has been partly implemented as a computer model (Kosslyn and Shwartz 1977) and developed by Pinker (1980). Because the array theory has the greatest explanatory adequacy for sketch function it is the one used as a framework here. This does not imply however, that all details of the theory are to be accepted without reservation. The theory which has been developed at the level of representation and process for "high level parsing" of visual cognition distinguishes the following components of imagery :-

1. **Visual representations** stored in long term memory are used as a source of information for imagery generation. In the Kosslyn and Schwarz model these are assumed to be propositional representations and also coordinate data used to preserve spatial structure. However there is evidence (Chapter 4 & Farah 1984) that the representations used for mental image generation are the same as those used for recognition. It follows that all the types of representational distinction postulated to explain recognition (Chapter 4) are equally relevant to mental image generation (see Kosslyn et Al. 1992). Thus Kosslyn & Koenig (1992) distinguish two kinds of representation of spatial structure, *categorical* (left hemisphere) and *coordinate* (right hemisphere). They also distinguish between *prototypical* representations of object categories from *exemplar* representations. By permutating the different possible long term memory sources, Kosslyn & Shin (1994) deduce that 24 different types of mental image are theoretically possible.

2. The **"Visual Buffer"** is posited to be a short-term array format medium of limited capacity in which images can be generated and transformed. For several years it was not clear how Kosslyn's "visual buffer" was related to the "visuospatial sketch pad" posited in the Baddeley and Hitch model of working memory (Baddeley 1986, Baddeley & Hitch 1974, Baddeley et Al. 1984). Recently however the two independent lines of research have converged and work by Logie et Al. (Logie 1986, Logie & Marchetti 1991, Logie 1995) throws new light on the role of working memory in mental imagery.

3. A **Generation Process** is also posited, which can retrieve visual information from unconscious long term memory and create, sequentially, component by component, percept-like, semi-depictive representations which are
mapped onto the visual buffer. The generation process is believed to convert unconsciously stored long term representations - "structural descriptions", object centred prototypes and exemplars represented as co-ordinates etc - to a view dependant spatial representation. It is thus a perfect example of the descriptive to depictive translation processes discussed in Chapter 3. Kosslyn posits a Refresh process, linked to image generation, which is necessary to regenerate sketch components which fade in the sequence in which they were generated. In the original model (Kosslyn 1980) image refreshment was supposed to occur within the buffer by reactivation of image parts as they are scanned or inspected. Kosslyn and Shin (1994) admit that "refresh" is a poorly understood component of the model and that it is uncertain whether it involves regeneration from long term memory or in situ reactivation of fading parts.

4. A set of Inspection Processes which are able to extract, organise and relate information which is only implicit in the spatial array, making it explicit and communicating it to other cognitive systems. The visual buffer is assumed (as already discussed) to be analogous to a labelled map with descriptive information about objects linked to spatial locations. A limited region of the visual buffer is assumed to be available in an "Attention window" for conscious access and detailed processing. Image inspection requires a spatial scanning operation similar to the movements of covert attention occurring in perception (Posner 1980 & Chapter 6). Also required is a mechanism for recording scanning distances and directions and the capacity to sequentially access and record descriptive information stored at specific locations. "Inspect" is clearly an attempt to model the introspective experience of the "mind’s eye" when we use mental imagery to answer questions. Paivio (1971b) gives the illustrative example of answering the question "How many corners does a block uppercase 'F' have?". Most people report that they have to mentally scan a mental image of the letter counting the corners one by one as they do so. Kosslyn believes that he has dispelled any "Infinitely recursive homunculus" problems raised by introspective "inspect" experiences by simulating image inspection in his model by machine pattern searching algorithms on a computer array. However, personally, I am not satisfied that such searching processes, however sophisticated, are an adequate description of the mental inspection experience. Missing, is an adequate process account of two crucial components of the mind's eye" - volition and conscious awareness. I discuss this difficult unsolved problem of psychology in Chapter 7.

5. A set of Transforming Processes are used to manipulate the contents of the visual buffer. The best known and widely researched of these is mental rotation. In the seminal studies of Shepard and Metzler (1971) and Shepard and
Cooper, L. (1982) the time to match or mismatch differently oriented shapes is linearly proportional to the angular difference in their orientation. The rotation of three dimensional objects (drawn in perspective), implied by the data from these and other workers, occurs as fast when it is in depth as when it occurs within the plane of the projected view. This implies that subjects can rotate stored three-dimensional models created from two-dimensional views. Other forms of transformation include changes of size (Bundesen et Al. 1975, 1981) and the synthesis of new shapes and patterns, by translating, rotating and combining separately imagined parts (Finke & Slayton, 1988, Finke, 1990).

This multi-component model is supported by a wide range of experimental performance data, the interpretation of which is to my mind persuasive (Kosslyn 1980, 1983, 1987, 1994, Finke 1989) but to some authors controversial (Pylyshyn 1975, 1981, Hinton, 1979). More recently Kosslyn has up-dated his model to make it consistent with new research in neuroscience and neuropsychology (Kosslyn, 1994). Whilst the early model was based upon a computer analogue, the later model is mapped more closely to visual pathways in the brain.

The model (in outline) is also supported by an analysis of individual differences in imagery tasks designed to separate its different components (Kosslyn et Al. 1984). Further, a careful study of selective imagery deficits resulting from brain damage, provides evidence for separate neurological resources for at least three of the postulated imagery components (Farah, 1984, 1988b). Thus there is evidence from patients with cortical lesions that image generation is mediated by the left posterior hemisphere (Farah 1985a), whilst movements of spatial attention in perception and imagery is dependant upon a visual region in the right parietal cortex (Bisiach & Luzzatti 1978). There is also evidence of selective damage to long term image representation and to the ability to "inspect" mental images.

5.2.3 Similarities of Structure between Images and Percepts.

Finke has researched and repeatedly reviewed the ways in which images are functionally "equivalent" to percepts (Finke 1980, 1985, 1986, 1989). Briefly he has argued that such similarities occur only at what are believed to be the "higher level" visual processes or those occurring in the later central stages of visual processing. Thus a finding, quite important to the arts, is the failure to duplicate colour contrast effects in imagery (Finke 1986), whereas other types of visual illusion thought to have a less peripheral cause have been duplicated in imagery (Wallace, B. 1984).

Studies by Finke and Kosslyn (1980) and by Finke and Kurtzman (1981a, 1981b) seem to show that the size and shape of the field of view and the way in
which acuity and spatial frequency discrimination vary within this field are equivalent in perception and imagery. In the former experiments, the ability to discriminate a pair of visual dots against a screen as they became progressively further from the fixation point was compared with a similar pair of imagined dots seen against the same screen. In the 1981 experiments a similar procedure was used to map the imagery field using horizontal and vertical gratings of different spatial frequencies.

These results are discussed separately by Kosslyn (1980, 1983) and Finke (1989) as evidence for the equivalence of the spatial medium of perception and imagery. However some qualifying points are necessary.

1. The experiments purport to measure the size, shape and resolution of a single point of the visual field with gaze fixed in a single direction. Thus they suggest an internal retinotopic map, with highest resolution in the foveal direction, which is shared by visual and by imagined information. This cannot be the whole of Kosslyn's visual buffer, however, since one can perceive and imagine a wider field by integrating successive saccadic glances (Hochberg, 1982, 1984, Jonides et al. 1982). It is useful here to distinguish the visual field with fixed gaze from the field with eye movements only, and the even wider "head field" with both eye and head movements as in normal vision (Sanders, 1963).

2. The experiments with both imagined dots and gratings were performed with eyes open, staring at a visual screen and fixation marker. Thus it can be argued that they measure the field of resolution of a hybrid visual and imagined image within a shared neurological medium.

3. Further experiments, reported by Kosslyn and Finke (Finke 1989), measured the visual angle at which imagined shapes are reported to overflow their imagined field of view. The results suggest that there is a roughly circular "window of attention" for mental imagery which is smaller than the total field of imagined view. It is further suggested by Kosslyn (1980, 1994) that this window of attention mirrors visual attention having a finite capacity in which its spatial size decreases with higher resolution.

There is evidence that the internal representations of percepts and mental images have a similar structure. This is a necessary but not sufficient requirement for the hypothesis that mental images are used to complete impoverished percepts. For example Shepard and Chipman (1970) compared peoples judgements of similarity of the perceived shapes of U.S. states with their judgements of similarity based on memory. A multidimensional statistical analysis revealed good correlation between the two with types of similarity clusters suggesting that the visual shapes compared in memory have similar visual features to those observed directly.
Similar results have been reported with familiar faces, numerical symbols and colours (reviewed Finke & Shepard, 1986). The latter case is particularly interesting since it has been shown that even subjects with colour vision deficiencies (protanopia and deuteranopia) construct the same standard colour circle based on imagined colour similarities that is produced by normally sighted subjects using visual judgements. Further studies showed that these results could not be easily explained in terms of an ordered search through descriptive statements.

Podgorny and Shepard (1978) compared the distribution of attention to letter forms in perception and imagery. In the perceptual condition the letter forms composed by filling squares of a grid were displayed with the grid and then immediately followed by a small probe spot. In the imagery condition subjects were instructed simply to imagine the letter forms projected on the grid. Subjects mean reaction times to decide whether the probe fell on or off the letter form varied in the same way across probe locations and types for the imagined and for the actually presented letters. The results are not easy to explain in terms of Pylyshyn's suggested "tacit knowledge" or experimenter bias (see Chapter 3). Podgorny and Shepard explained their results in terms of a sub-structure of horizontal and vertical bars possessed by both the perceived and the imagined forms. Shorter reaction times occur when the probe falls on or near more than one sub-component. I must point out that these experiments too, could be judged to be experiments with hybrid images in which the image components are stabilised by a constant perceived grid.

Further evidence that mental images have a similar substructure to percepts has been provided by Reed and Johnsen (1975). As the gestalt psychologists showed it is easier to perceive certain "good" substructures in an abstract pattern than others. For example it easier to perceive a Star of David as two large overlapping triangles or as an hexagon surrounded by six small triangles than it is as three overlapping parallelograms. Reed and Johnsen found that the parts which were easiest to detect in a mental image of such patterns were those detected in a corresponding percept or which matched an original structural description of the pattern before it was imagined.

5.2.4 Spatial Scanning of Images.

Some of the evidence that both percepts and mental images can be spatially scanned with an internally represented "window" of attention has been briefly discussed already in Chapter 3 as part of the evidence for depictive representations. The implication behind the experimental data is that percepts and mental images
are represented in a spatial array in view dependent format but with three dimensional depth also represented (as in Marr's "2½D sketch"). In the words of Finke (1989 page 61) "The spatial arrangement of the elements of a mental image corresponds to the way objects or their parts are arranged on actual or physical surfaces or in an actual physical space". Like maps images preserve metric spatial information (Kosslyn 1975, Kosslyn, Ball & Reiser 1979).

Pylyshyn (1973, 1975, 1981) in a series of papers has criticised the experimental basis for the spatial array theory of mental imagery. He claims that the results of experiments on the time to scan distances in remembered maps, for example, could be equally well explained by subjects using their "tacit knowledge" of scanning times for real maps. These criticisms have been addressed in experiments by Reed, Hock & Lockhead, (1983) who compared the times to scan different remembered patterns such as spirals with the times to scan the same patterns when actually displayed and with other subjects' estimates of the times to scan the same patterns. They found that while the relative scan times correlated well with the scan times for the same patterns when actually perceived, they differed in important ways with subjects' advance estimates as to what the relative scanning times would be. This weakens the "tacit knowledge" criticism.

In addition to the experiments already mentioned where subjects are told to scan an image, image scanning has been shown to occur spontaneously in order to extract explicit information represented only implicitly. For example Finke and Pinker (1982) investigated subjects ability to determine whether a small arrow pointing in an unexpected direction was pointing to one of a pattern of previously seen (but now absent) random dots. Even though imagery and image scanning was never mentioned the subjects response times increased linearly with increasing distances between the arrow and the dot.

These and other results are paralleled by evidence for covert (ie without eye movements) orienting of attention (Posner 1980, Tsal 1983) and the curve tracing results of Jolicoeur et Al. (1986) also referred to in Chapter 3.

An interesting finding in relation to chess perception has been described by Church and Church (1977). They found a linear relation between the time taken to decide whether the King was in check and the spatial distance between the relevant pieces on the board. Interestingly the implied scanning was faster for horizontal or vertical checks than for checks along the diagonal. This may have had something to do with the representation of the chess board which in this theory has a similar function to a sparse sketch, i.e. it is a percept which supports mental imagery.

Experiments on scanning provide evidence for the three-dimensional nature of the mental imagery medium. In a series of experiments by Steven Pinker (1980)
subjects learned the locations of five small toys suspended in an open box. With eyes closed subjects were then told to imagine a small black spot moving rapidly from object to object. Scanning times were proportional to the three-dimensional distance between the objects. However Pinker also showed that people could imagine the scanning of a two-dimensional perspective projection of the objects. In this case scanning times were proportional to the two-dimensional projected distances. In other experiments Pinker and Finke (1980) showed that subjects were able to correctly discover unexpected emerging two-dimensional perspective patterns between similar remembered objects by imagining the rotation of the space in which they were suspended.

It is clear that mental scanning is an important process for extracting and making explicit geometrical and topological relationships within an imagined object or scene. If, as the evidence suggests (Kosslyn 1980, Kosslyn et Al, 1983), mental images fade and must be refreshed, then a plausible function of the sketch is to provide stable percept landmarks to support the mental scanning of labile images. Mental scanning must be assumed to be performed frequently whilst artists or designers are sketching. For example in visual composition it frequently happens that one selects first the desired end points of a possible object contour. In order to decide whether such a contour line will end at the desired point an image must be generated and directionally scanned before the pencil is moved.

5.2.5 Image Inspection.

A problem which has traditionally been highlighted by critics of array theories of imagery (Pylyshyn, 1973 1984, Dennett 1981) is how we can "see" objects and their attributes in images without positing an infinitely recursive inner visual system. The answer provided by Kosslyn (1980, 1994) is that imagined objects are inspected by using the same mechanisms that are used to inspect an object during perception. During perception the visual buffer contains more information than can be processed at any one time. The literature contains much evidence showing that we can shift attention covertly (without eye movements) across space to selectively process limited regions of the visual field. There appears to be a "window" of attention which can be adjusted to take in limited regions of the visual buffer (Larsen and Bundesen, 1978, Treisman & Gelade, 1980). This attention window appears to exhibit a "scope/resolution trade-off" (Shulman & Wilson, 1987b, Humphreys and Bruce, 1989). Experiments by Navon (1977) suggest that we typically encode the global properties of an object before its details. Kosslyn further suggests that because of this trade-off, visual encoding of the overall shape of an object will often be of low resolution. Thus to "see"
How Sketches combine Perception with Imagery.

(visualise) a specific part clearly one will have to activate additional stored representations. Kosslyn assumes that image inspection involves the same separation between spatial and visual object processing that is shown in perception (discussed below). Thus a different sub-system may be used to image and attend to object details from that used to shift attention and scan mental images. This part of the Kosslyn model may be controversial, but it is broadly consistent with a model proposed by Logie (1995) for visuospatial working memory (discussed below) and with ideas suggested by Crick (1994).

There are other limits on image inspection which are important to sketch function. Chambers and Reisberg (1985) showed that a figure that is ambiguous when perceived, can lose its ambiguity in a mental image. Thus an outline drawing that can be seen alternately as either a rabbit or a duck, (Illustration XXXI b), is no longer ambiguous when imaged. The same was found for the "Necker " cube which does not show the usual spontaneous orientation reversal in imagery. The explanation usually offered for this (Brandimonte & Gerbino, Logie 1995), is that the perceived figure is automatically encoded descriptively labelled, so that only one of the possible descriptions can be used for image generation. The hypothesis outlined here suggests a slightly different explanation. The ambiguous drawing is an incomplete skeletal structure which elicits the superimposition of two kinds of mental image. It is the image component of the hybrid which determines how the drawing is recognised and it gates out any re-encoding in memory of the ambiguous parts of the drawing. Thus the percept changes by activating different mental images. When a mental image is generated without the percept, essential subtleties of the original drawing are not available.

This hybrid image explanation of the frequent non-ambiguity of mental images differs slightly from that offered by Reisberg and Chambers themselves. However their interpretation is consistent with the one above and has similar implications for sketch function. They report experiments (1991) to support their view that one can find targeted shapes in an imaged outline pattern only when the target shape is compatible both with the imaged geometry and with "how that geometry is organized and understood". This is compatible with the theory posited here that mental images consist of descriptively generated, spatial maps "looking for" impoverished stimuli. They have been generated from descriptive information in memory and are already structured by further links to descriptive hierarchies. Thus Reisberg and Chambers term "image understanding", which (they explain) includes perceived organisation such as orientation and figure-ground segregation, can be interpreted to refer to perceptually acquired descriptive structure. Images they suggest have an important role in thought by initiating a chain of further
How Sketches combine Perception with Imagery.

Images based on descriptive "resemblances". However they cannot be ambiguous (i.e re-interpreted) because unlike percepts they lack "the neutral stimulus with which perception begins". Here I wish to extend this argument in two ways,

1. If the primitive function of mental images is to complete impoverished stimuli then image restructuring in the absence of the appropriate stimulus match would be counter productive. Thus there may be a Darwinian reason why it easier to re-interpret outline drawings (eliciting new unintended images) than to re-interpret intended images. (But see Brandimonte & Gerbino, 1993 below).

2. The Structural re-interpretation of images is an important component of visual invention. Reisberg and Chambers (1991) provide useful (and overlooked ?) insight into the need by artists to support their mental imagery with neutral, ambiguous and unexpected stimuli.

5.2.6 The Transformation of Images and Mental Matching.

The classic study of mental rotation was conducted by Shepard and Metzler (1971). They presented their subjects with two perspective line drawings of three-dimensional objects which were at different orientations and were either identical in shape or mirror-images of each other. The mean time to determine whether or not the two shapes were identical was directly proportional to the amount of angular rotation which would be necessary to bring the two shapes into viewer-centred perspective congruence. This is true even if the imagined rotation would have to be in depth. The results are very robust and have been studied in great detail. Cooper and Shepard (in Shepard and Cooper, L., 1982) obtained similar results indicating mental rotation in matching rotated letters, numbers, random two-dimensional shapes and drawings of hands. To show the analog nature of the posited rotation the same authors performed experiments in which subjects were asked to start rotating an imagined shape. At some undetermined point during the mental rotation they were shown a test shape and asked to determine as quickly as possible whether or not it matched the imagined shape. The reaction time was fastest when the orientation of the test shape matched the expected orientation of the imagined shape calculated from a predetermined rate of mental rotation for that shape. Evidence for mental transformations of size, also to perform simultaneous shape matching, is provided by Bundesen & Larsen (1975) and for successive matching with a remembered shape by Larsen and Bundesen (1978) as mentioned in Chapter 4.

The implication of these results is that a form of "template" or spatial correspondence matching is necessary between corresponding views of shapes when recognition or matching by features, parts or descriptions fails (see appendix
to chapter 4 of Shepard & Cooper, L. 1982). However the theory that images originated as the top-down component of percepts (Waltz, 1979) suggests another related function for mental rotation and size scaling. If an impoverished or partly obscured perceptual stimulus is to be completed by mental imagery then the image component must be translated, rescaled and mentally rotated in order to superimpose it to the percept.

5.2.7 Mental Image Generation and Maintenance.

In the original Kosslyn and Schwarz model, images were supposed to be generated from two kinds of data in long term memory, (1) Lists of propositions about objects and part relationships and (2) Stored coordinates. These two types of information were posited to be used to position and map the structure of a viewer-centred image onto the "cells" of a limited capacity, and limited resolution spatial array. More recently Kosslyn (1994) has clarified this view to take account of later theories of recognition (see Chapter 4). Thus an object shape and detail encoding system is distinguished from the spatial location encoding system (Kosslyn & Koenig 1992), a distinction I discuss below. Moreover the spatial system is posited to consist of a separate categorical-relation encoding system and a coordinate location-encoding system, (Kosslyn et Al. 1986, Kosslyn et Al. 1992).

Image generation is not a simply a process of the retrieval of intact images from long term memory. Rather images are constructed from object and part descriptions and from the spatial encoding systems mentioned above. It has been shown for example that complex scenes take longer to image than simple ones (Beech & Allport 1978, Kosslyn 1980). Farah & Kosslyn (1981) compared the time taken to image subjectively large versus very small images of the same simple or complex objects. In earlier experiments Kosslyn (1980) had found that it took longer to generate large images of familiar objects than small ones. However Farah and Kosslyn found to their surprise that when subjects were given a mental part verification task (to check that they had generated an accurate image) it actually took longer to generate the small images than the large ones. They had expected to find a speed advantage for generating small images because small parts and details would be omitted, since resolution constraints of the buffer would make their generation difficult. However they were forced to conclude that, if motivated, subjects can with effort and extra time increase the resolution of the buffer to image the smaller parts. Kosslyn et Al. (1983) posit that generating an image of an object occurs in stages, in which an object "skeleton" with part locations is generated first followed by superimposed parts as and when more detail is required.
A series of experiments have demonstrated that mental images are generated in parts that reflect the structure of a stored description, and that it takes more time to generate an image with more parts. In one set of experiments (Kosslyn et Al. 1983) subjects were shown geometrical patterns which could be seen either as several small adjacent shapes (a "tiling pattern") or as a smaller number of larger transparent overlapping shapes. They were given verbal descriptions of the patterns and then asked to memorise them. The mean times to generate mental images of the patterns was linearly related to the number of constituent shapes and also to the type of verbal description used when they were first seen. Other experiments instructed subjects to mentally construct images of animals, which were presented in varying numbers of part drawings on separate pages of a booklet. After memorising six of the drawings presented in this way, subjects were asked to image named animals. To the authors surprise there was a linear relationship between the time to image an animal and the number of parts in which it had been first presented. It appears that although subjects mentally fused the separate parts of the drawings this fused image was not the memory representation used later for generating an image but rather the separately perceived parts. This result is consistent, I think, with the "late conscious construction but early access to recognition memory" theory of perception I supported in Chapter 4. It also seems to dovetail with theories of recognition by parts (Hoffman, D. & Richards, M., 1984, Biederman, 1987). One of the functions of image generation relevant to the sketch may be to synthesise or to re-synthesise, objects or scenes from separately stored parts, in order to encode new relationships in long term memory. The implication is that such a fusion of already encoded visual parts cannot occur without semi-depictive mapping onto a limited capacity spatial array. Again, implicit in the methodology (but not mentioned by the authors), is the use of percept-image hybrids in the initial mental construction process. Further experiments by the same authors showed that the ease of imaging depends upon how much material is included in each imaged unit and on how difficult it is to locate where the unit should be placed relative to the existing portions of an image.

In another series of studies Kosslyn et Al. (Kosslyn 1980 page 106-107) showed how the Gestalt laws of grouping could influence memory encoding and image generation time. Thus subjects were asked to mentally image a uniform matrix of letters described either as "three rows of six" or as "six columns of three". Generation time was longer for the "six columns of three" than for the other description even though the images generated were identical. The conditions were varied in such a way as to show that stored verbal/abstract information can be used in conjunction with depictive information in the generation process. I note
with satisfaction, in Kosslyn's description of these experiments another example of tacit, undisussed hybrid imagery. In the tachistoscope *descriptions were written at the very bottom of the card, leaving plenty of room to image the array*. So when imaging with eyes open from written percepts there must be "room" for the imagery to be superimposed (see below).

In a modification of the Podgorny & Shepard study described above, Kosslyn et Al. (1988) asked subjects to generate an upper case letter on a screen after cueing by its lower case equivalent. They measured the times to decide whether two small projected probe X marks were on or off the imaged letter. Decision times increased with the complexity of the imaged letter forms. They also depended upon which segment of the imagined letter the probes fell. The results suggested that different parts of the letter are generated at different times in sequential order. Subsequently they showed that this sequential construction of image segments mirrored the order in which segments of the letters are normally drawn. If this discovery can be shown to be general and include other objects, then it could have interesting implications for theories of the sketch. For example when an artist draws a new object he may be encoding it in sequentially structured parts. This sequence may in turn influence the way he imagines it before drawing it again in a later sketch. In a sketch with incomplete parts are the first components of the drawing sequence the most effective recognition primes for imagery completion?

Kosslyn's studies suggest that image parts continually fade with time and must be refreshed to maintain them. Thus a limit on the generation of complex images is the effort required to maintain the fading early image parts (Kosslyn 1980, Kosslyn et Al. 1983). If the visual buffer is a shared resource used for short term visual memory, then this finding is expected from studies of short term visual retention (Phillips and Christie, 1977a, 1977b). Little is known about the nature of image maintenance. It may be that allocation of attention to parts of the buffer during scanning or inspection may itself re-activate the image (Kosslyn, 1980, Logie, 1995). Alternatively the image generation mechanisms may simply be continually re-activated.

Recently Farah et Al. (1988) and Kosslyn et Al. (1993) have presented evidence that there are two types of imagery, one that is generated by allocating attention to mentally "draw" a pattern and one that is generated by activating stored visual memories. According to these authors the attention allocating form of imagery is more effective when combined with visual input from the eyes. Presumably vision provides a spatial framework which supports working memory in recording the pattern mentally "drawn" upon a spatial ground. In the theory presented here spatial imagery and object imagery were originally interdependent
and have only become separable in recent evolutionary history. It is assumed that
the attentional imagery system evolved to locate real object or scene positions in
need of superimposed memory completion. Later the brain learned to record
movements of the attentional imagery system for mentally generating new patterns
upon a perceived object or surface, without necessarily invoking stored object
descriptions. Here it is suggested that, whilst sketching, attentional imagery and
memory retrieval imagery are used collaboratively, in keeping with the "labelled
map" analogy discussed in Chapter 3.

In most of the experiments described above, images were generated
voluntarily (following instructions) with the awareness of the subjects. As Kosslyn
and Shin (1994) have pointed out however the possibility of automatic unconscious
image generation is not ruled out. It is surely a common experience that mental
images simply "arrive", resulting in awareness (and sometimes control) of the
products of generation but not the generation process itself (see Pani 1992). This
possibility would mean that image generation is analogous to the processes which
allocate attention (see Chapter 6). There is evidence that the directing of visual
attention has two separate mechanisms of control, an automatic involuntary process
and a controlled voluntary process (Jonides 1981). Here I adopt as a working
hypothesis, the idea that image generation has two analogous modes of control.
This helps to explain the apparently involuntary image generation that accompanies
some kinds of visual stimuli and is used, I suggest, to provide "top-down" support
for the construction of conscious percepts.

5.2.8 Mental Synthesis and Invention.

The ability to use mental imagery transformations to visualise new objects
and make unexpected discoveries is obviously relevant to sketch function. Many
forms of image transformation are implicit in a series of diverse experiments
demonstrating how new images can be mentally synthesised from memorised visual
parts or from verbal descriptions.

Shepard and Feng (1972) showed subjects various flat patterns of six
connected squares that could be folded into a cube. The task was to decide as
quickly as possible whether two different marked edges of a square would meet if
the pattern were folded into a cube. The time required to make such decisions
increased linearly with the number of folds which would have to be made if a real
piece of paper of the shape displayed were actually folded into a cube. Moreover
again we can make the claim that the subjects may have been manipulating a
hybrid image since the flattened cubes were actually displayed so that the mentally
folded shape could have been assisted by partial superimposition of an image upon
the flat patterns. Such evidence for the mental manipulation of perceived flat patterns into imagined three-dimensional shapes, is relevant to the claims made by Fashion designers, for example, that they can visualise the three-dimensional folds and movements of a fabric in an imagined garment (Bates et al., 1988).

Thompson and Klatzky (1978) have shown that people can fuse in their imagination the separately presented parts of a pattern in order to verify whether the synthesised parts match a separate pattern presently afterwards. They found that it was easier to fuse parts which resulted in simple closed forms than parts which formed open and less integrated shapes. The response times (for equivalent numbers of parts) were faster in the former case than the latter.

Finke, Pinker and Farah (1989) showed that people could detect emergent patterns in a mentally synthesised image and reinterpret what the image represents. The subjects followed verbal instructions to mentally combine (using rotation, translation etc) familiar alphanumerical characters or geometric shapes. They were then to report any new shapes or recognizable objects that emerged. When the transformations had been performed correctly (as assessed by subjects' drawings at the end of the trial) the "final pattern was recognised about two-thirds of the time. However when the transformations had been performed incorrectly, the final pattern was never recognised." This suggested they could not have used a strategy based on descriptive knowledge.

Using similar methods, Finke and Slayton (1988) explored the ability to make inventive discoveries. At the start of the trials subjects were shown a limited range of familiar geometric shapes and alphanumerical characters. On each trial subjects closed their eyes and a randomly selected collection of three of these shapes were named. They were asked to mentally join or over-lap all three of the named shapes in order to mentally construct any recognisable pattern or object containing the three shapes. After two minutes they were told to write down the name of any object they had imagined and then to draw it. A separate group of judges assessed whether the drawings were recognisable from the names and if so they rated them for creativity. In a later book Finke (1990) describes the results of these and similar experiments in the use of mental imagery to invent new forms. He also describes extended experiments to investigate the mental discovery of new three dimensional objects. Finke claims that the use of "pre-inventive forms", that is shapes selected for mental manipulation without any pre-determined objective, can act as catalysts for creative problem solving.

Relevant to the design of sketching systems is some research by Anderson, R. and Helstrup (1993). These authors showed that people could use imagery to imagine the results of novel combinations of patterns as effectively as they could
use drawings. Drawings helped to produce more combinations of patterns but not more recognizable or creative ones. Unfortunately these authors did not investigate the advantages or disadvantages of combining imagery with drawing.

There is considerable, if anecdotal, evidence for the role of mental imagery in discovery and invention (e.g. Miller, A. 1984, Shepard 1978, Shepard and Cooper, L., 1982 pages 5-11). Einstein's claim to have had an insight into Relativity theory by imagining what it would be like to travel on a beam of light is often quoted. Kekule's discovery of the benzene ring after dreaming or imagining a snake biting its own tail is also well known. Shepard also quotes as addictive imagers, James Clark Maxwell, Faraday, Helmholtz and Galton. Further accounts include Tesla's "hallucinatory vision" of the self starting, reversible induction motor and James Watson's account of the contribution of imagery in his and Crick's discovery of the structure of DNA. In fact in the book that Shepard quotes, Watson (1968) describes how he manipulated cardboard cut-outs of the nucleotide bases until he suddenly saw that an H-bonded "adenine-thymine pair was identical in shape to a guanine-cytosine pair". Thus his imagery was supported by perceptual aids.

For some unexplained reason most of the reported case studies of discovery by mental imagery are from the sciences. One would certainly expect, as I argued in Chapter 1, that visual mental imagery is a primary cognitive tool in visual arts. Possibly the scarcity of accounts of mental imagery by artists compared to scientists is the relative verbal reticence of artists about their thought processes. Henry Moore has stated that a sculptor "must be an expert in visualising three-dimensional form" and the sculptor James Surls reported that while lying on a couch he planned his next sculpture and "manipulated the image around in his mind. He saw it tumbling and rolling, took an arm off, put an arm on...". (Samuels and Samuels, 1975). In one of the very few reviews of the use of imagery in the arts, Lindauer (1983) is able to quote a number of surveys and studies which purport to show the importance of imagery to artists. Thus Henderson in an unpublished study (Lindauer 1983) sampled 115 autobiographical works by painters, musicians and writers from the eighteenth to the twentieth centuries. She found 501 statements of various aspects of creativity. Of these 19% were in an imagery category. Of these 46% were by painters with musicians and writers tied at 27% each.
5.3 How Sketch Percepts Support Images.

5.3.1 The Need for Perceptual Support: Preliminary Conclusions.

From the brief review above, it can be seen that mental image generation and manipulation is probably an essential cognitive skill for invention in the visual arts. Thus invention, in so far as it relies upon visual experience for its raw material must draw upon the stored visual information discussed in Chapter 4. However these are encoded by automatic processes that can only be inferred and are accessed during recognition without voluntary control. Mental image generation is the only known mechanism by which this vast store of visual knowledge can be retrieved for visually depictive manipulation and mental synthesis (Mandler, J. & Parker, 1976).

However because mental images fade quickly and are often limited in resolution and detail they need perceptual support. This fact is given tacit support by much of the imagery literature but only rarely given explicit recognition and discussion. Thus many of the experiments discussed above were performed with eyes open whilst looking at a marked screen or grid of lines. Farah introduces one of her articles with a typical hybrid imagery task - imagining the layout of luggage in the perceived boot of a car (Farah, 1988a). Kosslyn comments on the need to "leave room on the screen for the image". In one experiment Kosslyn, Cave et Al. (1988) found that subjects were faster to image letters on grids. "Presumably the grids were a kind of crutch" and subjects were not forced to use "coordinate spatial relations to construct the image". We can hypothesize that the contours in sparse outline sketches support imagery by acting as a similar "crutch". Moreover, if as Kosslyn suggests, the initial global shape generation in the image is low resolution then the high resolution of a self-occluding perceived contour line would appear to be an ideal partner. Used with skill an outline sketch may save resources used for the "skeletal" generation and part location finding processes whilst clear spaces allow "room" for fast testing of multiple alternative imagined parts and details. Work by Kelly & Freyd (1987) on the extrapolation of imagined movement from static percepts helps to explain how for example Toulouse-Lautrec is able to suggest a rearing horse (Illustration XXIX)

Whether or not my explanation of the finding that reversible ambiguity is difficult in imagery is correct, we can already understand more clearly the deliberate use of confused and accidental percepts to assist invention (Leonardo, Cozens, Ernst etc. Chapter 2). The confused stimulus for the hybrid-image must be a perpect and must be partly accidental to avoid mental dominance by a single preferred encoding used for image generation. This does not of course conflict with Finke's demonstrations of original invention by mental transformation. Here a
new descriptive encoding is possible but only after synthesis from easily remembered parts.

5.3.2 Evidence for Shared Neural Resources.

There is considerable convergent evidence that mental images use the same neural structures within the brain that are used for visual perception. (See Farah, 1988a for a good review). One source of evidence comes from direct measures of local cerebral activity.

Regional cerebral blood flow can be measured using injected radioactive Xenon. Using this technique Roland and Friburg (1985) compared an imagery task (imagining a walk) with a mental arithmetic task and a verbal task. The imagery task showed a greater blood flow relative to the other tasks in three cortical regions all of which are known to be involved in visual perception and to be anatomically part of the visual pathway. Using SPECT (single photon emission computer tomography) to measure regional blood flow, Goldenburg et al. (1987) showed high regional blood flow in visual cortex related to memorising high imagery words compared to low imagery words. Subjects also showed higher blood flow in the visual cortex (occipital, temporal and parietal lobes) when asked to listen to high imagery statements compared to low imagery statements.

Event-related Potentials (ERPs) (Recorded by placing sensitive electrodes on the scalp) have also been used to compare perception and imagery. Farah et al. (1988) asked subjects to image either an H or a T and then presented them with random faint H or T stimuli. They found that imagery selectively affects the visual ERPs to the stimuli. The effect was greatest in two visual cortical regions.

Kosslyn, Alpert et al. (1993) used positron emission tomography (PET) to study localised blood flow in the brain during visual imagery. They found very similar patterns of regional brain activity when analogous perception and imagery tasks were compared. Amongst the various visual areas (including parietal cortex associated with spatial vision and attention and Area TE probably associated with object recognition) that they found to be activated was a region of occipital cortex, probably area 17 shown by other PET studies to be retinotopically mapped in perception (Fox et al. 1986). This is an important finding because it suggests that primary visual cortex, areas V1 or V2 the earliest visual processing parts of the cortex are also activated by "top-down" information used in imagery. This paper is especially interesting because its authors found that their primary perceptual task (that used by Podgorny and Shepard (198) described above) activated primary cortex less strongly (i.e. caused less regional blood flow) than the corresponding imagery task. They speculate that the perception of clear unambiguous stimuli
How Sketches combine Perception with Imagery.

places less load on visual cortex than imagery because "edge detection and region organising processes" operate less effectively when one must recreate a pattern from remembered information which is typically incomplete. To test this hypothesis, they repeated the experiment using deliberately noisy and indistinct stimuli in the perceptual condition. This time they found that the perceptual task increased cerebral blood flow in primary visual cortex just as strongly as it did in the equivalent imagery condition. They state that these results (in agreement with the theory presented here) suggest that "visual mental imagery may be used in perception itself when one is faced with fragmentary input". In another set of experiments they investigated the effects on regional blood flow in primary cortex of imaging letter forms with the eyes closed so there could be no visual input. The authors point out that the area active in occipital cortex whilst probably area 17 (V1) is also close to the border with area 18 (V2) which cannot be ruled out as an alternative location. However both these regions are topographically mapped. Tantalisingly, they do not mention the possible implications of this discovery for the finding by Peterhans et Al. (1986) that cells in area 18 (V2) respond to cognitive contours. Nevertheless there is now physiological support for the theory that sketch indeterminacies exploit imagery mechanisms used by the human visual system to complete impoverished stimuli.

Brain damage which disrupts visual perception disrupts visual mental imagery in parallel ways. Lesions in the posterior occipital cortex can cause hemianopia (blindness in a limited region of the visual field). Kosslyn, Cave and Gabriel (1989) studied patients who had blind regions so that they could only see objects which subtended a small visual angle. They found that their patients showed a similarly limited imagery field when asked to image objects against a screen and point to the edges of the imagined object. Farah et Al. (1992) used Kosslyn's method (Section 5.2.3) to measure the visual angle of the mind's eye in a female patient before and after unilateral removal of the occipital lobe of visual cortex. Her results were also consistent with the claim that imagery and perception are retinotopically mapped in a similar part of the brain.

A region in the parietal cortex is known to be associated with spatial perception and shifts of visual attention. Damage to the right parietal cortex can cause visual neglect of objects in the left half of the field of view. Patients do not notice and cannot attend to objects in the left half of the field (mapped in the right half of the brain) even though it can be shown that isolated objects can still be seen there. Interestingly it can also cause lack of conscious perception of the left half of objects even when they are scanned (Marshall & Halligan, 1988). Bisiach and Luzzatti (1978) examined the imagery abilities of two such patients with unilateral
visual neglect. They asked them to imagine the cathedral square in Milan, a place well known to both. When asked to imagine they were standing on the steps of the cathedral looking away from it and report what they could visualise, they only reported shops and objects on the right. They reported shops and objects on the other half of the square only when asked to imagine themselves facing in the opposite direction (towards the cathedral). Other tests showed a similar parallelism between visual and imaginable neglect. In a further study Bisiach et Al. (1979) showed that unilateral neglect affects the ability to integrate and distinguish abstract shapes when they are seen as a succession of parts through a slit in a screen. Their results and similar ones obtained by Ogden (1985) suggest that parietal cortex is the location for a spatial representational medium (or "Schema", Bisiach et Al. (1979)) which is used to integrate and inspect (by attentional shifts) both percepts and images.

Evidence that lesions causing content specific recognition disabilities also cause a matching deficit in imagery ability has already been mentioned (Chapter 4). This evidence that recognition and image generation draw on the same stored representations is important to sketch theory. It makes it likely that whenever a sketch percept "part" is recognised, the memory information needed for image superimposition is also primed.

Inferred image generation deficit is one kind of brain damage to imagery which is not shared with perception. Farah (1988b) reviews eight such cases showing that the predominant site of damage was the left occipital hemisphere. Grossi et Al. (1988) report a case of a patient with a left occipital lesion who was able to copy pictures and to match drawings of faces but not to draw a simple object without a model. He could recognise objects even from unusual perspectives but was unable to generate an image of an object not physically present.

5.3.3 Interaction between Perception and Imagery.

As might be expected if visual perception and visual imagery share the same neural resources, incompatible or unrelated perceptual and imagery tasks interfere with each other if they are in the same mode.

In a series of studies Brooks (1968) asked subjects to visualise a block letter, scan the corners, and to respond "Yes" if the corner was on the outside of the figure and "No" if it was on the inside. Brooks found that subjects responses were slower if they had to respond by pointing to a printed word than if they had to speak their response. This result contrasted with a complementary linguistic task where subjects had to say whether or nor the successive words in a phrase were nouns. In this task responses were slower if they had to respond verbally than if
they pointed to a printed response. Brooks concluded that looking and pointing compete for resources with the scanning of a mental image, whilst speaking competes with thinking about a phrase. In another set of studies subjects had to imagine placing numbers in a visualised matrix according to instructions about which squares to use. Their retention of the matrix was worse if they had to read the instructions than if they only had to listen to them. Baddeley et Al. (1975) followed this up by showing that a concurrent perceptual-motor tracking task also significantly disrupted by the Brooks matrix material although it had no effect on the retention of a set of nonsense sentences. (These results are relevant to the design of sketch system interfaces where pointing with a stylus to system menus would be expected to interfere with the users' spatial images of their primary task).

Mode dependent interference between mental imagery and perception was also demonstrated by Segal and Fusella (1970). They showed that sensitivity to a visual or auditory signal is reduced by instructions to subjects to form unrelated mental images, and that this reduction (measured by detection rate) is greater if the image is in the same sensory mode than if it is cross-modal. Thus instructions to form a mental picture of a tree reduced the visual detection rate of a small arrow, more than instructions to imagine hearing, for example, a phone ringing. The same instructions had an opposite effect on the ability to hear a harmonica chord played into ear-phones. This time it was the acoustic imagery that had the greatest effect on the stimulus detection rate. The authors suggest that their data support a model in which interference depends upon the need to discriminate the internal representation of a percept from a memory generated image rather than a simpler channel competition model.

Interference between perception and imagery is certainly relevant to the design of sketching systems. Indeed it is part of my thesis that some existing machine visualising systems are likely to cause perceptual interference with the users mental imagery. However, the main point of this chapter is the hypothesis that, in skilled hands, sketch stimuli support mental imagery whilst avoiding interference. In the above experiments the perceptual tasks were irrelevant to or incompatible with the imagery task. Under what conditions can perception be used to facilitate mental imagery?

The effects of image and percept compatibility were studied by Peterson and Graham, S. (1974) who investigated the effects of compatible and incompatible cuing, with and without imagery instructions, on the ability to detect embedded pictures masked by visual noise. They discovered that compatible cueing increased the detection rate of target pictures whether or not the group received imagery instructions. Incompatible cueing was more effective in increasing errors
with the imagery group than with the control group. Although these results are open to more than one interpretation, they are consistent with the assumption that the cueing generated a mental image which facilitated picture perception in proportion to the degree to which the image and the picture were compatible.

There is evidence that faint percepts can be absorbed into a mental image, sometimes without the subject noticing. Perky (1910) asked 24 students to form mental images of familiar objects against a two-way mirror. Unknown to them, she then gradually projected from behind a faint actual image of the same object. None of her subjects noticed that a real image was present and attributed characteristics of the real image to their mental images. Unfortunately these early experiments were not well controlled for the effects of the task demand characteristics.

However better controlled studies by Segal and Nathan (1964) and Segal (1971, 1972) also showed that faint visual percepts may be assimilated into a mental image. For example, asked to image a city skyline, but shown a tomato (inside a hood with a concealed projector), several subjects described the skyline with a round sun behind it setting in a red sky. Segal (1972) interprets these, and other results, as evidence that percepts can be fused with images and even that the image may be constructed around an undetected physical signal.

Other experiments demonstrate that compatible mental images can facilitate percepts. In a rather neat experiment, Martha Farah (1985) asked her subjects to memorise and then mentally image either an uppercase H or a T. Whilst doing this with eyes fixated they were given a forced choice detection task of similar 'H's or 'T's in random order. Farah found that the detection rate for either letter was facilitated if the same letter had been imaged, and moreover that this facilitation was dependant on the image and the stimulus letters being in the same position in the visual field. The fact that the facilitation was dependant upon location suggests (to me) that an alignment match between imaged letter and stimulus letter preceded the formation of a percept-image hybrid. It was the increased saliency of the hybrid which caused the facilitation. In a later paper Farah (1989) examined the facilitation of small threshold stimuli which either fell on or off the mentally projected image of a letter form. Stimuli falling on the image were detected more often than stimuli falling off the image. When the facilitation was analysed in terms of signal detection theory, it was found to consist only of criterion lowering and not of enhanced sensitivity. Thus false alarm rates increased as well as hit rates. Qualitatively, the effects of imagery in this experiment resemble the effects of attention and not passive perception. Farah interprets her findings to support a theory of Uric Neisser (1976) that mental images are "perceptual anticipations" or "plans for the future pick-up of perceptual information". Strangely I cannot find
support for her claim that Neisser (1976) linked this theory to attention or that it is an attentional theory of imagery.

Another type of facilitation has been shown by Freyd and Finke (1984). Subjects were presented with cross-shaped patterns consisting of a vertical line bisected by a horizontal line that was slightly different in length. Their task was to say which of the two lines was the longer. On half the trials subjects were asked to imagine either an outline square or a diagonal cross centred over the two lines. The sides of the squares were slightly shorter than either of the two lines. An actual square displayed over the lines was shown to facilitate the line discrimination making it easier to detect the difference. The results showed that the imagined square also facilitated the discrimination of the line lengths whilst the X pattern did not. This is clearly an experiment in the formation of a hybrid image although not presented as such.

5.3.4 Evidence for Percept - Image Hybrids.

The above evidence that images can interfere with perception and that perceptual acts can interfere with imagery tasks is persuasive. So also is the more limited evidence that imagery can be used to facilitate perceptual recognition and discrimination. However in order to strengthen the claim that untidy sketch attributes, (as used by Leonardo and others see Chapter 2), support imagery, we need evidence of image facilitation by percepts and that imagery can be combined in spatial register with a percept. Very surprisingly there is much less experimental study of perceptual support for such "projected" imagery. We have seen that prominent work on imagery shows tacit recognition for perceptual support, without clarifying its nature. Thus Kosslyn remarks on the "need to leave room for the image" when subjects are following imagery instructions whilst viewing cards containing written material. Similarly, Farah refers to the "hybrid" image of her car boot. Moreover hybrid imagery is tacit in all experiments on imagery which use perceived grids or screens.

An exception is a series of studies by John Hayes (1973) of imagery in elementary mathematics, which has been neglected (or mis-reported) in most relevant reviews (e.g. Kosslyn 1980, 1983, 1987, Farah 1985, Finke 1985, 1989). Skemp's (1971) suggestion that mathematical notations promote the use of imagery in problem solving led Hayes to explore the nature of such notation related imagery in mathematics. In one study Hayes used protocol analysis to record his subjects reported use of imagery. He gave his subjects cards containing diagrams of elementary mathematical problems which they were asked to solve and then to later report any imagery which they used. 17 out of 19 subjects reported that they
used imagery to solve the problems and most that the images were superimposed on the cards in the same plane as the stimulus symbols. In solving card 4 (below) three subjects complained that the line drawn round the symbols interfered with their imagery. One said that he had to mentally remove the "bars" to move the symbols and two others felt constrained to move the symbols as if they were tiles in a frame.

\[
\begin{array}{c}
5 + z = 9
\end{array}
\]

Card 8 caused even more trouble

\[
\begin{array}{c}
589 \\
+734
\end{array}
\]

Subjects complained that there was not room on the card to place (mentally) their answer. One subject reported that to solve the problem he mentally "scotch-taped a piece of paper underneath to write the answer on".

In solving the cancellation problem

\[
\begin{array}{c}
4 . C . S . R \\
R . W . C
\end{array}
\]

most subjects reported that they generated either two or four imagined slash marks to cross out the two R's and C's. Three of the subjects said that they solved the problem by examining the hybrid image thus formed and reading off the unmarked letters (This agrees with my own introspection).

In the addition and subtraction problems most subjects worked from right to left, processing each column in turn and storing the results as images. One subject commented that the "important thing about generating imaginary figures is that they stay while you move to the left to do more calculation". Three subjects volunteered the fact that they had to read the answer quickly because the image faded rapidly. Others said that they went back to "refresh" the image of the partial answer before proceeding with the rest of the calculation.

These reports agree with the subjective experience that in sketching it is important to omit denotational marks which may hinder imagery movement, and
that it is necessary to leave space within the sketch for its imagery components. Millet's "The curtain of trees" (Illustration XXIV) is a good example of a sketch which generates imagery as much by what it omits as by the skeletal marks that it contains.

In another study Hayes (1973) obtained more objective evidence that hybrid images really are generated in spatial register with the supporting percept. He noticed that European-trained subjects found a division problem was more difficult by placing the 2 figure divisor to the left of the 3 figure dividend (as is usual in the U.S. and the U.K. e.g. as 151673) because they had been trained to place the divisor on the right (e.g. as 673115). Hayes considered two models of the solution process. In the first the subject generates an image in the familiar format and then solves the problem using the generated (but not a hybrid) image by the usual manipulation. In the second model the subject might solve the problem with a hybrid image correcting the unfamiliar to the familiar format digit by digit. Thus the effects of the unfamiliar versus the familiar format would, if the first model is correct, be expected to be limited to the first stage of the solution process, whilst the second model predicts that the effects would be distributed digit by digit throughout the solution process. He tested this prediction with a group of 3 European and 9 American subjects each with 12 familiar and 12 unfamiliar 2 digit into 3 digit division problems. The difference in the means of subjects times to compute successive digits of the answers for the familiar and the unfamiliar formats were distributed throughout the solution process confirming the hybrid model. It is a pity that comparable studies, (perhaps using the Finke (1990) methodology) on hybrid imagery in visual invention are not yet available.

It has already been suggested that sketches (e.g Picasso's studies for "Les Demoiselles D'Avignon" Illustrations XI, XII) provide substitutes for Kosslyn's posited image skeleton and thus reduce the cognitive effort for image maintenance, whilst leaving space for fast image transformations. A possible analogy for this function is the role of the chess-board and pieces in chess (Chase & Simon, 1973). Most of the moves considered in a chess game are played only in the minds of the players. According to Reuben Fine, a chess grandmaster who is also a psychologist, chess skill depends upon making mental images of possible positions (Holding, 1985). In a number of studies Chase & Simon compared chess players of different strengths in their ability to remember chess positions, to perceive threats, find good moves etc. They concluded that the thought of skilled chess players relied upon "perceptual structures" which they retrieve from long term memory. These perceptual structures are used to generate mental images which can be superimposed upon the "external representation of the chess board". They
conclude that "This capacity to construct an image combining perceptual structures from internal memory with sensory features from external memory is probably one of the very basic cognitive processes." They call the meeting point the "visual vestibule". This is perhaps an intuitive 1973 anticipation of the "visual buffer" (Kosslyn, 1980) or "visuospatial working memory" (Baddeley, 1986, Logie 1995).

One of the problems of "forward search" in chess is that in visualising the position several moves ahead the unchanged actual positions of the pieces can interfere with the attempt to imagine their new positions. It is easier to position an imaginary piece upon an empty square than upon one already occupied by another piece (see discussion of the chess "blind spot" in Kotov, 1971). One sometimes sees chess masters in tournaments shielding their eyes from the board as they think. This suggests that the board and pieces is a necessary but far from perfect crutch for the "mind's eye". By analogy, some perceptual attributes of sketches (and other visualising aids) may also cause unintentional image interference where facilitation is needed.

The evidence available for hybrid imagery in sketching is circumstantial, but I believe persuasive. This case would be stronger still, if it could be shown that normal perception has an imagery component whenever it is impoverished or ambiguous as suggested in Chapter 4. Does the hybrid image hypothesis have explanatory value in normal perception?

5.3.5 "What" and "Where" in Perception and Imagery.

In Chapter 3 a case for the relevance to sketch function of the distinction between spatially "depictive" and "descriptive" forms of visual representation was made. It was also pointed out that there are many intermediate types of representation such as the "labelled map" or "pictorial description" (Ullman, 1989). In Chapter 4 evidence was presented that both spatially mapped (depictive) and descriptive (encoded parts and features) representations are used to perform different types of recognition task.

This conceptual distinction can be related to evidence about the modularity and division of labour within the visual pathways of the brain. Neural signals from the retinæ pass first to paired midbrain structures the Lateral Geniculate Nuclei (LGN) on their way to the primary visual cortex. Here two types of cells, large (magnocellular) and small (parvocellular) are neatly segregated in six alternating layers. Responses in magnocellular layers are transient, sensitive to movement (have a high temporal gain), sensitive to small luminance contrasts (high contrast gain) but have a low spatial resolution (large receptive fields) and are insensitive to colour. Cells in the parvocellular layer have lower temporal and contrast gains but
much higher spatial resolution and carry colour information. There are at least 32 separate areas of the primate brain (mainly in the folded outer layer the neocortex) responsible for visual processing (Braitenburg 1985, Van Essen 1985). Many (but not all) of these contain organised columns of cells with receptive fields which topographically map the retina. It is known that different visual cortical areas are specialised for different aspects of visual analysis (see Zeki, 1993) although this specialisation is imperfectly understood and there is much apparent functional overlap. Moreover the cortex is subdivided into a least six functional layers (with further sub-divisions) and in those visual areas which have been studied in most detail there is a parallel segregation of function between layers (Hubel & Livingstone, 1987). From the LGNs the signals pass to visual area V1 where the retina is mapped at the highest resolution. Here, parvocellular and magnocellular separation is preserved, in parallel, in different layers and a further division of labour has been documented. Cells responding to moving edges and stereoscopic depth can be traced to the magnocellular stream whilst those responding to oriented edges and to colour but not edges (the cytochrome oxidase blobs) have a parvocellular origin. V1 connects to V2 where these three streams are further segregated. Livingstone & Hubel (1987) have collected psychophysical evidence to support their contention that at least three separate visual channels are used for (a) movement, depth and low resolution form (magnocellular path), (b) high resolution form (parvocellular path) and (c) colour (parvocellular path). Livingstone (1988) discusses how physiological properties of these three pathways explain aspects of art. For example why it is possible to paint a good likeness in unusual colours, provided luminance contrasts are preserved, whilst colour contrast at equiluminance fails to provide pictorial depth. Relevant to the sketch is the suggestion that colour which overflows the boundary of a form appears to be attached to that form because cells in the colour channel have a lower spatial resolution than those responsible for shape.

The magnocellular (M system) - parvocellular (P system) distinction overlaps with a distinction based on the effects of localised brain damage in human and non-human primates. It has been shown by Mishkin et Al. (e.g. Mishkin, Ungerleider & Macko 1983) that lesions in the parietal cortex of the macaque monkey disrupt the ability to spatially locate objects whilst leaving the ability to recognise them intact. In contrast, lesions in the infero- temporal cortex damage the ability to recognise objects whilst leaving intact the ability to locate and grasp objects. A similar double dissociation between object recognition and spatial perception has been found in human patients with lesions in the corresponding parts of the cortex. Mishkin and his colleagues postulated two parallel visual
A diagram showing neural connections between various visual regions of the macaque cortex. Modules on the left roughly correspond with the "dorsal" or parietal pathway and those on the right to the ventral or "recognition" pathway, but as can be seen there are many cross connections as well (Modified from Van Essen 1985) (see text).
pathways. One is **Dorsal**, the "Where" pathway concerned with spatial vision, connects Primary cortex (V1) through areas V2, V3 and V5 (known to be specialised for movement perception) to the parietal cortex where the visual cortex is less well understood but is known to have separable functions related to visual attention, spatial perception and spatial memory (Ellis, A. and Young, 1988). The other a **Ventral** route, the "What" pathway, connects primary visual cortex through V4, (shown to be important for colour vision), to inferotemporal cortex, a region critical for object recognition.

Zeki (1993) (an authority who has pioneered the study of visual specialisation within the cortex) has criticised, what he calls the *"What and Where doctrine"*. He points out that the visual cortex cannot be organised upon only two such pathways. Zeki accepts that the magno and parvo pathways are concerned with different attributes of visual processing. However he points out that they are richly interconnected and that there are many more than two routes from V1 to other visual areas. Moreover he claims that certain visual areas have both "What" and "Where" attributes. For example areas V3 and V3A contain orientation selective cells which are commonly *"gaze-locked"* (they respond only to a particular orientation if the monkey is looking in that direction). Thus according to Zeki, they would be classified for both spatial vision (where) and for form (what). He himself summarises the evidence as indicating four pathways.

*"We can therefore speak of two form pathways, one derived principally from the M system, and much more concerned with dynamic form, and the other derived principally from the P pathway and much more concerned with form in association with colour. In addition to these, there is a principally M-derived motion pathway and a principally P-derived colour pathway. There is ample opportunity for the P and M signals to mix in the cortex, so that the input to the specialised visual areas may consist of signals from either source."* (Zeki 1993, original in italics)

Diagram II shows the connectivity of some of the more important visual regions of monkey cortex. The hypothetical "where" pathway (M-system) passes from the eyes to V1 and then dorsally through V2, V3A, MT (movement), 7A (spatial attention) to TF (in the frontal cortex). The hypothetical "what" pathway (P-system) passes from V1 through V2, and diverges through VP, V4VA, and V4 towards a number of infero-temporal regions believed to be responsible for object recognition and also to area TF. It is interesting that TF is close to a region believed to be necessary for the conscious control of attention (Posner & Rothbart 1992 see also Chapter 6).
Despite Zeki's reservations the terms "where" and "what" will be preserved here for the following reasons. The distinction between the cerebral representation of spatial structure in association with movement (used for orienting attention and for guiding overt and imagined movement) and the representation of form in association with colour (used for recognising and categorising objects) has considerable explanatory value for the theory of sketches. However the terms "what" and "where", as used here, are not intended to imply isolated processing pathways but different, neurally implemented forms of visual representation. There is convergent evidence that the dorsal "where" pathway includes one or more topographical spatial maps which can be linked, possibly by movements of focal attention (Treisman & Gelade 1980) to expandable hierarchies of descriptive information about possible objects and object parts (the ventral "what" pathway). For all well studied visual areas in the cortex the number of ascending neural connections is matched by an equal number of descending re-entrant connections (Braitenburg 1985, Van Essen 1985). It is amusing that originally neuroscientists were puzzled by this and found it difficult to suggest functions for so many returning connections. Edelman (1989) and Zeki (1993) have suggested that such re-entrant axons are used for integrating separately processed visual attributes in consciousness. The theory of perceptual completion and hybrid image formation in sketches also requires the existence of such re-entrant connections. An elegant advantage of having separate representations of spatial location and object form is that it may facilitate fast preliminary movements of attention (Posner et Al. 1994) to locations of interest (where system) without burdening working memory with details about what is represented at each primed location until required.

There is evidence that the same modular distinctions found within the visual system are mirrored by the mental imagery system, as predicted by the shared resources hypothesis. For example Levine et Al. (1985) studied two patients with impaired perception and imagery due to bilateral posterior cerebral lesions. One who had impaired recognition for faces and animals and lacked colour vision showed matching imagery deficits. The other had visual disorientation combined with an inability to describe spatial relations from memory (spatial imagery deficit). Thus "what" and "where" are dissociable systems in imagery as well as perception.

Using similar evidence Farah et Al. (1988) argue for a similar dissociation between "spatial" and "visual" imagery systems. In my opinion this distinction is rather muddled. It is based on the premise that "spatial" cognition is not only visual but multimodal since we can perceive space auditorily or by combined movement and touch. There are two other sorts of evidence. One consists of
How Sketches combine Perception with Imagery.

Evidence for imagery processes in the congenitally blind using tactile analogs of visual stimuli. Thus the blind can perform mental rotation tasks (Carpenter et al. 1978, Marmor & Zabeck, 1976), mental scanning tasks (Kerr, 1983) and imagery mnemonic tasks (Jonides et al. 1975, Zimler & Keenan, 1983). Whilst this may well be due to the fact that visuospatial and tactile-spatial task share resources it does not seem to me to justify a distinction between "spatial" and "visual". It is possible that either, (a) separate visual and tactile spatial regions of the cortex can perform analogous operations, or (b) that a shared resource which evolved for high-level vision is intact and functional in some blind people and is used for unconsciously visual-like imagery processes. Indeed some blind people report imagery like experiences and can even paint using imagery (Finke, 1986, 1989).

Another type of evidence is provided by studies of the interfering effects of spatial tasks on spatial imagery and of visual but non-spatial tasks on non spatial imagery. Thus Baddeley & Lieberman (1980) (following up Baddeley et al. 1975) showed that the Brooks matrix task (which is disrupted by concurrent visual tracking of a moving target) was also disrupted by concurrent tracking of an auditory moving pendulum by blindfolded subjects. On the other hand performance of the matrix task was not affected by a concurrent brightness judgement task. They concluded that the Brooks matrix task drew upon a "spatial" rather than a "visual" resource which is involved in spatial perception and also in motor-control. In a series of studies Logie (1986) demonstrated that visual imagery tasks which are not disrupted by concurrent tracking are disrupted by a concurrent visual task. Logie used performance at a memory by association technique known as the "one-bun pegword mnemonic" as a measure of non-spatial visual imagery (described in Paivio 1971b). This is a technique for remembering lists of words based on forming mental images of objects in unusual associations. Subjects were presented with concurrent visual patterns which they were told to ignore as far as possible whilst concentrating on remembering a list of words by the imagery method. Concurrent visual patterns were found to disrupt recall of the words when the imagery technique was used but had no effect upon recall using rote rehearsal. A number of other studies (reviewed Logie 1995) have also demonstrated a disassociation between the two types of interference.

The above evidence elegantly demonstrates a parallel distinction between two types of visual and two types of imagery process. However I find the fashion for referring to this as "spatial" and "visual" rather illogical. There is surely good evidence that the spatial resources shared by visual perception and visual imagery tasks are part of the visual system (even though they may also be used by other
How Sketches combine Perception with Imagery.

modalities). Surely also, the process of perceiving or imagining the form of visual objects and patterns has a spatial component? Thus the terms "spatial" and "visual" do not refer to disjoint sets. In the context of visuospatial working memory, Logie admits that the distinction between "spatial" and "visual" is unclear (Logie, 1995 pages 77-79). He makes the interesting suggestion that the term "spatial" can be interpreted to refer to a representation that involves movement in a broad sense. This would include not only physical movement and imagined movement but also movement planning and movements of attention or mental scanning.

I suggest that an single explanation for all the above distinctions lies in the existence of the two different formats for representing visual information discussed in Chapter 3. An array format, depictive representation of spatial structure lies in the dorsal parietal pathway and a descriptive hierarchical representation of object colour and form in the ventral, infero-temporal pathway. There is a necessary connection between spatial depiction and movement because spatial depiction in an array is the form of representation from which implicit distances and directions between arbitrary locations can be extracted. Also movements of attention are implicated in extracting implicit geometrical information from spatial arrays (Chapter 3). As already discussed, the descriptive representation of object attributes is necessary to explain object recognition, categorisation and abstraction. Spatial tracking interferes with the Brooks matrix task because the latter requires movements of covert attention and because both tasks compete for the same spatial array. Visual patterns and irrelevant pictures interfere with the imagery mnemonic tasks because both compete for a short term descriptive buffer in which object attributes can be accessed but movement or spatial scanning is not possible. The two systems are interdependent, with necessary connections between spatial location and descriptive information making an harmonious combined descriptive-depiction.

It is now possible to relate the hypothetical hybrid imagery function of sketches to their intermediate depictive-descriptive status. Spatially represented components of the sketch (which support overt scanning) will amplify and combine with spatially represented components of the users imagery which support covert mental scanning ("Where" system). Descriptive components of the sketch, (principally object colour and form) will elicit and combine with descriptively represented imagery components. Commonly both types of hybrid representation will be interacting in the users mind during the acts of generating and inspecting a sketch. However, as I shall suggest in Chapter 4 some sketches may differentially support spatial (depictive) imagery whilst others differentially support object (descriptive) imagery components.
5.3.6 Working Memory and Mental Imagery.

A crucial cognitive resource for planning and designing imaginary objects in the visual arts is an easily manipulated temporary memory system with fast access. Involuntary temporary storage is implicit in the late construction theory of visual perception. Voluntary temporary storage is necessary for reasoning and imagining and for planning and monitoring overt motor activity such as drawing. Psychologists have long distinguished such a working memory system from long term memory containing stored knowledge about the world (semantic memory) or of experienced events (episodic or autobiographical memory). In working memory, information must be quickly accessed but can be discarded after a few seconds or minutes. In contrast in long term memory, information may last a life time but cannot be consciously manipulated and is not always easily accessed. Working memory is necessary for both perceptual and imagery components of the act of sketching. When drawing from a model the shape for an intended line must be remembered whilst the eyes are transferred to the paper. When correcting a contour whilst drawing, memory for either imagined or recently erased contours must be available for comparison. To develop and correct a chain of visual thought one must remember the immediately preceding stages of the chain.

The Baddeley and Hitch model of working memory has been developed as the result of several years experimental study of memory capacity (Baddeley & Hitch, 1974, Hitch 1980, Baddeley, 1986). The original model had three main components which are distinguished by studies of performance. A general purpose, modality independent, resource the Central Executive, used in reasoning and planning is posited to be able to initiate and monitor a number of sub-processes (Baddeley 1986). This is the least understood of the model's components, but is similar in function and attributes to the "supervisory attention" control system posited by Norman and Shallice (1980) in a theoretical model of voluntary and automatic behaviour (discussed further in Chapter 7).

Two slave memory systems, co-ordinated by the Central Executive, are postulated in the model. Most studied has been the articulatory loop, responsible for the temporary retention of speech based material. There is evidence that it consists of at least two components, a phonological or acoustic store and an articulatory rehearsal system, a sort of "inner voice" which is used to refresh or maintain verbal material in the phonological store. The evidence in support of this part of the model is reviewed by Baddeley (1986) and Salame and Baddeley (1982, 1989). Relevant to the function of the written notes frequently found on artists' sketches (see Chapter 2) is the evidence that the articulatory loop is used in reading and verbal reasoning (Baddeley, 1986). This suggests another function for written
notes. It is possible that despite their visual format some (not all) written notes have the function of releasing internal visuospatial resources for more spatially dedicated tasks. These two components are sometimes referred to as "the inner ear" and the "inner voice".

The other slave component of working memory the visuospatial sketchpad (Baddeley, 1986) or simply visuospatial working memory (Logie, 1995) is especially relevant to sketch function. For many years the Kosslyn model of mental imagery and the Baddeley and Hitch model of working memory were developed independently, using different types of evidence and making few if any references to each other. It was therefore not clear to what extent Kosslyn's "visual buffer" could be identified with "visuospatial working memory". Recently however Logie has attempted to bridge this gap (Logie, 1991, 1995). As described in the preceding section, there is persuasive evidence, based on dual task performance, that visuospatial working memory has at least two sub-systems a "spatial" or location memory and a "visual" or object attribute memory responsible for colour and object form. However although he recognises that they share many features, Logie is not convinced that the "visual buffer" is identical with "visuospatial working memory". He quotes evidence that imagery generation and manipulation is the responsibility of the central executive. Logie views Kosslyn's "visual buffer" as "more like a workspace for visual imagery than a buffer as such". He suggests that visuospatial working memory, by analogy with the phonological loop, consists of a passive short term visual cache larger in capacity than the conscious mental image and a spatial memory system "the inner scribe" which refreshes its contents by active rehearsal or mental scanning. The visual cache idea makes sense of Kosslyn's otherwise puzzling theory of the mental "window of attention" into which parts of the image can be moved for detailed processing. For example, the question "how is the image within the attention window refreshed?" is unanswered (Kosslyn and Shin, 1994). If the visual cache suggestion is accepted, then sketches may be viewed as supporting not only the currently conscious mental image but also a larger short term memory in which overt eye movements are used to support covert movements of attention to parts of the visual cache. Thus the superimposed imagery components of the percept-image compound can be expected to fluctuate in piece-meal fashion with spatial movements of attention.

A connection between the spatial (depictive array ?) component and memory for movements has been demonstrated, which could be important to sketch function. Thus in a series of experiments, Quinn and Ralston (1986) had subjects move their unseen arms around a square matrix taped to a table, concurrently with performing the Brooks matrix task. They found that movements
An outline model of processes used in sketch hybrid image formation suggesting a symmetry and shared resources between perceptual processing and memory retrieval processes (see 5.3.7).
which were incompatible with the mental movements used in the matrix task interfered with retention of the matrix whereas compatible movements did not. Compatible arm movements were those which mirrored in direction the instructions in the Brooks task such as "In the starting square put a 1, in the next square to the right put a 2, in the next square down put a 3" etc. Now, in normal sketching on paper, movements of the pencil whilst drawing can be assumed to be compatible with internal mental movements whilst the artist is considering a particular form. It seems possible, that such movements of the pencil actually facilitate corresponding movements of spatial thought. However with many computer sketching systems forced arm and hand movements, to menu functions for example, are incompatible with the users covert imagery and are thus likely to interfere with visual thought.

5.3.7 Towards a Model of Percept - Image Interaction.

5.3.7.1 Hybrid Images: A Global Outline.

It is now possible to outline a hypothetical model of percept-image interaction within the sketch. Diagram III shows a global model of the processes posited in the formation of percept-image hybrids (after Fish 1991). The model may be criticised as too complicated. However it is difficult to imagine how existing knowledge can be summarised more simply. In truth the model already contains several simplifications of detail and an exaggerated symmetry. Here I defend only its broad outline fully aware that it may contain faults or omissions of detail. It is an attempt to relate the known attributes of sketches with the "late construction" theory of perception, the Kosslyn model of imagery and existing data on percept-image interaction. The model is intended to explore the hypothesis that there is some symmetry (resulting from evolution and shared neural resources), between visual perception and visual memory retrieval. This is reflected in the diagram about the vertical axis. A different symmetry in working memory related to that suggested by Logie (1995) between descriptive and spatially depictive processes is reflected about the horizontal axis. In the diagram the ellipses symbolise representational structures whilst the arrows and rectangular boxes symbolise processes manipulating the information represented. Circles symbolise combined representation and process systems. Two artificial and probably exaggerated symmetries have been emphasised. On the far left of the diagram are the cultural depictive and descriptive visual stimuli used in the sketch interacting with hand and eye. These are symbolically mirrored on the far right by the corresponding stored representations in long term memory. Here I should perhaps have made the descriptive depictive distinction of non-verbal information clearer and I could perhaps have incorporated Kosslyn's distinction between categorical
How Sketches combine Perception with Imagery.

and coordinate representations of spatial structure. But that would have made the model even more complex. Processes to the left of the central vertical axis are stimulus derived visual processes divided (with Zeki's reservations in mind) into dorsal route spatial and movement depictive systems (lower half of diagram) and ventral route object attribute descriptive systems (upper half of diagram). Most of the arrows are assumed to be physically implemented in the ascending, outwards connections between specialised areas of the visual cortex. Diagramed to the right of the central axis are corresponding representational systems derived from long term memory, with again the lower half showing memory derived spatial representations and the upper half memory derived object descriptions. These processes are assumed to share cortical resources with the stimulus derived processes and account for the existence in almost all cortical visual areas of the rich descending, re-entrant connections. The whole of the large central circular area of the diagram represents processes and storage components which are assumed to be part of a multi-component "working" or cache memory system. However only information which has passed the depictive and descriptive gates (the two triangles in the diagram) can contribute to visual awareness and volition (see Chapter 7).

Both percepts and images are simultaneously spatially structured and descriptively interpreted. Thus there has to be a mechanism within the brain for linking the separately resourced representations of depictive spatial layout and descriptive structure. In the diagram I have followed the suggestion of Treisman and Gelade (1985) and assumed that focal attention to objects and object parts, provides one mechanism by which separately represented object attributes and features become integrated at specific spatial locations (see also chapters 6 and 7). However it is surely necessary to suggest also a mechanism by which important or typical spatial-descriptive links are learned and re-accessed for fast perceptual completion or image generation. For this reason Kosslyn and Koenig (1992) suggest an "Associative memory" with this specific function. As the diagram is already complex I have avoided adding yet another module. I suggest that some kind of "associative memory" is implicit in the diagram where I have drawn ascending and descending arrows between the long term memory representation of descriptive hierarchies and "spatial memory".

5.3.7.2 Percept-Memory Comparison Buffers and Gates

Most of the relevant literature (eg Farah 1985, Kosslyn & Shin 1994) attempts to explain the data on percept-image interference and facilitation by assuming that images and percepts simply share resources. Whilst this might
How Sketches combine Perception with Imagery.

explain Farah's (1985) results on image facilitation of percepts, it does not explain the interference data and offers no mechanism to control the proposed functional construction of percept image hybrids. As Finke & Shepard (1986) have pointed out, existing theories of percept-image interaction do not easily explain why instructions to image sometimes causes interference with perception and sometimes facilitate it. Moreover as Neisser (1976) has pointed out the shared "visual buffer" hypothesis does not explain why we rarely confuse mental images with percepts.

For these reasons I suggest in the model the following refinement to the shared visual buffer hypothesis. Instead of a single shared buffer I suggest that memory derived and stimulus derived part representations are temporarily mapped separately, side by side and in the same format, in special twinned Percept-Image Buffers in short term memory. The two buffers are connected to Gates which feed selected components of the two comparison buffers into a shared visual Integration Buffer. Depending upon the results of a comparison between corresponding components of the stimulus derived and memory derived buffers, the gate will either suppress one or other component or - when memory image and percept image match both spatially and descriptively - the gate will pass elements of both forming a fused image in the Integration Buffer. I term this combined structure of three physically interleaved representations a percept-image "Comparison Buffer". Because the three representations in a Comparison Buffer must share the same information format I suggest there must be distinct types of comparison buffer corresponding to spatial (depictively mapped) and object attribute (hierarchically descriptive) working memories located in parietal and infero-temporal cortex respectively. In view of the evidence already mentioned (Fox et Al. 1986, Kosslyn et Al. 1993, Farah et Al. 1992) that imagined and perceived visual receptive fields are congruently mapped in occipital cortex (areas V1 or V2) there is also the fascinating possibility that there might be a comparison buffer in "primary" cortex. This would have the function of matching high resolution object detail at the fovea with expanded stored information, provided by the re-entrant connections, and of providing a pathway for integrating high-level and lower-level spatial and object information (as hinted by Zeki 1993).

In the model I propose, an array structured 2½D spatial "Comparison buffer", houses percept (stimulus generated) and image (memory generated) location maps. In the diagram I have tentatively drawn small "foveal" or "focal" centred maps for detailed comparison within larger low resolution head field maps (Sanders, 1963) formed by saccadic integration. This is roughly equivalent to Kosslyn and Koenig's "attention window". The diagram shows only a logical structure. Physically the neuroscience suggests that these different types of spatial
representation are anatomically separate (see Stiles-Davis et al. 1988). The term "comparison" is preferred to "matching" because the processes by which compatibility between sketch percept and sketch image are determined cannot be based only on forming a match. Initially (I suggest) landmark features are used to perform Ullman type local alignments between sketch contour fragments and the edges of candidate images, which have been generated and suitably transformed from encoded descriptions pre-linked to corresponding locations of the two maps (Ullman, 1989). It is suggested that the spatial comparison buffers have overlapping receptive fields of different sizes, corresponding to parallel spatial frequency filters as are known to exist in areas V1 and V2. The template fitting process might then be facilitated by matching locations in sequential order from low frequency to high frequency. Thus rough contour and position matches would precede more refined ones. (A similar application of spatial frequency filters has been suggested for visual position grouping algorithms by Watt (1988) and to binocular matching (Arditi, 1986). The existence of independent spatial-frequency tuned channels for binocular fusion and rivalry has been demonstrated by Julesz and Miller (1975). See also Mayhew & Frisby 1980.

Where there are non-matching empty spaces in either the stimulus map or the image map (weakly activated receptive fields), a search of the sketch schema (activated in descriptive working memory) will elicit candidate image parts in the complementary buffer. Thus the model assumes that a process that can compare nodes in twinned descriptive hierarchies of scenes or objects is used in close cooperation with the alignment test for compatibility between the spatial maps. The corresponding twinned descriptive buffers are drawn separately in the upper part of the diagram. It is posited that the descriptive image in the buffer is assembled from long term memory by two possible mechanisms. In "descending" assembly, a global super-ordinate sketch schema (or anticipatory hypothesis) is used to search for associated sub-ordinate scene or object component parts. In "ascending" assembly a sub-ordinate object or part is used to search for associated super-ordinate objects or scenes (see "Recognition by Components" 4.3.4).

Thus according to the model two parallel and richly cooperative matching processes are necessary to support the hybrid image theory of sketch perception. Moreover there must be a fast and easily manipulated "map and pin" type link between the spatial location maps and descriptive detail. The rich interconnections between relevant parts of infero-temporal and parietal cortex is available for this purpose. Such connections are not a problem for the model, but a necessity. Probably the initial perceptual descriptive tree will be matched with its most likely anticipatory schema from memory at their respective "roots". This scheme then
becomes the global model which drives the location matching processes in the spatial buffer. Mismatches at particular locations can then be resolved by looking up alternative attributes in branches of the descriptive schema and re-matching locations. As the spatial and hierarchical resolution of the two matching processes increases, so the depictive and descriptive information converge with Kosslyn's "categorical to coordinate" translation process being used to "zoom in" on detail in the memory half of the spatial comparison buffer. The descriptive gate has the function of keeping the hybrid (part perceived - part imagined) description consistent. Simultaneously the gate in the spatial array buffer must resolve spatial conflicts between image and percept even if the descriptions match. If different descriptions are found at the same location the gate will select the most strongly activated location. Normally this will be the percept and the weaker conflicting image component will be inhibited. However where the image is very strong and the percept component weak or confused the percept component may be gated out (Perky 1910, Segal 1972). In ambiguous sketches, as the description oscillates, so we may notice or not notice inconsistent components at some locations (the eyelash in Boring's "Wife or Mother-in-Law" figure is an example. (Illustration XXXI a)). At a location where image and percept components match to a sufficient degree the gate allows both signals to pass to the Integration Buffer resulting in a fused percept-image part with combined signal strength. The resulting stronger activation at that location accounts, I suggest, for the position dependent facilitatory effects found by Farah (1985, 1989). The extra processing which must precede the inhibitory gating out (blocking from awareness) of an inconsistent image (or percept) part and the possibility of position dependant percept-image competition for the integrative buffer explains the percept-image interference results of Peterson & Graham, S., Segal and Fusella and others. Another function of the gate distinguishes the comparison gating process from a relatively simple recognition match. It must often happen that a spatial location is empty (ie. inactivated in the map) in the sketch percept but not in the corresponding image location or vice versa. In this situation I suggest that a relatively weak image (or percept) signal will be uninhibited (no competing signal) and pass to the integration buffer but remain relatively unstable. This has important consequences for the function of blank spaces (as opposed to crowded ones) in the theory of the sketch. Empty spaces in a sketch have the important function of providing "vacant spaces" for image generation. However the model predicts that the images on such empty spaces will be relatively unstable compared to those locked to sketch contour fragments or other visual components with assigned descriptive labels.
An unresolved problem in the model is that of determining what process controls which parts of the Comparison Buffers are accessible to conscious attention. In the normal perception of impoverished real world stimuli the theory posits that we attend to the contents of the Integrative Buffers and may be unaware of the involuntary image components of our perceptual experiences. Many image attributes of sketches may be equally inseparable attentionally from their stimulus frameworks - for example the three-dimensional volumes often suggested by simple line drawings. Evidence that mental image generation can become automatic with practice is provided by Pani (1982) (See also Chapter 7). However when we consciously superimpose imagined forms upon sketch stimuli, voluntary control is available for the image components at least and it seems likely that we can switch attention at will between the contents of the three representations within the Comparison Buffer. Conscious control of the gate seems less likely (see the binocular rivalry analogy below). At least the triple structure of the proposed spatial Comparison buffer answers Neisser's criticism that the shared resources theory does not explain why we rarely confuse our percepts with our mental images. We do sometimes confuse them, but only when we are attending to the hybrid contents of the Integration buffer.

Although many details about how such Comparison Buffers might work are obviously unclear I claim that the above model has greater explanatory adequacy than simpler shared resources models such of those of Kosslyn, Finke, Farah and others upon whose much quoted work it is built.

5.3.7.3 Plausibility: A Comparison with other Visual Gates.

The proposed gating mechanism with its triple representation buffer is a variation of the "shared resources" hypothesis which (I claim) helps to explain the experimental data on percept-image interaction and to provide a mechanism for percept-image hybrid formation. However its physiological plausibility may be questioned. What other types of high level visual gating does the brain perform and do they illuminate by analogy percept-image interactions? In the above model, it is implied that at least three (or more) specialised modules in the visual cortex, are able to process memory derived signals from their re-entrant fibres simultaneously with stimulus derived signals from the ascending fibres. Because, within a particular processing module, the information from the two sources concerned share the same visual attributes and format (spatial arrays or descriptive trees), they may either compete (causing interference), or reinforce each other (causing facilitation and fusion), depending upon their compatibility. Unfortunately the
downwards (memory derived) component of this competition cannot be directly observed. It is therefore of interest to examine as an analogy how the visual system handles two spatially overlapping, but structurally separate, sources of visual information which, although they are both stimulus derived, provide well documented evidence of a similar combination of compatibility dependent rivalry or facilitative fusion. Two examples demonstrating such a process have been studied experimentally.

In the first example visual stimuli from two different sources are superimposed, using a half-silvered mirror or video imaging techniques. Thus Kolers (1972) designed a headgear with a half-silvered mirror which when he wore it, superimposed the visual world in front of him and the visual world behind him in his binocular field of view. Thus two incompatible scenes of similar contrast and salience compete for attention. Kolers found that he could attend either forwards or backwards with the unattended of the two scenes "disappearing". It was simply not possible to perceive both scenes at once. Similar results have been reported by Neisser & Becklen (1975) who demonstrated the ability to attend selectively to one of two videotaped "games" in full visual overlap. These experiments provide strong evidence for the "late construction" theory of conscious perception and show that a gating mechanism must be available which can selectively suppress from consciousness meaningfully segregated parts of complex scenes.

The other example is the well studied binocular rivalry or binocular fusion which occurs when different images are presented to the left and right eyes. In normal binocular vision the two eyes present slightly different laterally disparate images to the brain which are then fused and the degree of disparity between corresponding points used to compute stereoscopic depth. However retinal fusion is a process distinct from stereoscopic depth perception. Different but similar images presented to the two eyes can be fused to a hybrid image without added depth perception and stereoscopic added depth can be obtained without fusion of the two images (Anderson, B. & Nakayama 1994). Developing a discovery made by Austin, A.L. in 1877 (see Ross 1976), Engel, E. (1958, 1961) studied the effects of presenting two different faces, one to each eye, in a Wheatstone stereoscope using over 100 subjects. For the majority of the subjects a hybrid fused face was perceived. The binocular face was usually described as dissimilar to and more "attractive" than either of its two monocular components. Although it would be rash to push the analogy too far it is possible that further investigation of non-stereoscopic binocular fusion will suggest further mechanisms for the unconscious completion from memory of impoverished stimuli. For example position dependent Image-Percept facilitation suggests a neural mechanism analogous with that
underlying the superiority of binocular over monocular performance in a host of visual tasks (Blake, 1989). According to Blake, it is generally agreed that this binocular superiority reflects the operation of a process of neural summation between the two channels. By the analogy made here neural summation between the stimulus and memory derived channels may account for some of the types of image facilitation of perception and of stimulus facilitation of imagery. Another problem which may be analogous with binocular fusion is that of finding the corresponding parts between stimulus and image which as already indicated may well be a Darwinian *raison d'être* for mental rotation. When binocular matching is assumed to use simple primitives (as in random dot stereograms) then the binocular correspondence problem can only be solved using probabilistic constraints (Marr, 1982). However as Anderson, B. and Nakayama (1994) have pointed out correspondence matching (eliminating the false targets) is much simpler if structured high level primitives are used. In the model proposed here each activated location points to a structured part description and matching these will produce fewer "false targets" than would a correspondence process based on simple luminances or even edge primitives (Marr and Poggio, 1979). Anderson, B. and Nakayama (1994) show that information from half occluded parts (parts visible to one eye only) is used at an early stage of the stereoscopic process suggesting that structured primitives are indeed available for binocular fusion.

When incompatible stimuli are presented to the two eyes, either one image is completely suppressed or a piece-wise rivalry occurs, with patches of the field of view oscillating between dominance for the left eye or the right eye (Panum 1940, Blake 1989). With incompatible, meaningful stimuli, the most expected or attended stimulus can completely suppress that presented to the other eye. For example Engel, E. (1956) showed his subjects a correctly oriented face to one eye and an inverted face to the other. Subjects found the normally oriented face was dominant and typically required several repeated observations to even notice the inverted face. However there is evidence that despite its suppression from awareness, a subliminal stimulus in the currently suppressed eye can still be visually analysed (Walker, 1975). The suppressed stimulus may also cause interference with the dominant field, especially if movement or change in one field causes attention to waver (the "Cheshire Cat" effect, Duensing & Miller, B., 1979). The following is part of a verbatim transcript from one of Engel, E.'s subjects "I see a boy - a boy's face, but there is something else there that I can't see. Something interfering with this". If this analogy is valid, then we would expect incidental, incompatible percepts in sketching systems to similarly interfere with candidate images for sketch hybrids.
With semantically neutral stimuli, such as incompatible grids of oblique lines, the period of dominance of one eye increases with factors such as greater edge sharpness (Kulikowski, 1978), brightness, colour contrast, colour saturation and movement (Levelt, 1965). If by analogy these factors are applied to percept-image rivalry, then it must be expected that in most cases the stimulus derived (percept) signal will be the dominant one. The evidence suggests that when unmatched imagery is concurrent with perception, it is dimmer, less detailed and more blurred than the percept (Kosslyn & Shin, 1994).

Fascinating physiological evidence that binocular rivalry is associated with a neural gate which controls visual awareness has been provided by Logothetis and Schall (1989). These authors trained monkeys to respond differentially to either moving horizontal lines or moving vertical lines. When the horizontal lines were presented to one eye and the vertical lines to the other the monkeys' responses showed the same evidence for visual rivalry that is experienced by humans. The response indicated that they "saw" either vertical or horizontal lines alternately in sequence as one or other eye dominated. Logothetis and Schall then combined this response study with single cell recordings from cells in visual area V5, a topographically mapped part of visual cortex known to respond to direction of movement. Over 200 cells were studied. They found that many horizontal movement cells responded correctly whilst other "vertical" cells were simultaneously responding to the other eye. However there was another set of cells which appeared to respond only to the dominant eye, either horizontal cells to one eye or vertical cells to the other, in a way which correlated with the monkeys' behavioural indication of which type of movement it was actually seeing. If the brain gates visual awareness (see Chapter 7) between competing information from the two eyes, then an analogous process for gating rivalry between images and percepts is at least feasible.

From the analogy with binocular interaction, we can predict that the variety and unexpectedness of the imagery component of sketches will increase with the ambiguity and incompleteness of the sketch percepts. Very sharp high contrast detailed stimuli whilst supporting one image are likely to suppress the formation of others. Faint, blurred or confused stimulus sketch parts are more likely to facilitate new or unexpected imagery.

A final point concerns the neuro-anatomical plausibility of the postulated Comparison Buffers. If, as is suggested, Working Memory contains separate representations of stimulus derived and memory derived images, as well as a shared hybrid image, then why does the neuropsychological literature contain so few examples of patients deficient in a perceptual ability but not the corresponding
imagery ability or vice versa? The neuropsychological inseparability of corresponding perceptual and imaging skills is after all widely quoted as evidence for the single shared resource hypothesis. The analogy with binocular vision answers this question. Corresponding receptive fields of the two eyes are mapped in primary visual cortex in uniform columns of neurons side by side so that the left half of the visual fields of each eye are exactly interleaved in the right occipital cortex and vice versa for the right half of the visual field (Hubel & Wiesel, 1979). In layer 4 the cells in a single 0.5 mm column responding to a small region of the visual field are paired and monocular, that is one responds only to the left eye and one to the right eye. In other layers the cells become binocular, responding equally to the same visual field in both left and right eyes. There are also a series of cell types which respond with varying degrees of dominance for one or other of the two eyes. Thus there is physiological support for physically interleaved, overlapping, computational field maps of the left eye, the right eye and a fuzzy series of fused, hybrid binocular fields. As Cowey (1981) has argued it is likely that the cortex of the brain is organised in such a way that cells which need to communicate with each other frequently, are close together because long axons are biologically "expensive". There are numerous other examples of such interleaved computational maps. Well documented, are those in the cortex of the owl which correlate separately sensed audio frequency and phase lag differences between the two ears in order to locate and map the spatial positions of sound sources (Knudsen et Al., 1987). The idea that corresponding visual attributes from sketch stimulus and memory sources are processed in analogous computational maps in the visual cortex is speculation, of course, but is, I claim, plausible and consistent with known physiology.

5.3.8 Percepts as Hybrid Images: Further Evidence.

There are a number of phenomena in the recorded literature on perception "going beyond what is given", which are usually interpreted as demonstrating the role of expectancy, "set" or unconscious inference on perception. However such explanations seem to me to be deficient in that they address only the question of how the added information is derived and say little about how it is represented. What is the mechanism by which knowledge, inference etc. alter the stimulus derived representation from which conscious experience is derived? If it is possible that hybrid images are routinely generated, unconsciously in perception then much of the traditional literature can be interpreted in terms of the above model. A few examples must suffice.
It is well known that when ambiguous or neutral figures are presented
tachistoscopically for fractions of a second the shape that subjects report or record
can be distorted by the way prior expectancy is manipulated (Carmichael et Al.
1932, Vernon, 1952). Bruner and Postman (1949) asked subjects to identify
ordinary but hand painted playing cards exposed for a fraction of a second in a
tachistoscope. Unknown to the subjects a few of the cards had the incorrect colour
for their suit, black diamonds or hearts, red spades or clubs. The authors recorded
four types of response to the incongruous cards. The first they call a "dominance"
reaction in which the subject reports with assurance that a red six of spades is
either a normal (black) six of spades or is the six of hearts depending on whether
the colour or shape is dominant. Another type of response they call "disruption" in
which the subject fails to resolve the stimulus and does not know what it is
(although at the same exposure time he can identify normal cards). In a third type
of response the incongruity is "recognised". However in a fourth type of response
there is a "compromise". For example, a subject may report that (a) the red six of
spades is either the purple six of spades or the purple six of hearts; (b) the black
four of hearts is reported as a "greyish" four of spades or (c) the red six of clubs is
seen as the "the six of clubs illuminated by red light". The authors conclusions are
simply summarised as "perceptual organisation is powerfully determined by
expectations based upon past commerce with the environment". Whilst this is
undoubtedly true, it is interesting how neatly the results also fit a triple
representation, percept-image gating mechanism, provided it is admitted that
perception can involve automatic imagery completion. By the above model the so-
called "compromises" are hybrid images formed when a weak elicited mental
image and a weak stimulus percept are sufficiently matched to allow fusion, red
plus black generating purple or gray. When the image is stronger than the percept a
dominance reaction is recorded. When the percept is much stronger than the image
the incongruence is noticed. "Disruption" occurs, I suggest when both stimulus and
memory derived image representations are equally strong and cause mutual
interference without time for resolution. Sketches, such as the Goya Procession
(Illustration XXII) would probably cause disruption if presented tachistoscopically.
However for longer perusal, stimulus-memory conflict may be useful, forcing the
cognitive system into imaginative searches for inventive depictive-descriptive
matches.

Experiments demonstrating the influence of knowledge or experience on
perceived colour also fit a hybrid image explanation. Thus Bruner, Postman &
Rodrigues (1951) coloured a tomato, an orange and a banana an identical
intermediate shade of yellow-red. Their subjects were asked to match the colour of
each fruit in turn using an adjustable colour mixing disc also within their field of view. The results showed that the visual matches were clearly biased towards the expected or remembered colours of the fruit.

The hybrid image model also accounts for the perception of the "hard to see" figures such as the Porter's (1954) "Hidden Man" (Illustration XXXII). The head and shoulders of a man is often unrecognised until its outline is pointed out. The old explanation of this perceptual experience is an unconscious inference like change of "set" or expectancy. However this does not explain why the figure changes not only descriptively but spatially as well acquiring an illusory volume that was not seen before. This illusion results I suggest only after a template match between mental image and flat stimulus-percept in the comparison buffer has generated a hybrid-image with depth and volume added from memory.

More controversial is a hybrid image account of spatial perception. It is well known that three-dimensional visual experience cannot be easily explained by the two-dimensional information contained within static, monocular retinal images. Most of the debate has concerned the relative weights to be assigned to information "pick-up" from changing stimuli (Gibson, 1979) or to unconscious hypotheses about the real world (Gregory, 1980). However given that the visual system extracts or infers three-dimensional depth information one way or another, how is this incorporated into the internal "2½D" representation available to visual experience? Experiments conducted long ago by Thouless (1931) indicate that under normal viewing conditions the perceived sizes of objects at varying distances from the eye and the perceived shapes of objects tilted at varying angles to the eye do not remain constant as is implied by the terms size and shape "constancy" but that they "regress" in visual experience away from the stimulus towards their expected real relative dimensions. In one experiment subjects were asked to match the perceived shapes of circles or diamonds viewed at an angle against a chart of possible shapes. Thouless found that perceived shapes were matched to a compromise shape that was intermediate between the known shape (eg. a symmetrical circle or diamond) and the true stimulus a perspectival ellipse or flattened diamond. Thouless termed this result, "phenomenal regression towards the real object". His results also fit a image gating theory of perceptual construction where a flat elliptical stimulus and a circular memory image are fused after gating into an intermediate hybrid.

The relevance of the above hybrid image interpretation of perception to the sketch is as follows. Even if it is accepted that traditional sketch attributes assist invention by acting a specialised templates for superimposed imagery, it does not follow that such support for imagery is always necessary or even desirable. On the
other hand it may be that due to its perceptual origins, imagery by its nature functions best with visual support. If imagery is an automatic component of the perception of real objects, which is primarily evoked when the stimulus is noisy or impoverished, then a knowledge of percept-image interaction will be relevant to all visualising media however far removed they may appear to be from the humble sketch. It does not follow that all the attributes of the sketch are necessarily useful and must be retained in a new medium, nor does it follow that new stimulus attributes in new media as yet untried, will not prove to be powerful visualising tools. It is possible that the three dimensional component of sketches is a consequence of hybrid image formation. Anyone who has been taught life drawing will remember that at first even after much painful effort a drawing will remain obstinately flat. Then one day quite suddenly a simple outline drawing with no shading or other "cues" to depth will acquire illusory volume and solidity. According to the model this experience is due a successful hybrid image match and the sudden passing of an unconscious image by the "guardian" of the gate of the Comparison Buffer.

5.4 Conclusions

5.4.1 Biology of the Sketch.

The above theory makes a narrow, specific claim and a broader but more general claim for the explanatory usefulness of the hybrid image model. The strong claim is simply that the imagined component of artists sketches and the "Image in the Clouds" theory of the viewer's contribution to art (Gombrich, 1962) can be made more specific and more clearly understood by applying recent knowledge of percept-image interaction to the traditional attributes of sketches. We can roughly divide artists' sketches into two classes according to way in which their attributes support mental imagery. Sparse sketches provide fairly unambiguous but incomplete, recognition cues with appropriate spaces for superimposed imagery. An example is Millet's "Curtain of Trees" (Illustration XXIV). (Note that the original contains very faint lines representing branches which do not duplicate well - perfect partners for memory images without rivalry). Noisy sketches on the other hand initiate multiple "recognition by parts" search loops which in turn elicit repeated image-percept match attempts (e.g Goya's "Procession" Illustration XXII).

Some of the reasons why inventive imagery may need perceptual support are summarised below.

1. The attention shifting type of image generation benefits from a visual framework against which imagined patterns can be held.
2. The retrieval of stored information for image generation is facilitated by perceptual object and scene "parts" which act as look-up keys to long term memory.

3. Mental images fade and need continual refreshment. This is facilitated if a visual stimulus "skeleton" is present, upon which they can be superimposed.

4. Unlike percepts, mental images are rarely ambiguous and are hard to re-interpret. Hybrid image theory explains why the stimulus component is necessary to represent multiple alternatives.

5. Mental images are either low frequency (blurred) or contain a smaller range of frequencies to normal percepts. The percept component may be necessary to supply the missing spatial frequencies.

6. The spatial and temporal storage capacity of the resources used for mental imagery are limited and need support from visual stimuli.

The more general claim is that both the advantages and the disadvantages of mental imagery as a visualising tool in the arts, stem from the perceptual origins of imagery. For our hominid ancestors, day to day survival by hunting, plant gathering and avoiding predators was probably more important than long distance planning. The use of visual memory in the visual system evolved primarily to complete confusing percepts. Imagery capacity developed in the brain, from top-down re-entrant signals. These provided stored information, which was used by our hominid ancestors to supplement, rather than to replace, information from the immediately sensed environment. In hunting, for example, glimpses of game would often be of short duration, poorly illuminated and partially concealed or camouflaged. The ability to spatially superimpose memory images upon such a glimpse would have obvious advantages in guiding a spear throwing arm. Gradually mental imagery became separable from visual stimuli to be used for forward planning and visual invention. The image precursors used for perception were automatic and unconscious. In order for imagery to become useful for invention, the brain had also to evolve mechanisms providing voluntary control of imagery. It is possible that the evolution of this ability was facilitated by the Baldwin effect (Richards, R. 1987). This is a postulated Darwinian mechanism by which a learned behaviour pattern advantageous to a species can accelerate the evolution of neural development. Individuals which possessed neural connectivity which made learning to use imagery easier would be favoured by natural selection. In this way the uncontrolled top-down components of perception gradually evolved a secondary function for controlled visual planning. However because the biological time-scale for this development has been brief (Leakey & Lewin, 1992), imagery has never lost its dependence on primarily visual mechanisms and neural
resources. The first steps in the Baldwin effect would have been the observation of stones and marks on the walls of caves which accidentally invited perceptual completion into objects, animals or human figures (Windels, 1949). From these developed the first line drawings, percepts which provoke and later anticipate imagery. According to this theory paleolithic cave paintings are Baldwin effect imagery catalysts. The earliest known sketches are percepts which function as cultural stepping stones in human cognitive evolution. Today despite enormous advances in the technology of creating images, the widely variable human ability to imagine non-existent visual situations still depends for support upon a medium little different from the early charcoal marks on cave walls. One reason for this could be the belief that visual invention needs better physical images. This may not be true. In one sense sketches are not images at all but specialised visual stimuli which support images.

This third function of sketches is summarised below as Hypothesis Three.

5.4.2 Hypothesis Three.

Sketches are percept-memory hybrids. The incomplete, physical attributes of sketches act as a stimulus for percepts which invite completion from memory and imagination. This results in the generation of transient mental images which after a matching process are transformed and spatially superimposed upon an internal representation of the sketch.

The mental construction of such semi depictive percept-memory compounds assists visual invention by two related mechanisms.

1. Sparse or simple sketches act as a "screen" with spaces reserved for projected mental imagery. The physical (light derived) components stabilise memory components and reduce the cognitive effort required for refreshing them. This supports and amplifies those imagined components of an idea which need rapid mental manipulation for refinement.

2. The attributes of ambiguous and crowded sketches elicit a two phased recognition mechanism which results in a stream of temporarily superimposed mental imagery. In the first phase it is proposed that sketch indeterminacies initiate a search loop for position and size invariant stored object (and scene) descriptions (see Chapter 4). In the second phase it is conjectured that interleaved, twined maps of spatial location enable provisional "object hypotheses" to be tested by an alignment match between sketch percept and appropriately scaled and orientated mental images generated from the stored representations of objects and object parts. Depending on the extent of the match, this leads by a gating mechanism to
either image suppression and replacement or to the synthesis of compound percept-image hybrids.
6. VISUAL SEGREGATION AND SELECTIVE ATTENTION.

"Everyone knows what attention is. It is the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought. Focalization, concentration, of consciousness are of its essence. It implies withdrawal from some things in order to deal effectively with others, and is a condition which has its real opposite in the confused, dazed, scatter-brained state which in French is called 'distraction'."

William James. (1890)

6.1 Introduction

Hypothesis Four states that "Sketches amplify inventive thought by isolating and representing separately those attributes of visual experience that are of special relevance to a particular task. This assists the user to attend selectively to a limited part of the task, freeing otherwise shared components of cognitive capacity and reducing the complexity of preparatory visual processing."

Evidence has been produced so far to support the idea that sketch attributes are used -

1. to translate between descriptive and depictive forms of representation in a cycle of additive enrichment,
2. to facilitate multiple access to long visual term memory,
3. to provide perceptual stimuli which provoke and support superimposed mental imagery.

Design or composition by successive refinement, implies the solution of difficult visual problems in stages. As the difficulty and inventiveness of a task increases, it is to be expected that an artist will wish to focus her or his mind selectively on various partial aspects of the total task. Sketches support this need by segregating and representing separately components of the objects in which the artist is interested, before they are later integrated to simulate the appearance of more complete objects.

The use of lines to represent the self-occluding contours (silhouette edges) of objects is a familiar example. In retinal images and photographs lines do not exist at object edges. In the mentally processed construction of visual experience we may be aware of the location of object edges if we attend to them, but they are not normally experienced separately. Rather they have become submerged in the constructed experience of solid objects. For our hominid ancestors, the recognition of objects and their three-dimensional shape was probably more important for survival than the relatively fleeting two-dimensional views presented to the retina. Thus even though the extraction of such contours might be an important stage in
early visual processing (3.4.6), their explicit representation in the final percept is not necessarily biologically useful. However artists and designers who need to imagine and manipulate view dependent object shapes and their relationships, must learn to attend to the very information that the visual system has extracted and used, but suppressed from consciousness. It has already been mentioned (3.4.6) that some psychologists believe that it is not by accident that outline drawings resemble the output of the postulated early stages of human vision (Frisby, 1979). Thus it is worth considering whether one of the functions of linear outlines in sketches is to preserve for selective thought some of the results of such unattended early stages of visual processing.

The limited representational scope of sparse sketches facilitates the concentration of visual thought on one part of a problem at a time. Later of course such separated partial solutions will have to be integrated and re-assessed within a fuller context. However the advantages of this segregation are not always obvious. It is likely that most line sketches are used to think about form without (temporarily at least) worrying about colour. However, it is also possible that line drawings are used as convenient spatial frameworks on which to mentally manipulate imagined colour. For example because working line drawings represent neither surface tonality nor hue, it cannot be assumed that their creator is not also thinking about (and imagining) colour. Turner's spidery line drawings covered with written notes about colour show this (Wilkinson, 1975). We cannot resolve such questions by reading an artist's mind. It is necessary again to compare visible attributes of sketches with existing theories of visual processing and ask the question "Do the segregated, visually incomplete attributes of sketches really facilitate focused visual thought or are they simply an unhelpful consequence of the needs for convenience and speed?". Two a priori considerations suggest that in some circumstances at least, they do indeed support selective thought.

The first consideration is that according to the late construction theory of perception (4.2.6) much of the visual analysis which our brains perform is not directly available to visual experience. Unstructured patches of colour and luminance are presented to the retina complete with perspective distortion. They are experienced as visual objects separated from their backgrounds with the perspective replaced, or partly replaced, by their true sizes and shapes with depth and three dimensional shape added. In addition luminance and chromatic variations are transformed and segregated into reflectance colours, firmly attached to surfaces, with separately perceived illumination bathing those surfaces (4.2.1). This is the information that our hominid ancestors needed from their visual systems. However, artists who may wish to manipulate new visual experience or
designers who must manipulate the appearance of an imaginary object, need access to those segregated components of images which the brain's visual system normally conceals. As already suggested, the history of human image making from the discovery of perspective to today's chemical and electronic media has been largely devoted to the creation of artificial stimuli that mimic or capture the unstructured light pattern presented to the retina by real objects. When realism is experienced it is a product of the human visual system and is not explicit in the image (2.3 & Gombrich, 1962). The visual system evolved to meet the needs of our hominid ancestors, not the needs of contemporary artists. For example, in order to generate and manipulate illusory space by perspective, it is necessary to perceive and manipulate the projected distortions on the retina which the visual system suppresses. To illustrate to others the intended effects of illumination upon a coloured surface, painters must learn to consciously perceive and manipulate that segregation of image luminances into perceived surface and incident illumination which is normally assigned to the pre-conscious brain. Thus there is a need in the visual arts to make commonly unattended products of vision available for conscious control and selective thought. This is an effortful skill that must be learned (see Pratt, 1984 and 3.4.6).

The second consideration (discussed below) is that taken as a class across different artists and styles, working sketches can be sorted into categories which mirror to a remarkable degree what is known about the modular division of labour within visual system. If this is not accidental, then it suggests at least that the attribute segregation found in sketches, shares some functional purpose with perceptual processes. However the early division of labour in the visual pathway is believed to be automatic and with different processes operating in parallel (Treisman, 1986). It appears to occur quickly and effortlessly in perception before the separately analysed attributes are reconstructed into the conscious experience of real objects. Therefore we must not make the mistake of assuming that the function of separated visual attributes in the sketch is simply to by-pass or simplify these early perceptual analyses. It seems more likely that their function is to make perceptual attributes that are normally unattended salient and more easily available for conscious manipulation. Whether the attributes emphasised in a particular sketch are derived directly from memory or from the effortful examination of real objects, sketches, unlike the images of "realistic" media, record such selective attributes for mental manipulation. Otherwise they would have to be continually re-extracted every time they are examined.
6.2 Objects and their Attributes

The fundamental end products of conscious perception are objects and events. The traditional media of sketches do not support temporal change (movement) so event structure is not represented. However it is likely that future sketching systems will allow event representation. It is therefore worth noting one of the most obvious forms of perceptual segregation - the freezing of a moment or a series of moments in time. The normal percept of an object, even a static one, as Gibson has emphasised, is constructed from a moving flow of retinal images. The very fact that a sketch (or a finished work of art) does not move allows a kind of contemplative visual analysis that would otherwise be impossible. On the other hand, for film and many kinds of object design (such as Fashion) visualising movement is crucial and, according to claims outlined in this chapter, new sketching media (eg inter-active video) will require new ways of representing movement extracted from object variability.

The perceived attributes which objects possess can be classified in many different ways. Treisman (1986) has reviewed in detail attempts to relate such classifications to experimental evidence about perceptual analysis or decomposition. Her terminology will be respected here. Any decomposed product of perceptual analysis, conscious or pre-conscious will be referred to as an attribute. Object attributes are further divided into three kinds, dimensions, features and parts.

6.2.1 Dimensions

"A dimension can be defined as a set of discrete or continuously varying values of which only one can and must characterise any non-composite stimulus to which the dimension applies (i.e. it is in the relevant sense modality and at the right level of description)" (Treisman 1986 page 35-3). Treisman distinguishes simple dimensions which cannot be further decomposed from "emergent" dimensions which are themselves derived from simpler dimensions. However it is not at all clear what are the rules by which simple dimensions can be combined to make new emergent dimensions which are "psychologically real". Dimensions such as those that characterise shapes commonly occur in nested hierarchies. For example points have locations and intensities, lines have in addition orientation and length, surfaces have, in addition to these, shape, area and texture. Some dimensions, such as those of colour vary continuously, others (according to this definition) may comprise a large number of discrete values.
6.2.2 Features

Sometimes objects appear to differ in some binary attribute (such as open or closed contour), or they either possess or do not possess a particular attribute. Such attributes are here termed features. Treisman makes the point that continuous dimensions of physical stimuli may be categorised by the visual system into discontinuous domains, constituting features in perception. An important example is the perception of the "unique" colour sensations red, green, blue, yellow (Hering, 1964). These are often treated as if they were dimensions, but as Hering pointed out they are really features (Hering, 1964). Hue dimensions, which are the "emergent" perceptual constructions from the wavelength dimension of the stimulus, are derived from varying mixtures of these features, red-blue, blue-green, green-yellow, yellow-red as described in the Natural Colour System (Swedish Standards Institution, N.C.S. Colour Atlas 1978). The failure to recognise how little colour experience may correlate with the measured attributes of the physical stimulus is sometimes called the "stimulus error". The history of the stimulus error in the colour theories of artists is discussed by Fish (1981). The point made in that paper, relevant to this Chapter, is that whilst artists are interested in attributes of the perceived colour, in traditional media they are often forced to manipulate (and hence attend to) non-matching attributes of the stimulus. The "stimulus error" is embodied in natural language when we frequently refer to colour experience by the names of pigments. I suggest that having successfully identified the perceived colour attribute of interest, a different notation, in a sketching system, unrelated to the stimulus would assist visual thought.

6.2.3 Parts

Parts are the spatially localised or articulated components from which most objects are constructed. Parts differ from dimensional values, in that parts are neither mutually exclusive nor obligatory. (A perceived chair must have a size but can optionally possess arms). Object parts are usually articulated according to certain topological rules (arms and legs of the body) or located in specified spatial relationships (eyes, nose and mouth on a face). A point important to sketch function that Treisman omits, is that, in picture perception especially (and probably all perception), the dividing line between what constitutes an object and what constitutes a part has a variable locus in the object-part hierarchy. Where this locus lies depends upon the current scope of attention. When a painter composes a landscape the whole scene is perceived as a single object, in a portrait the object is the head, to an oculist the eye etc. The importance of what constitutes a whole object in a sketch lies in the evidence that when attending to object detail it is
6. Visual Segregation and Selective Attention

easier to make simultaneous judgements about two attributes of a single object than it is to make simultaneous judgements about two objects (Duncan 1985, Kahneman & Treisman, 1984). It is suggested here that one function of the sketch is to declare (avoiding the attentional conflict present in real scenes) the object-part relationships appropriate to specific tasks.

6.3 Integral and Separable Dimensions

Amongst the many visual tasks that sketches are used for, we can safely include that of making judgements of similarity and difference (studying relationships) between specific attributes of colour or spatial structure in an imagined scene or object. Even if it is possible to make such judgements by attending to the attribute of interest when it is combined with other irrelevant attributes, there are grounds for believing that the irrelevant attributes will make the task harder (see next section). Segregation of such attributes in a sketch could then be regarded as helpful but not essential. However if it is found that the perceptual system combines some attributes in a holistic way so that one attribute alters judgements made about another, then segregation for accurate visual comparison becomes a necessity rather than a luxury.

Shepard (1964) noted that some physically distinct unitary dimensions interacted perceptually in just such a manner. Similarity judgements on such dimensions fit a Euclidean metric. This means that judgements of similarity when two objects differ in two such dimensions is proportional to the root mean squares of the differences on each of the individual dimensions (the shortest distance between two points in Cartesian coordinates). In contrast analysable dimensions contribute independently to judgements of similarity and can be represented by a city-block metric, (i.e. with no diagonal short-cut between them.)

Garner (1974,1978) has elaborated the distinction further. He distinguishes integral dimensions (Shepard’s unitary stimuli) from separable dimensions and also from what he calls configurable dimensions. Garner’s integral dimensions interact so that performance is improved when the dimensions are perfectly correlated and unimpeded selective attention to either is impossible, causing interference when the dimensions vary orthogonally. Separable dimensions can be attended to separately but judgements of similarity are not facilitated when both dimensions covary. As a measure of the separability between two dimensions Garner used a speeded classification task. The mean time taken by subjects to sort stimulus cards according discriminatory levels of a specified dimension was measured in three conditions. Either the stimuli varied in one dimension only or both the specified dimension and the irrelevant dimension covaried and were correlated or the two dimensions varied orthogonally (i.e. each level of one
dimension was combined with each level of the other dimension). If the stimuli were **separable** (it was reasoned) the sorting times would be equal in all three conditions. On the other hand if they were **integral** sorting times with the correlated irrelevant dimension should be shorter (facilitated) than with no irrelevant dimension but should be longer (hindered) in the orthogonal condition. Using this criterion (Garner & Felfoldy 1970, Felfoldy & Garner 1971) found that Munsell colour Value and Chroma (Lightness and Saturation) and colour Hue and Value were integral dimensions. They also produced the expected Euclidean metric in dimensional scaling measures. Also integral were the horizontal and vertical positions of a dot. On the other hand the size of a circle and the angle of a line drawn across its diameter were separable. Using the same criterion Gottwald and Garner (1975) **claimed that colour and form were separable.** The orthogonal condition is described as a "filtering" task because the subject attempts to filter out variation in the irrelevant dimension. This finding has implications for sketch function, since any non correlated variation in an irrelevant integral dimension can be expected to require filtering and hence to cause interference. For example irrelevant, uncontrolled variation in chroma or hue can be expected to interfere with judgements of colour brightness ("tone value" or "lightness") in a chromatic sketch. However, the inverse conclusion for separable dimensions, that for example comparisons of form should not be hindered by irrelevant colour variation, must be qualified.

The criteria for dimensional separability depend upon the task. For example, it is found that the ability to filter out an irrelevant dimension in the speeded classification task, is not always repeated in simultaneous comparison (same-different) tasks. Now although sorting is relevant to sequential sketch comparisons, simultaneous dimensional comparisons within a sketch are of greater interest. For example despite the separability found by Gottwald and Garner for colour and form, they have been shown to show dimensional interference in a simultaneous comparison task (Egeth, 1966, Hawkins & Shigley, 1972). Santee and Egeth (1980) found that when subjects were required to make speeded classifications in the Garner conditions, they were able to attend selectively to form when the irrelevant dimension of size or shading was varied. However subjects were **not** able to filter out irrelevant size or shading disparity in a comparably designed simultaneous comparison task. In another experiment they demonstrated a monotonic increase in "same" reaction times as a function of irrelevant disparity in a successive comparison task. The authors conclude that dimensions which are separable by the Garner criteria, can still show mutual interference in simultaneous matching. This shows that the ability to attend selectively to a particular stimulus
dimension is not simply a function of the particular dimensions employed, but is also a function of the particular task and processing strategies required of the subject.

6.4 Normalisation and Visual Comparison

The evidence that the time to match random two dimensional shapes is directly related to irrelevant differences of size or orientation, has been discussed in chapter 4 (4.3.6, Bundesen & Larsen 1975, Bundesen et Al. 1981). So also has the well known evidence that the time to match perspective views of three-d shapes is linearly related to the angular difference in orientation (5.2.6 Shepard & Cooper, L., 1982). The common explanation of these results is that different values of an irrelevant dimension of the internal representations of the comparison stimuli have to be normalised (or equated), by a time consuming transformation process, before the same-different judgement can be made. It was suggested that this was evidence for a "template" or spatial array type of matching process, one of the many processes available for object recognition in sparse sketches (4.3.6). Dixon and Just (1978) have suggested that such a necessary normalisation of irrelevant attributes may be a principle which can be used to explain interference in a much wider range of selective "filtering" tasks from Garner's orthogonal "speeded classification" to mental rotation and other types simultaneous comparison. "We suggest that irrelevant disparity may produce these various interference phenomena because subjects mentally equate the two stimuli on the irrelevant dimension before deciding that they are identical along the relevant dimension." The authors investigated the time to make "same" judgements about the height of two ellipses with width as the irrelevant dimension and the time to make hue judgements with tint as the irrelevant dimension. Their results supported their normalisation hypothesis. However Santee & Egeth's results mentioned above shows that the suggestion that speeded classification and simultaneous comparison involves a common process cannot be true. However these authors agree that their results support normalisation of irrelevant dimensions even when: (1) the levels of the irrelevant dimension may be accessed easily through a set of encoding features (such as the names of geometric figures) and (2) the stimuli are presented in a successive comparison paradigm.

The important conclusion is that, in sketches, eliminating variation along irrelevant dimensions should facilitate several types of visual judgement about a dimension of interest. There are two types of reason for this advantage. Either the relevant and irrelevant dimensions are integral making selective attention to one difficult (Garner implies "impossible" but the results do not support this extreme conclusion) or the irrelevant dimension must be mentally normalised before a
judgement is possible. Whether a particular irrelevant dimension will cause interference depends not only upon the dimensions concerned but also upon the task. Simultaneous similarity judgements within a sketch are particularly susceptible to "filtering" interference.

It is of course quite true that in the visual arts there may be few if any "irrelevant" visual dimensions. For example, in painting, colour and form are usually equally important and if they interact then it is especially important to visualise the projected object with both attributes present. This is of course true of the final synthesis. However it is argued here that in the early inventive stages of visual design, the need to attend to difficult parts of a task singly and the need to preserve options in some dimensions makes those dimensions temporarily irrelevant. Evidence in favour of this position can be found in traditional working practices.

6.5 Theories of Selective Visual Attention

The sketch is both a product of and a support for visual attention. The ability to concentrate the mind on a difficult component part of a visual problem implies the ability to mentally manipulate, analyse and re-interpret information which has already been perceived or imagined. The need to attend selectively is undoubtedly part of this problem.

Selective attention is distinguished in the literature from divided attention, by the task. In experiments on divided attention two (or more) tasks must be attended to at once and the interference which occurs is often used to provide evidence about whether the two tasks compete for shared resources in the brain (5.3.2 & 5.3.3). With selective attention the task is to filter out, ignore or avoid interference from irrelevant or distracting information. The remarkable but imperfect capacity of the brain to focus visual processing on a limited part of the information provided by the eyes (perception) or by memory (imagery), has another obvious function. Many of the neural resources used by the visual system have limited capacity. Therefore it is necessary to have a mechanism which can select important or interesting parts of the information in the sketch hybrid image for detailed processing by those resources which are in short supply. Attention is also important as a mechanism which allocates information for voluntary control and "conscious awareness". What this means is considered separately in Chapter 7.

I believe that sketches have special attributes which support such focused attention at different processing levels within the brain and that, in this respect at least, they are superior to photographs and other realistic images.

There has been an on-going debate about whether visual attention is based primarily on
1. Visual attribute discrimination (e.g. Allport, 1971, 1980)

2. Spatial selection based on processing a single limited part of the visual field at one time (e.g. Hoffman, J. and Nelson 1981, Posner 1980, Posner et Al. 1982, Tsal, 1983)

3. Selection based on objects, with objects defined hierarchically as discussed above (e.g. Neisser, 1967, Treisman and Gelade 1980, Kahneman & Treisman, 1984, Duncan, 1985)

Currently the evidence seems to support selection based on all three of these aspects of vision but in different contexts and for different purposes. (For reviews see Humphreys and Bruce 1989, Kahneman & Treisman, 1984, Duncan 1985, Kosslyn and Koenig, 1992, Posner et Al. 1994).

A visual attribute theory of attention has been suggested by Allport based on the idea of a number of analysers working in parallel. If two discriminations depend upon the same analyser then mutual interference occurs, otherwise attention can be divided between the attributes. Allport's theory is actually not incompatible with either spatial or object based theories of attention. The problem of selectively attending to attributes and their relevance to the sketch has been discussed already (6.3).

To understand the relevance of spatial and object based theories of attention to the sketch and how they may be mutually supportive it is helpful to consider the functions of different attention mechanisms. The fovea is the only part of the retina capable of high resolution. Overt visual attention involves movements of the eyes, usually in saccadic jumps, in order to target interesting, but necessarily limited, parts of the visual field upon the fovea. Covert orienting of attention occurs when limited parts of the visual field receive preferential "focused" treatment without eye movements, drawing upon an internal scanning mechanism with limited resources, the "window" of attention (Kosslyn, 1994 233-234). Such "orienting" of spatial attention can be selective not only to direction within the visual field but also to different perceived distances within a given direction (Downing and Pinker 1985). Object recognition and detailed analysis appears to occur only after such a focussing of resources. Logically an efficient attentional mechanism demands two quite different processes. Given that the whole of the visual field cannot be processed in depth simultaneously, the nervous system provides a mechanism which can perform just enough processing in parallel over a wide field to select a smaller region of interest for identification and further more detailed analysis. The visual system is capable of directing not only overt movements of the eyes to a selected spot, but also covert movements of focal attention across an internal map of the visual field. But when visual capacity has been thus focused a different
attentional capacity is required which can analyse in greater depth objects of interest and can also control scanning movements of attention for information extraction. Over many years, evidence has accumulated for two such attentional systems. A fast low resolution, automatic system is driven by the periphery of the visual field. It selects without identification, targets for further processing. Then there is a slower high resolution system which allows inspection and recognition of objects within a limited spatial region and which can be moved with voluntary control. Convergent evidence from several sources suggests two corresponding channels controlling input information - a sustained and a transient channel. During a saccade initiated by the transient channel, the sustained channel is inhibited to prevent overlapping signal interference from the slower sustained channel (Breitmeyer & Ganz, 1976). These two channels correspond partly with the magno and parvo cellular pathways which later in the processing sequence segregate movement, depth and perhaps figure-ground processing, from colour processing and high resolution object detail (5.3.5, Hubel & Livingstone 1987, Livingstone & Hubel, 1987). There is evidence that there are also two separable modes of control over such movements of the "mind's eye" (Posner & Boies 1971, Posner & Snyder 1975, Jonides, 1980 1981) a fast automatic exogenous system and an endogenous one which can be voluntarily controlled.

Recently separate anatomical attention networks in the brain which fit this distinction have been identified (Posner & Rothbart 1992). The posterior attention system involves portions of the parietal cortex (also associated with visual neglect - see 5.3.2) and parts of the thalamus and mid-brain, known to be used in the control of eye movements. This system is involved in directing attention to relevant locations as in visual search, in binding information to spatial locations to produce object perception and in selecting a relevant scale for examining visual input.

The anterior attention network involves areas of the mid prefrontal cortex that together appear to be active in a wide variety of situations involved in the detection of events. In experiments with humans this detection is often signalled by a verbal response. However this is simply one form of evidence that, in contrast to the posterior system, the anterior system is associated with conscious experience of the information attended to. Thus the anterior system is thought by (Posner at least) to be responsible for conscious awareness of an attended object or location. As studies of visual neglect have shown, awareness of a signal depends upon the posterior system (parietal cortex) which must feed attended information to the anterior system (The two systems are richly interconnected). However it not the posterior system itself which produces awareness. This is demonstrated by studies which dissociate the act of orienting to signals from their detection. Thus it
is after attention has been engaged by the posterior system to a location of interest in a sketch, that the anterior system is activated in a mode which supports awareness and is necessary for the discrimination of colour, form and meaning (Petersen et al., 1989). The distinction between automatic and voluntary is discussed further in Chapter 7 (7.3). The point made here is that the functional separation between automatic and voluntary modes of processing may be an important reason for segregating sketch attributes. Even if the anterior system allows the sketch user to discriminate between relevant and irrelevant attributes in unsegregated images, the posterior system can cause involuntary processing and movements of attention of which he/she may not be aware.

Posner has used the metaphor of a spot light with a "zoom" lens for guiding research into spatial aspects of attention. Movements of attention define sequential locations of limited area within the visual field (or an imagined field) to which object attributes can become bound. As with a spot light the area selected is preferentially activated for further processing. It is likely that this provides the mechanism by which the links between stored spatial structure (the "where" system) and stored object attributes (the "What" system) are established in sketch hybrid images (see 5.3.5). There is evidence that in addition the system can control the size of the attended field (the zoom lens analogy) at the expense of spatial resolution. Thus a wide attended field has a low spatial resolution whilst a smaller attended field (1 degree or less) supports a higher resolution. Watt (1988) and Kosslyn et al. (1992) discuss the advantages of a system which can use, in sequence, a fast low resolution wide field information followed by higher resolution information from a more limited field, for analysing position and spatial structure. Only one location can be attended to at a time. Not only is information at that location processed in greater depth but information at other locations is suppressed (Posner, 1992).

It is interesting to consider an implication of Posner's account of the Posterior Attention Network for the percept-image hybrid theory of sketch structure presented in Chapter 5 (5.3.7). When the "spot-light" of attention moves covertly over a mental image or to a peripheral part of the field of view as the result of a foveal cue (as in some of Posner's experiments) it seems inevitable that the spatial sensitivity of such "mind's eye" movement will be handicapped by the low resolution and poverty of the framework against which the movement must occur. It is possible that the perceptual skeleton of the sketch hybrid image provides the necessary landmarks for accurate overt eye movements, which in turn ensures more accurate "mind's eye" movements over mentally superimposed image components. In fact Brandt et al. (1989) provide evidence that subjects move their
eyes when scanning mental images internally. Thus skeletal sketch percepts may provide a mechanism for harnessing overt and more accurately guided movements of attention to mental images (as in Illustration XXIV).

Object based theories of attention have been supported by Neisser (1967), Treisman (Treisman & Gelade, 1980, Treisman & Schmidt 1982, Kahneman & Treisman 1984) and Duncan (1985). In outline these theories agree that perceptual analysis of a visual scene takes place in two successive stages. According to Neisser (1967) the first "pre-attentive" stage analyses aspects of the visual field segmenting it into candidate objects on the basis of Gestalt laws of grouping, figure-ground separation etc. The second stage, focal attention, analyses an object in more detail and attempts a recognition match. The pre-attentive stage occurs in parallel across several objects in the whole visual field, whilst focal attention is limited to a single object and region at one time and can only analyse in sequence objects at different locations. According to Treisman’s feature integration theory (Treisman & Gelade, 1980), the early pre-attentive stage consists of a number of parallel processes which extract simple features of objects such as orientation, colour, size and shape. In Treisman's model, these features are represented by separate maps, presumably corresponding to specialised modules in the visual cortex. Whilst individual features can be detected in parallel, conjunctions of features must be searched for sequentially. The second stage consists of sequential attention to objects at different locations with the function of binding separately detected features to specific locations. Treisman suggests that focal attention opens an object "file", which can be used for comparison with stored object descriptions, for recognition. There are problems with the details of Treisman's theory. For example there is evidence that more than simple features can be detected in parallel or without focal attention. Humphreys et Al. (1989) have shown conjunctions of form elements can be detected in parallel when the distractors are homogeneous. They suggest that whether or not sequential focal attention is needed to detect conjunctions of features depends upon the way in which background objects are grouped.

In support of object based theories, it has been shown that when subjects are given a task which requires attention to one of two overlapping objects they have difficulty reporting or remembering attributes of the un-attended object even though both objects occupy the same spatial field (Rock & Gutman, 1981). Duncan (1985) used brief, small foveal (1 degree) displays, consisting of two overlapping objects. He showed that two judgements concerning the same object can be made without loss of accuracy, whereas judgements which concerned two different objects cannot. He showed elegantly that neither the similarity nor the difficulty of
the required discriminations, nor the spatial grouping of the relevant information could account for the results. It appears that focal attention deals, if not with whole objects, then at least with chunks of information grouped by "Gestalt like" processing into candidate objects. Duncan concludes, modestly, that the ability to attend simultaneously to several sources of information, is at least in part a limit on the number of separate objects that can be seen.

Also relevant to the problems of attending to objects in sketches is the discovery of negative priming effects by Tipper et al. (Tipper 1985, Tipper & Driver, 1988). These authors showed their subjects briefly displayed line drawings of two common objects superimposed upon each other in different colours (red and green). They measured (by successive increments) the minimum time needed by their subjects to correctly name the red objects. Subjects were instructed to ignore the green objects. Their results showed that this procedure not only facilitated the later naming of pictures of objects of the same category as the attended objects (ie showed "priming") but that it also interfered with the naming of the ignored picture categories ("negative priming"). In a final test they showed blank cards immediately after the red and green test pictures and asked subjects the surprise question "what category was the green (ignored) object?". Subjects' responses were only correct at chance level, showing they were not aware of the ignored object category. The authors explain their results by a spreading inhibition of response at the level of a symbolic, stored representation of object meaning. Whatever the explanation of these results the fact that such negative priming for unattended parts of complex images can occur is another reason for segregating components of a complex visual task. It implies that although it is possible to attend selectively to some components within a complex, or detailed sketch whilst ignoring others, the ignored components may limit, without our awareness or control, the flexibility and inventiveness with which we may wish to re-interpret those same components at a second viewing.

The relevance of object based theories of attention to sketch attributes also lies in the implication that some visual processing of attributes which are perceived as belonging to the same object is obligatory. Interference with selective attention to the attributes of interest would then be unavoidable due to the impossibility of filtering out distracting or irrelevant attributes of the attended object. Evidence for attentional interference from the results of such automatic object processing is discussed in Kahneman & Treisman (1984).

The well investigated "Stroop" task is often cited to illustrate automaticity of certain visual processes. The subject is asked to identify the colour of the ink in which a word is printed. If the meaning of the word is inconsistent with the colour
6. Visual Segregation and Selective Attention

of the ink (i.e. the word "RED" printed in green ink) then naming the colour of the ink is delayed compared to when the word is consistent with the actual colour or has no colour meaning. Petersen et al., (1989) have shown with PET studies that in the inconsistent condition Stroop interference activates the Anterior (i.e. conscious) Attentional System in the brain more than in the non-interfering conditions. However much subjects may practise ignoring the irrelevant meaning of a word, it continues to interfere with their response. The shape of the word is believed to quickly and automatically activate the node representing its stored meaning before the rival colour meaning node can be accessed by the perceived colour. Stroop type conflict has been demonstrated in a wide variety of other tasks. For example, Paivio (1975) measured the times to decide which of two objects or animals was the larger in real life size. In the congruent condition subjects were shown sketches of the two objects (e.g. a zebra and a table lamp) with the larger object depicted as larger. In the incongruent condition the larger object is depicted smaller than the other. The results showed that decision times were slower for the incongruent than for the congruent pairs. The involuntary processing of irrelevant perceptual information makes it harder to make judgements based on memory. Paivio (1986) quotes these and similar experiments to support his dual coding theory of memory representation. However the results also suggest an involuntary mode in the hybrid image model of sketch perception I developed in Chapter 5. In the incongruent condition involuntary scaling of generated images to fit the sketch percepts would distort the relative sizes of the memory images. (This is not Paivio's suggestion but my own.)

Kahneman & Treisman (1984) quote other experiments which show that Stroop interference is greatly reduced when the colour and the word are separated and not seen to belong to the same object. The authors quote a number of further experiments which also demonstrate the relative ease of filtering out irrelevant objects and the difficulty of filtering out parts or properties of an attended object. They state "Attention to irrelevant parts of objects is obligatory" and that "The major conclusion is that it is essential to distinguish selection of inputs or objects from selection of properties. As we have seen observers are capable of efficient rejection of irrelevant objects, but the irrelevant properties (and perhaps parts) of an attended object cannot be prevented from contacting their nodes and from activating irrelevant responses. The distinction between selection of objects and selection of properties (or analysers) seems salient and fundamental; yet is often ignored in psychological research and theory".

This quotation is given because of the bearing it has on my suggestion (6.2.3) that sketch attributes declare the object structure. I mean by this that the
Gestalt laws of grouping and the figure-ground rules are used explicitly or intuitively to emphasise, for selective attention, a particular structure in the object-part hierarchy. For example the Cezanne landscape sketch (Illustration XVIII) demonstrates the linking by similarity of separated curved and angular parts of trees and houses. One of the several functions of such drastically segregated visual rhyming is, I suggest, to declare to the automatic attentional processes of the posterior neural network, that the whole landscape is to be considered as a single object. This according to object based theories of attention should facilitate the simultaneous attention to the parts of different trees, rocks and houses which are at this stage not declared as separate objects. At the same time it interferes with the selective attention to individual trees, houses etc., which was, presumably, of less importance to Cezanne at this stage of his thoughts. In contrast the two faces of Cezanne's son (Illustration XIX) are obviously "declared" as separate objects. In many sketches of course the object-part structure is more ambiguous and depends upon the mental bias of the sketch user.

Another way of thinking about the visual "rhyming" of colour dimensions or components of form is to treat it as a kind of normalisation. It would then be tempting to view the normalisation of irrelevant visual dimensions and the difficulty of filtering irrelevant integral dimensions, as a consequence of the same pre-attentive processes which group parts into objects. The prediction which would follow is that, if we want to compare corresponding features in similar objects, then we should remove or equate ("rhyme") corresponding irrelevant features in our sketches. Duncan's results with foveally imaged boxes and lines suggest that this is so. However this conclusion would be premature. We need more research into visual comparisons between the kind of object parts and features that are typical of representational sketches. In general the theory of object based attention seems to reinforce the idea that, whenever an artist needs to attend selectively to a component attribute of an object, irrelevant attributes will involuntarily hinder visual thought.

The capacity to generate representations of segregated visual attributes in a sketch shows that, with skill, selective attention to such attributes in natural objects is possible, although as we have seen, there is usually a cognitive penalty. I believe that the parallels between the attentional division of labour found in, for example, a series of sketches of the same subject and the kinds of modular division of labour recently discovered within the visual system, is no coincidence (see Livingstone 1988). Cowey (1985) has even suggested that one of the biological reasons for having many modular visual regions in the cortex may be to facilitate attentional
6. Visual Segregation and Selective Attention

switching between visual attributes. Some of these parallels will now be briefly considered.

6.6 Separating Colour and Form.

Perhaps the most obvious characteristic of sketches as drawings is that, like the primate visual system, they commonly separate the representation of colour and form.

The word "colour" is used ambiguously in common speech (and by some scientists !) to either include or reject black, grey and white. ("Is that a colour film?" answer "No its black and white", but "What colour is coal?" answer "Black"). Here, to avoid this confusion following the usage of Colour Science, the meaning of the word "colour" will include the hueless monochrome colours and the terms Chromatic and Achromatic used to distinguish colour experiences containing hue (redness, greenness, yellowness, blueness etc.) and optionally blackness and whiteness, from those which only contain varying proportions of black, white and grey (Hering 1964). There is a debate about how many dimensions of perceived colour there are (Evans, R. 1974, Hunt 1977, Fish 1981). There is agreement that there are at least three dimensions of variability for surface colours under constant illumination with a fourth variable (Brightness) related to changes in illumination. However some authorities believe that there is a fifth dimension - "Brilliance", (Evans, R. 1974) or "Colourfulness" (Hunt 1977). (For discussion of the relevance to art see Fish, 1981). However in addition colours have "modes of appearance", opacity, surface, volume, lustre, mattness shininess etc. (Katz 1935, Beck 1972). The point is that to the normally sighted all real surfaces must have a colour and this must have a value at least of lightness (or black-whiteness) and a brightness (related to how strongly it is perceived to be illuminated) and it must have a mode of appearance. Only values on the dimensions related to its chromaticity such as Munsell Hue and Chroma are optional. Therefore it is the cultural familiarity of drawings which conceals from us the intrinsic oddness of a representational system for visual surfaces and three-dimensional form, which offers no values for any of these colour dimensions or attributes (Rawson, 1969). Three functions for the contour lines in drawings have already been suggested. They act as intermediate descriptive-depictive signs to catalyse the mental translation of representational type. They highlight fragments of object contours which act as look up keys to memory by "recognition by parts". They provide skeletal frame works which provoke and support superimposed imagery. Now a fourth function emerges. By representing only the edges of surfaces or silhouette fragments of objects they avoid all representation of the illumination and reflectance of surfaces. Yet still they describe form.
Of course, as the hybrid-image theory of the sketch suggests, outline drawings may have imagined colour superimposed in spatial register with them. However such colour thoughts are likely to be optional or fleeting and unlikely to cause the type of involuntary competition discussed above (6.5). Watkins and Schiano (1982) have investigated this voluntary super-imposition of imagined colour onto black and white outline shapes. They found that such imagined colours can facilitate or hinder the later recognition of the same shapes presented with real colours. The implication is that colours, although only imagined on a sketched object, will be stored with that object in long term memory and may influence the later ability to recognise and to think about similar object representations actually coloured.

The inverse segregation can occur in colour sketches where any representation of form is vague and even the spatial location of colours is hard to determine. Turner's justly famous watercolour sketches furnish many examples (Illustration XXXV).

Often the segregation between colour and form occurs only in the temporal succession of sketching acts so that it is not apparent in the finished sketch. Thus it is common practice to manipulate spatial structure by drawing before later adding colour. As Livingstone (1988) has pointed out, colour added to a drawing need only roughly fit surface outlines yet still appear to adhere to perceived surfaces. She points out that this is a predictable consequence of the fact that the parvocellular colour pathway in the cortex (e.g in the cytochrome oxidase blobs in V1) has a lower spatial resolution than the parallel form processing path from the interblob regions of area V1.

Sometimes the opposite temporal segregation is used. Thus the painter Dufy frequently mapped out his first ideas in spatially vague areas of colour, later drawing on top (Werner, 1987).

Dufy believed in the value of attending separately to colour and form. "Once when I was sketching by the harbour I noticed a bright, red blob moving in a line across my field of vision. Only later did I realise it was a small girl in a red dress .. I have ever since been convinced of the possibility of thinking separately about colour and form" ("Dufy" Channel 4 Television documentary 1982).

To an artist or designer the fine adjustment and mental manipulation of spatial and chromatic relationships are each very demanding tasks. They may be facilitated in the early stages of visual composition by attending to them separately. The working methods of many artists and designers suggest that they know this intuitively. Ruskin in his advice to students of drawing put it in the following words
"You may, in the time which other vocations leave at your disposal, produce finished, beautiful and masterly drawings in light and shade. But to colour well requires your life. It cannot be done cheaper. The difficulty of doing right is increased - not twofold nor threefold, but a thousandfold, and more - by the addition of colour to your work. For the chances are more than a thousand to one against your being right both in form and in colour with a given touch: it is difficult enough to be right in form, if you attend to that only; but when you have to attend, at the same moment, to a much more subtle thing than the form, the difficulty is strangely increased, - and multiplied almost to infinity by this great fact, that, while form is absolute, so that you can say at the moment you draw any line that it is either right or wrong, colour is wholly relative." Later of coloured sketches "Give up all the form, rather than the slightest part of the colour" (Ruskin 1907)

We have already seen that Garner regarded form and colour as separable dimensions but that studies using a different task suggest mental normalisation must occur on the unattended dimension before making simultaneous comparisons of form with irrelevant colour or vice versa (Egeth 1966, Hawkins & Shigley, 1972). Slightly different results were obtained by Schroeder (1976) following a lead proposed by Garner (1974). Garner had suggested that colour should interfere with form more than the reverse because a "form cannot exist without a colour but the reverse is not true". In a series of experiments Schroeder required subjects to match form or colour alone when both dimensions varied. The results were clear. When colour was made irrelevant it interfered with the ability of subjects to make form judgements, but the reverse was not true.

6.7 Segregating Surface Colour Dimensions

When surface colour is represented in a sketch it may be helpful to segregate some of its dimensions. Theoretically artists might want to think selectively about any of the 3-4 dimensions of surface colour, about illumination induced surface brightness, or about any one of many aspects of colour modes of appearance.

6.7.1 Luminance Contrast without Chromaticity

Typically monochrome studies in pen and wash or charcoal for example show combined attention to surface tonality and light and shade. Because Munsell value, hue and saturation are integral dimensions, separation of value from hue and saturation, is almost obligatory for selective thought. Yet there has been very little published discussion of the possibility that achromatic studies may enhance
6. Visual Segregation and Selective Attention

the perception of the subtleties of, for instance, reflected light and shadow and form from shading.

Livingstone and Hubel (1987) produce psychophysical evidence that luminance contrast without chromaticity supports (besides stereoscopic depth and movement perception) perspective depth perception and figure ground separation in pictures, but that all these perceived attributes evaporate when only hue contrast is present (ie at equiluminance). This they ascribe to the fact that these processes are performed by the "colour blind" magnocellular pathway in the brain. Whilst some of their claims need validation, there are interesting, testable, consequences for designing new sketching systems. Their results imply, for example, that coloured lines in sketches need strong luminance contrast with the ground to depict spatial structure. Portraits with greatly altered hues (as in Fauvism), succeed in depicting their sitter only because the luminance relationships are preserved.

6.7.2 Other types of Colour Attribute Segregation

Clear examples of other types of colour attribute segregation are harder to find in rough sketches. However in post-impressionist painting, (which has been influenced by sketch attributes) there is evidence, of an emphasised attention to colour hue, by the suppression of variation in colour saturation and tonality (Wilenski, 1944). One reason for this may be that paint has only a relative reflectance range of about 1 to 27 (3% for ivory black; 80% for titanium white), whilst in nature the range is far greater. Before Impressionism colour saturation of the darker colours was lost by adding black. Since impressionism, figurative painters have learnt to heighten chromatic experience by compressing the total range of luminance and saturation at the expense of shadow.

6.8 Segregating Spatial Attributes.

There are several ways in which visual attributes related to spatial structure may be segregated in sketches.

6.8.1 Segregating "what" from "where"

The distinction in the visual system between seeing where something is in three-dimensional space and recognising what it is has already been discussed (5.3.5). The distinction is supported by evidence about the components of visual working memory (mainly by studies of divided attention), in evidence about the resources available to, and shared by, perception and mental imagery and in neurophysiological evidence about the visual pathway.

Experiments by Sagi and Julesz (1985) suggests that knowing "what" even a single feature is requires time consuming processing and focal attention. Only
knowing "where" a target is mediated by automatic parallel processes. This distinction is reflected in the visual pathway within the brain (5.3.5).

The analogy of a labelled map with flexible pointers to large stores of information, some of which can itself be mapped in greater detail, was made earlier. It has also been pointed out that there are elegant functions for this separation.

1. A single stored object representation can, with suitable transformations be mapped repeatedly onto many different locations, like leaves on a tree or figures in a crowd. (This is also an aspect of descriptive to depictive mental translation).

2. Fast movements of attention can be implemented to activated, but only partly analysed, regions of the visual field, before slower detailed analysis and memory matching is necessary (focal attention). It also follows that it is possible to contemplate the nature of an object without simultaneously worrying about its spatial context.

3. Two kinds of mental image generation are made possible. Generation by attentional scanning of a spatial representation and generation by retrieving stored object attributes from memory (Kosslyn & Shin, 1994).

4. The separation allows us to reconcile otherwise conflicting evidence for spatial based and object based attention, with two anatomically distinct attentional networks in the brain (although this is still speculative).

Rueckl et Al. (1989) have made an interesting computational model to investigate why location and object identity are processed by separate visual systems. They designed a "neural net" computer model which was set the task of learning to recognise simple shapes placed at random locations within a two-dimensional grid. The model performed this task much more efficiently if the hidden layer of "neural" units (connecting input units to output units) were divided into two types, those responding only to location and those responding only to object shape, than if all the hidden layer units were of the same type responding both to location and to shape.

Of course, in the brain, the two systems must be flexibly linked. Otherwise there is an explanatory problem. Object recognition often depends upon the spatial relationships of parts. I suggest this is solved by redefining the object-part boundary as discussed above. After descriptively represented part relationships are mapped into a "zoomed in" coordinate system on the spatial location map, the parts can be scanned at new object locations.

Not surprisingly, sketches can often be classified according to whether they emphasise what the subject matter is, or whether they describe where the subject matter is within the plane of the picture or within the volume of represented space.
Where spatial location is of primary interest, object identity is neglected. In the contrasting type an artist appears interested in describing an object and the features by which they are recognised, but is not interested in their spatial location. This distinction is illustrated in some of Cezanne's sketch books. The lines in his sketches for landscape composition describe with precision the locations and the spatial intervals between edges and fragments of rocks and trees with little attention to the objects themselves (Illustration XVIII). On other pages of his sketch books well observed details of faces and other objects float in random juxtaposition on the page (Illustration XIX).

Thus it is reasonable to suggest that the capacity to plan and visualise spatial composition is increased by the generation of sketches which outline spatial structure, omitting object detail. An inverse advantage can be postulated for easily located object detail. It may be easier to describe, recognise and re-describe an indeterminate object if involuntary spatial search processes are avoided. Computer graphic systems based on the vector method of representation typically also separate the representation of location and object shape (Hopgood et al. 1983). However the advantages to sketching have yet to be fully explored (e.g. for depictive to descriptive translation (9.4.4)).

6.8.2 Sketching and Spatial frequency Channels

Another form of spatial segregation, which is sometimes found in sketches, also has parallels in the visual system. It is the segregation of different spatial frequency ranges of luminance contrast. In realistic images the rate at which changes of light and dark varies with distance varies widely. It can occur slowly and gradually over wide areas (defined mathematically this is a low frequency luminance contrast). Alternatively changes between light and dark regions may occur more abruptly over smaller regions, corresponding to higher frequency luminance contrast. The existence within the visual system of a number (at least six and perhaps as many as nine) of parallel spatial filters, selective for frequencies with a bandwidth of about 1.5 octaves and sensitive to orientation, is well documented (see Ginsburg, 1986). Low frequency channels respond to gentle changes of luminance with distance and have wide receptive fields. High frequency channels respond to sharp contrast changes and fine details and are associated with smaller receptive fields. The evidence for orientation tuned band-pass spatial filters in the mammalian visual system is reviewed by Robson (1975) and Ginsburg (1986). In normal percepts and realistic images a continuous range of spatial frequencies are super-imposed upon each other and we do not normally notice different spatial contrast frequencies separately. However there are times when an
artist or designer will wish to attend selectively to either the sharp high contrast
detail, at edges for example or to the gradual blurred contrast changes which define
tonal masses.

Cognitive scientists have tended to concentrate on simulating by computer
the functions these parallel spatial frequency channels may have in the early edge
extraction and grouping stages of visual analysis (Marr & Hildreth 1980, Watt
1988). The different algorithms used by these and other authors is beyond the
scope of this study. Both the above authors assume that the outputs of band-pass
filters at different spatial frequencies must be combined in the early stages of
vision to find edges but their theories about how this is achieved are different.
Thus it must be pointed out, that while fine line sketches may capture higher
spatial frequencies than less detailed, smudgy ones, it is necessary to use both high
and low frequencies to distinguish object edges from the many other types of
luminance boundary present. Mathematically "Convolving" a black and white
photograph with a plain Gaussian filter (Watt 1988) produces an image very
similar to a Seurat drawing. A high frequency and low frequency spatial
mathematically filtered photograph of Groucho Marx is illustrated for comparison
(Illustration XXVII).

There is evidence that it is possible to attend selectively to different spatial
frequencies (Graham, N. Kramer & Haber, 1985, Shulman & Wilson, 1987a,
1987b) but with a penalty. The hypothesis here is that frequency segregation makes
it easier, and that it frees attentional resources for other cognitive tasks useful for
visual invention.

In the array of luminances presented to the eyes by normal scenes high,
medium and low spatial frequencies are mixed and are processed in parallel. In
sketches high frequency segregation manifests itself in delicate line drawing
representing sharp edges and fine detail with little or no representation of gradual
shadow or global luminance contrast (example). Low frequency segregation is
beautifully demonstrated by the tonal studies of Seurat. Surface texture and local
detail are suppressed whilst the global subtleties of tonal structure and broad
masses created by gradual gradients of luminance contrast over wide areas is
emphasised.

6.8.3 Forests or Trees ?: Global or Local Precedence

Because a low spatial frequency filtered image selectively represents
gradual changes of luminance over wide areas, smudging out fine detail, it can be
looked upon as one way of attending to global characteristics of a scene as opposed
to local detail. However spatial frequency provides no information about the
perceived (descriptive) hierarchy of meaningful objects and their parts discussed above. There is thus another type of global-local distinction which depends upon whether or not we wish to attend selectively to the overall shape and structure of an object or to its perceived parts. As suggested above (6.5) the preliminary perceptual structuring of the sketch stimulus field into candidate objects and parts is involuntary, but can be manipulated by the way the sketch component marks are grouped and related. However it is also worth considering how sketch attributes can facilitate (without completely separating) global or local judgements with minimal interference from temporarily irrelevant structure. There is a large but rather complex literature which can be drawn upon.

In a widely quoted series of experiments Navon (1977) investigated how global and local attributes interacted in pattern perception. He showed subjects letters or simple geometric forms which were themselves formed by grouping together smaller versions of the letters or geometric forms. In a typical experiment upper-case 'H's were made by grouping either smaller 'H' s or 'S's and 'S's made by grouping smaller 'S's or 'H's. The stimuli were thus of two types, \textit{consistent} in which the local parts had the same form and name as the global shape and \textit{inconsistent} in which the global and local attributes differed in their form and name. Subjects were asked to make speeded identifications of such stimuli under two instructions (a) to identify the global form and (b) to identify the local form. Navon found that identification of the \textit{global} forms was always faster than that of the local forms. Further, although it made no difference to global identification whether the local elements were consistent or inconsistent, local identifications were significantly slower when the global form was \textit{inconsistent} with them than when it was consistent. Thus there was a sort of Stroop type interference of irrelevant global form on attention to local parts but not vice versa. Navon had considered that his results supported a general principle of "\textit{global precedence}" in which the processing of global attributes precedes, or occurs faster than, the processing of local features. However this conclusion has since been shown to be premature.

Navon had considered the possibility that global precedence might only hold for stimuli within a limited size range so he ran another set of experiments using stimuli of two sizes with identical results for both sizes. However Kinchla and Wolfe (1979) using a target search task showed that global judgements were faster than local only if the global stimuli subtended a visual angle less than 6°-9°. For larger global stimuli and when local stimuli were a constant fraction of global size, local judgements were faster. Kinchla and Wolfe suggested that there is an optimal size of about 2° for which forms will be preferentially processed with
higher or lower level forms processed later. Their results support a principle of optimal size precedence. Ward (1982) has suggest that these results can be related to the work of Larsen and Bundesen (1978) suggesting that size normalisation precedes shape matching judgements (4.3.6). The principle would be explained if the frame used for shape feature detection has a preferred size (about 2°) and shapes are accordingly mentally adjusted to that size. However, in another set of experiments Lamb and Robertson (1990) showed the visual angle at which the transition from a global to local advantage occurs, varies with the average angle in a block of trials. Thus the transition takes place at a larger visual angle for a large stimulus set than for a smaller stimulus set. It is as if the visual system can, with practice, set itself to favour a certain visual size for global object perception. The authors conclude that attention plays a major role in determining the relative speeds of processing global-level and local-level information.

The principle of optimal size precedence is relevant to the question discussed in 2.6 about how the physical size of a sketch should relate to the real or imagined size of the object it represents. Artists probably instinctively adjust switches between desired global and local levels of attention, by varying the visual angle of a sketch as the eye is moved towards or away from the plane of the paper. However because the plane of the traditional sketch surface is so salient it is possible that the visual system will compensate for such distance related size changes with "phenomenal size regression" (4.2.1) thus counteracting the required attentional level change. The rapid "zoom" facility, possible with machine sketching systems, highlights the importance for system designers of understanding the connection between visual size, object attention and global or local precedence. However the question remains open until controlled trials in realistic contexts are performed to confirm or refute the applied theory.

Another limitation on the principle of global precedence was shown by Martin, M. (1979). She found that Navon's global precedence results were only obtained if the global features were made of many small local elements. When she tried stimuli in which the global features were made of fewer and somewhat larger local elements the opposite effect was obtained. Local judgements were faster and unaffected by the identity of the global form, whereas global judgements were interfered with by an inconsistent local form. Ward (1982) suggests that her data support a principle of "conspicuity precedence" in which the most conspicuous or salient features in a visual scene will be processed first. Here conspicuity is used in a sense defined by Engel, F. (1971, 1974). Thus object features that are brighter, that stand out more from their background, that are seen in more detail (ie foveal) or are better formed in the Gestalt sense will attract attention and be preferentially
processed. One of the many ways in which sketches differ from more realistic ("stimulus bound") representations, is that they usually provide added salience (conspicuity) to objects, parts or attributes of interest using Engel, F.'s factors - enhanced contrast, greater isolation etc. Here it is useful to make the distinction between visual segregation between sketches as discussed above and visual segregation within a sketch which is implied here.

The above factors affecting the precedence level of perceived structure are characteristics of the sketch stimulus itself. The implication (to me at least) is that they bias attention by their effects upon early involuntary processes which direct attention (Posner's posterior attention network). There is another line of research which demonstrates the influence of the observer's cognitive state. For example Hoffman, J. (1980) using a memory scanning task obtained results which also disagreed with the global precedence hypothesis. He found mutual interference between global and local levels of visual structure regardless of which level was attended to. He also found that when subjects had to divide attention between two levels, responses to targets at both levels were made equally quickly, but more slowly than when attention was allocated to a single level. Hoffman, J. concluded that the ease of processing a given level of figural detail depends both upon the "quality" of the information at that level and on the processing capacity allocated to that level. Following this finding Ward (1982) investigated how a previous level of processing (in the global to local continuum) biases the current processing level. He found a robust "level-readiness" effect whereby processing at a given level is faster if previous processing has also been at this level. He suggested that interference between levels depends upon a conflict between the degree of salience at one level and the allocation of attention to another. He proposed a two-state model of attention switching, to explain his results. These results seem to me to be relevant to the strategy pursued by a sketch user when thinking about a design or work of art at a particular level of spatial detail.

As already pointed out there is a distinction (related to the depictive-descriptive distinction) between global and local levels defined in terms of spatial frequency (degrees of fuzziness) and defined as here in terms of object structure. However in practice, selective attention to a particular spatial frequency and attention to a particular level of object structure will over-lap, since, for example, attention to low spatial frequencies which blur fine detail is also likely to facilitate attention to global structure. Shulman and Wilson (1987), set their subjects the Navon task of identifying either the global or local letters. However on some trials they superimposed a sine wave grating of a particular spatial frequency on the local-global display. After the trial subjects were asked to make a present/absent
judgement on the probe. The results showed an increased detectability of low frequencies during global processing and of high frequencies during local processing. This suggests another explanation of some of the global precedence findings, because there is evidence that the low frequency channels are faster (higher temporal frequency) than high frequency channels (or are associated with receptive fields which have faster channels (Graham, N. 1981)). When an image is progressively blurred, the highest spatial frequency components are removed first, with progressively lower frequency components being removed as the blurring continues. This suggests another way by which a machine sketching system might assist visual attentional precedence control, which is worth testing. Not many sketch users, after all, have Seurat's powers to filter out unwanted spatial frequencies (Illustration XXVII).

6.9 Discussion and Conclusions

6.9.1 Selective Attention and the Sketch

There is ample evidence that the visual system in man and other primates segregates different attributes of perceived and remembered visual information using dedicated resources which interact in complex ways. It is also clear that the brain possesses a number of neuro-anatomical resources that serve different parts of the task of attending selectively to one component only, of the available visual information (Posner et al. 1994). We can attend selectively either to different objects, to different visual attributes or to limited regions of a visual field. The systems which control such orienting of attention can behave both as a moving "spotlight" facilitating limited but fixed areas or as a "zoom-lens" processing selective levels of the global to local hierarchy of visual structure. This hierarchy may be defined spatially in terms of band-width spatial filters or in terms of a descriptive structure of objects and parts. In addition we can attend separately to colour and to the various dimensions and attributes of colour. Much of the literature can be fitted to a framework in which attention mechanisms have evolved to select objects or candidate objects.

Evidence has been presented which shows that irrelevant dimensions, attributes and objects interfere in different ways and to different degrees with such selective tasks. The main hypothesis is that sketches by separating out attributes and parts, by removing variation in irrelevant dimensions, re-scaling sizes, controlling salience etc, can lighten the cognitive effort necessary for selective thought. Attention theorists still debate whether the need for selective visual attention is due to capacity limits, with the early need to filter out irrelevant
information, or to the need to resolve processing response conflicts caused by involuntary parallel processes. There is evidence for both.

Whilst in the early stages of composition or design, it may be necessary to think selectively about one or more partial attributes of a visual problem, it is also necessary to integrate all the visual attributes and parts into a unified whole. Where previously separated attributes interact as they usually do (e.g., colour and form) further manipulation will be necessary. In painting, for example, it is rare to find that a composition developed in colourless sketches does not change its form when colour is used (see "Les Demoiselles D'Avignon", Illustration XIII). Fashion designers sometimes design a single garment shape to be implemented in fabrics of different colours. Sometimes also in the Fashion industry the buyer (Bates et al. 1988, Makirinne-Crofts et al. 1992) wants to see a particular garment style in a different fabric. On the other hand Fashion designers frequently tell you that their ideas start with the fabric which they need to handle (see also 2.4).

Both "forests" and "trees" can have precedence in early sketches. Picasso's global ideas for "Les Demoiselles d'Avignon" were influenced by new local ideas for representing human figures (1.7.1). One must not forget either that in figurative sketches there is often an implicit structural redundancy which can be interpreted at different levels of structure. If obscure lines in Goya's "Procession" represent a figure (global) then it probably has a head, arms and legs (redundant local part). If a mark is seen as a face (local) then it is equally likely to belong to a figure (redundant global part) with trunk, arms and legs (Illustration XXII).

In such cases integrating visual attributes implies a many to one relationship between an attribute selected early in the thought process and one selected later. The role of selective attention in inventive visualising can perhaps best be understood in stages:

1. Effortful selective attention is paid to source material to segregate, by memory or observation, visual attributes of special interest or importance.

2. Records of selected attributes (sketches etc) are used to facilitate selective manipulation and visual thought. A variety of mechanisms are available for instance:
   - Attribute isolation or filtering (colour dimensions, spatial frequency etc).
   - Normalisation of irrelevant attributes for comparison
   - Global / local level selection by optimal size, grouping, object - part salience etc.

3. Separately considered components are re-integrated, usually in stages, for a final synthesis. Thus early idea drawings by the late designer Jean Muir described global attributes of a garment. Later drawings added details about
buttons, collars, pleats stitching etc, integrating global and local levels of attention (Muir 1984).

4. Through this process the artist's mind may "back-track", returning to reconsider an earlier decisions. This may then lead to further stages of selective segregation and re-integration until both global structure and local detail of the whole design can be visualised together. (Do late changes to detail then implicate further modifications to the global design?).

The need both to visualise attributes selectively and to integrate all components into a satisfying whole suggests that neither the traditional impoverished attributes of sketches nor very complete and "realistic" representations solve the whole problem. Especially relevant to the sketch is the concept that the capacity to divide attention between features, depends upon whether those features are perceived as parts of the same object. Visual composition in the arts is largely a question of manipulating the relationships between the separate components of an image or imagined object until they appear to cohere into a meaningful whole. Perhaps this mysterious process can be seen as a an unconscious cultural adaptation of innate object construction mechanisms in the brain.

The ideal system would facilitate the types of attribute segregation discussed above and their separate manipulation whilst at the same time it permitted a later more detailed image resynthesis to occur, in successive stages, as partial decisions are taken. This possibility is considered further in Chapter 9.4.

We are now in position to re-state hypothesis four.

6.9.2 Hypothesis Four

Sketches amplify inventive thought by isolating and representing separately those attributes of visual experience that are of special relevance to a particular task. This assists the user to attend selectively to a limited part of the task, freeing otherwise shared components of cognitive capacity and reducing the complexity of preparatory visual processing.
7. METACOGNITION AND CONSCIOUS CONTROL
"The beast does but know, but the human knows that it knows" John Donne (1628)

7.1 Introduction

**Hypothesis Five** states that "Sequences of sketching acts support the conscious awareness of one's own cognition. This assists creativity, providing voluntary control over highly practised mental processes which can otherwise become stereotyped. Unforeseen percepts from untidy or accidental stimuli can elicit unconscious processes which break the mould of habitual thought, whilst a temporal record of recent ideas makes it easier to change one's mind at appropriate stages."

**Metacognition** is a term used by psychologists to refer to a knowledge of and awareness about one's own cognitive processes. It is sometimes divided into a number of sub processes (Matlin 1989). For example **Meta-Attention** is the awareness of how one's attention is allocated or divided and **Meta-Memory** refers to one's awareness and knowledge of one's own memory. One can know something without being aware of it and one can be aware of knowing something without being able to access the information (the "tip of the tongue" phenomenon). Many of an artist's personal visual skills are based on knowledge acquired from long experience, but which is not available to conscious expression. Such "**tacit knowledge**" influences visual invention but is not available for verbal report or even voluntarily controlled action. Studies of the working methods of designers (Lawson, 1990, Makirinne-Croftes et Al. 1992) and of fine artists (Arnheim 1962a) confirm the importance of such tacit knowledge.

**MetaCognition** is a form of consciousness. The nature of consciousness is such a controversial subject that even now in the post-behaviourist era with concepts of "mind" and internal representation at the centre of cognitive psychology frequently the word is not mentioned in the index of text-books (e.g Eysenck, 1984, Kosslyn 1980 Baddeley 1990). However the unconscious-conscious distinction is implicit in many of the topics discussed. Thus in the preceding chapters it was found to be impossible to avoid the explicit or implied distinction between conscious and unconscious processes. For example

- The late construction theory of perception is actually a theory of perceptual consciousness. But what internal operation or attribute distinguishes the visual processes we are aware of from those that we infer to occur outside awareness?
In theories of visual recognition, the distinction between those automatic memory access processes which occur unconsciously and those which are initiated with conscious intention is unavoidable.

The most interesting but least discussed attribute of mental images is that they are usually conscious (or at least often conscious) visual memories which can be voluntarily controlled. This distinguishes them from the reservoir of unconscious stored visual representations of which we are not aware but which we infer to explain the facts of recognition. Kosslyn (1980) gives credence to Freud's theory that one of the functions of mental imagery is to make memories conscious, but he takes the idea no further (see 5.2). If it is true, then it follows that since mental imagery is important to invention then conscious awareness must be also. (But there is also evidence for unconscious imagery 7.2.9)

In the following sections I will treat **awareness** and **volition** (as in voluntary versus automatic processing) as connected but separable topics. I first discuss briefly some existing theories of consciousness, including my own theory of consciousness as a "cognitive monitor". Then I review some relevant literature about the difference between automatic and voluntary cognition and discuss the reasons why the related pairs of opposing concepts (1) automatic versus voluntary processing and (2) conscious versus unconscious processing, are relevant to visual invention and the sketch. This will lead to what is, I hope, a plausible extension to current theory, in order to explain how sketch attributes support the complex but limited capacity processes of visual awareness. It is postulated that the (unsolved) problems of the nature of visual creativity and the nature of visual awareness are interdependent.

The core of this chapter is that sketches are cultural inventions which facilitate an interchange between conscious and unconscious types of visual knowledge and that this internal dialogue is necessary for visual planning and creativity.

### 7.2 Consciousness: Theoretical Issues.

#### 7.2.1 The Imperial Psychology's New Clothes.

"Was it too early then to talk about the Imperial Psychology's New Clothes?" Mandler, G. (1975)

One reason why many (but not all) psychologists have been reluctant to use consciousness as an explanatory concept in cognitive theories is the historical failure of conscious introspection to provide reliable experimental data about mental processes and the difficulty in finding objective criteria for conscious awareness. Miller, G. (1962) recommended a ten year moratorium on the use of
the words "consciousness" and "voluntary" in psychology, long after the demise of behaviourist strictures on "mind". For most theorists the ban has lasted much longer, although many discussions of attention (Treisman & Paterson 1984, Posner 1980a, Julesz 1985) and some of short-term memory (Marks 1983, last chapter of Baddeley, 1986) all but admit they are also about the conscious - unconscious distinction. With the exception of a few bold theorists (e.g. Shallice, 1972, Mandler, G. 1975, Posner 1980) consciousness had been regarded as a philosophical rather than a scientific problem (e.g. Ryle 1949, Nagel 1974, Churchland, 1984, Rosenthal 1986, Dennett 1991). More recently however (from the eighties onwards), the ice has started to thaw over the topic. Consciousness is attracting the attention of many more cognitive psychologists (Pope & Singer (Ed.) 1978, Underwood & Stevens 1979, 1981, 1982, Davidson & Davidson (Ed.) 1980, Marcel & Bisiach (Ed.) 1988, Harnad 1982, Johnson-Laird 1983, Marcel 1983b, 1988, Mandler, G. 1985, Jackendoff 1990, Milner & Rugg 1992, Baars 1988). Despite the rebirth of scientific interest in the subject (it was a serious topic in the time of William James, W.), no commonly accepted functional theory of consciousness has yet emerged. There are probably as many theories as there are theoreticians.

As Marcel (1988) has pointed out psychology needs to treat consciousness seriously if only because it is implicit in many of the experimental methods on which it is based and in many cases it is consciousness which is actually being examined. For example if verbal introspection is disallowed, button pressing certainly is not. Yet button presses are often used to perform the same function as introspective speech acts conveying fragments of consciousness "this is when x looks the same as y" etc. Marcel concludes that "In sum...we should acknowledge that what we are studying either is or involves consciousness".

7.2.2 Distinguishing Awareness from Volition.

There is no agreed definition of consciousness and many theorists regard it as a collection of concepts. For example Wallace, B. and Fisher, L. E. (1983) list seven different definitions.

One blurred but important distinction is especially relevant to this discussion. Conscious awareness (as opposed to "non-awareness") and Conscious control (ie voluntary as opposed to automatic processing) are related concepts which are frequently discussed separately. Awareness presupposes the representation of information which is either sensory in origin (perceptual experience) or internal in origin (metacognition, memory recall). Both types of awareness are difficult to investigate objectively because they are either purely
private experiences or un-provable hypotheses about other peoples experience. Models of awareness which avoid the trap of an infinite series of minds, but allow some regression within the mind are not impossible to construct (see Johnson-Laird, 1983). Thus one can imagine that Robert Welch in designing the Yamazaki Collection (Crawford, 1983) was aware that he was planning grooves on a coffee-pot and aware that he was aware he was planning grooves and perhaps aware of his strategy that led to this awareness but probably exited from the mental loop here. It may be the very existence of automatic processes and the limited capacity of working memory which provides the necessary exit from such a regress.

A little regression is not without its uses. Being aware that one is aware allows an extra, higher level strategy for making decisions which would not otherwise be possible (see Johnson-Laird, 1983, 1988). One can illustrate levels of awareness recursion (fuzzily) with some purely hypothetical reflexive inner dialogue by Robert Welch (1986)

Level 1. question "I need grooves in this table-ware to collect the light. What shape of profile should the grooves have?" (Illustration XIV)
Answer "They could be sharp or rounded. I will sketch them vaguely and add a little visual note to show two alternatives."

Level 2. question "Shall I continue to think about grooves (level 1) or do I now know enough to shelve the problem until I have clearer ideas about global shape?"
Answer "Yes probably but I mustn't forget that the grooves will alter the look of the global shape and vice versa, so I may have to think again about them before long" (Illustration XIV)

Level 3 question "Is my whole approach to this design problem correct, (level 2) or should I on principle clarify a set of global shapes before worrying about local surface structure?"
Answer "Sometimes that might be true but with this project surface relief is going to be too important to leave to the later stages" (exits to level 2 again)

The concept of voluntary as opposed to automatic cognition is a distinction more open to experimental attack. However it also involves conceptual difficulties (see Allport 1980). Awareness of a cognitive process and voluntary control of a process appear to be separable in so far as we can be aware of a mental event such a headache or a bad smell from our neighbours drains without being able to exercise any voluntary control over the inner experience. Less obviously, an artist may be conscious of an ingrained thought pattern when designing but cannot necessarily interrupt it voluntarily when required. The separability between conscious awareness and conscious volition may be more
apparent than real. If it exists it is certainly uni-directional. For it is difficult to see how either an overt or covert cognitive actions can be "voluntary" if we are not also aware of them. Tentatively then we can say that "voluntary" actions are those which are controlled and monitored by something referred to as "awareness". Yet we may be aware of some processes we cannot control and which are therefore automatic. However it has frequently been claimed (Miller, G. 1962, Posner 1980a, Johnson-Laird 1983) that we are aware only of the products of thought processes rather than the processes themselves.

In this connection, Hilgard's (1980) distinction of two basic kinds of conscious awareness the passive mode and the active mode may be useful. The passive mode according to Hilgard includes perceptual awareness when no response is required, daydreaming, listening to music etc. In contrast the active mode involves the need to plan, to make decisions and to act upon those decisions. Clearly purposive sketching is more likely to be classed as active mode awareness rather than passive mode but the distinction is blurred. There is surely a grey area in which the mode depends upon how we define planning and decision taking. Thus the apparently passive contemplation of a sketch can conceal volitional decisions related to judgements and mental imagery which are hard to classify as passive or active on Hilgard's criteria. Moreover they may be an important part of the concealed processes used in creation (7.4.3).

7.2.3 The Problem of Criteria

Allport (1988) has claimed that there are no general criteria for consciousness, because (he claims) there is no such general phenomenon. Later in the same paper however he seems to relent slightly and in suggesting that consciousness may be understood by studying the way behaviour is integrated he tacitly admits that there is such a phenomenon. It is true that although we know when we are visually aware ourselves, we cannot use the kind of objective criteria for other peoples awareness that are traditionally demanded for other types of experimental observation. Allport discusses three criteria that have been suggested in the literature. The criterion of potential for action attempts to distinguish conscious from unconscious information by whether or not it is available for overt actions (such as speech or button pressing) as well as internal thought. A memory criterion suggests that the ability to remember events is evidence for awareness of those events. According to the "confidence criterion" (Cheeseman & Merikle, 1985) the subject must not only be able to indicate correctly a stimulus identity, he must also indicate confidence in his or her own report. The trouble with all these (and other) criteria, Allport claims, is that there are always cases where they can
be shown not to be reliable and moreover they can be dissociated from each other. There is even evidence for speech without awareness (Milner & Rugg, 1992).

Clearly in order to make any progress with the study of awareness, it is necessary to accept fallible criteria. Marcel (1988) suggests that the only criterion we can and must use is an individual's verbal statement of awareness of their own conscious experience. He points out that this is often unavoidable. For example, in medicine it has long been recognised that a patient's introspective accounts of different kinds of pain can sometimes provide a better diagnostic guide than the use of any instrumental measurements (Marcel 1988). Moreover, now that more is known about which processes are unavailable to conscious self-inspection, it is possible to reconsider how introspection can be better managed as a way of gathering data. It is understood that in many cases, this can provide access only to the results of cognitive processes and not to the processes themselves (Johnson-Laird, 1983). In recent work on introspection and protocol analysis, what is verbalisable is sometimes treated as equivalent to what is conscious (Ericsson & Simon, 1984). Although there are many problems with this (for example visual experiences may not be verbalisable), carefully used protocol analysis (getting subjects to report conscious stages of their thoughts) has yielded useful insights into problem solving which could hardly be achieved in other ways (Newell & Simon, 1972 also Goldschmidt, 1991 see Chapter 3).

Objective criteria for awareness are unimportant to an individual artist in the act of sketching. However they are relevant to the problems faced by those who are seeking evidence about the thought processes of artists and designers, in order to design improved sketching tools. Here we are concerned not only with distinguishing tacit from explicit visual knowledge, but also unconscious and automatic thought processes from those under voluntary control.

7.2.4 Functions of Consciousness

Amongst the various functions for consciousness that have been suggested there are some which are directly relevant to the inventive functions of the sketch.

Mandler, G. (1975, 1985) has proposed that consciousness has a special role in mediating the activation of stored schema for actions and in selecting between possible alternatives and competing hypotheses. Because of its limited capacity, Mandler, G. states "very few preconscious candidates for actions and thoughts will achieve this additional consciousness mediated activation". If one were aware of everything one knows, he points out, one would "be swamped with information and unable to act". Consciousness intervenes whenever one must re-assess potential action choices in the light of a new situation. It is also used to
modify and "interrogate" long range plans, to initiate (but not to implement) memory retrieval and to "comment on one's current activities". Lastly it has a special role in "trouble-shooting" structures not normally represented in consciousness. This last function is echoed in an otherwise curiously limited proposal by Kosslyn and Koenig (1992). They suggest that consciousness is a sort of "Parity check" which assesses whether or not the mental system is functioning properly.

Marcel (1988) considers the causal status of consciousness in the sense that when someone does something for a reason, that reason is one of the causes of the action. Beliefs and intentions he implies must become conscious if they are to modify behaviour. Phenomenal awareness enables a certain kind of self-monitoring. For example in the condition known as "blind sight" patients who are apparently blind in part of their visual field can nevertheless correctly "guess" the position and sometimes the shape of stimuli presented in their "blind" field. Their lack of visual awareness prevents them from knowing that they can see and therefore of using (unconscious) visual information for any intentional goal related actions. Moreover "if one cannot consider courses of actions consciously, then one would be almost unable to carry out planning of action". Marcel points out that we need access to a moment-to-moment representation of our environment, our ourselves and our immediate past. Without such a representation we would be unable to make plans which are not rigidly dependent upon our immediate circumstances. Marcel even hints in an off-hand manner at the fact that such representations may include cultural inventions such as number systems and writing. Another function discussed by Marcel (1988) is the role played by metacognition in learning. Knowing that one knows and what one knows, Marcel claims, is essential for reflecting on phenomenal experience or in selecting between competing cognitive processes. "Unless one can pay attention at an appropriate level to one's own sensations or actions, one will have great difficulty learning a new essentially arbitrary, skill which is based on that level of representation".

Some of the most interesting work on metacognition has related to learning in children. It may illuminate by analogy the problem of visual invention and novel problem solving by mature artists. Thus Leslie (1987) has proposed that infants acquire an internal representation of the world in two stages. In this view the infant starts by acquiring a first-order representation as described by its sensory apparatus. Only later does it develop a second-order representation which includes knowing that it knows or perceives. This ability frees the child from the belief that what it perceives or imagines necessarily represents reality. It underlies, according to Leslie, the capacity to play with symbols where possible non-existent worlds can
be imagined and in which one thing can stand for another. It also, in this view, allows the child to develop a theory of mind and to conceive possible models of other peoples minds and beliefs. Many of the behavioural characteristics that distinguish humans from other primates are neotonous (primate infant characteristics adapted by evolution for adult purposes (Morris, 1967)). Thus we can look on sketching in the visual arts as "playing" with imaginary worlds for adult purposes. The child's "second-order" representation is now facilitated by cultural sign systems in order to access and manipulate tacit visual knowledge.

In agreement with Mandler, G., Marcel believes that cognitive tasks which are not biologically given, must be made conscious in order to be acquired. The "intention to perform a non-habitual or non-stimulus-driven task needs to be conscious". Even if many of the mental processes used to perform a task are unconscious, the constraints and instructions which define it, together with the intended result needs to be made conscious otherwise slips of action (automatisms) will result (Reason, 1979).

Oatley (1988) considers the ability to change one's mind is an important function of consciousness. He classifies consciousness into four aspects. His "Helmholtzian" consciousness refers to perceptual experience. According to the late construction theory of perception - supported here - conscious percepts are constructed from the "conclusions" of inference-like processes driven by sensory and memory data as Helmholtz suggested. Oatley's Woolfian consciousness is equivalent to James, W.s' (1892) "stream of consciousness" consisting of an undirected stream of mental images, so well described by the writers James Joyce and Virginia Woolf. (In Chapter 5 I have, in effect, proposed that Helmholtzian and Woolfian modes of consciousness have a common origin and that in sketches they are used collaboratively as a single mode.)

Vygotskian consciousness is Oatley's term for that aspect of consciousness in which one uses mental representations to solve practical tasks. Drawing on Vygotsky's studies of children talking to themselves as they solve problems (Vygotsky 1978) Oatley draws attention to a mode of consciousness in which inner speech and mental imagery combine to provide a "simulation space in which we can plan and try things out". He agrees with other theorists that it is the goals and results of inference processes that are manipulated in consciousness rather than the inference processes themselves. We cannot manipulate the goals of unconscious processes. This overlaps with Marcel's version of self-monitoring and task analysis. Oatley's fourth mode he calls Meadean consciousness. The term refers to Mead's theory that consciousness has a social origin. According to Mead, tacit continuous monitoring goes on whenever we speak. We understand what we are
7. Metacognition and Conscious Control

saying to others because as we speak we are also speaking to ourselves. He postulates that consciousness is an internalised representation of what started as external relationships with other people. Humphrey (1983, 1986) has developed a similar theory (but with the reverse emphasis) as the result of his studies of social behaviour in primates. Humphrey believes that a primary function of consciousness is that by making an individual animal aware of its own mental processes it is better able to understand and predict the goals and behaviour of other individuals within a group. (Scrivener and Clark, S.'s (1992) study of the dialogue which occurs between two individuals during collaborative sketching is relevant to this theory)

Oatley suggests that these four modes of consciousness are used collaboratively "to create new pieces of cognitive structure". This need arises because human intentions commonly have multiple goals which can conflict with each other. It follows from this (if I understand him correctly) that a special type of cognitive process is needed which can back-track on mental actions in order to resolve conflicts by re-assessing goal priorities and forming new plans. In his own words "What is needed is a cognitive process for rewriting plans, or for rearranging goal priorities, when the unexpected occurs" and later "It is the voluntary ability to use the mind as a simulation space to try out possible actions before committing a new piece of plan to action that is hinted at by the Vygotskyan and Meadean types of consciousness.". Oatley does not discuss the limited capacity of consciousness to perform this function nor does he discuss visual alternatives to Vygotsky's inner speech or how plans and goals might be represented mentally.

Although there no agreement about all the functions of consciousness, amongst theoreticians who treat consciousness as a serious topic, there is at least a consensus of opinion that without consciousness the human mind can deal only with familiar, routine situations without the unexpected or the novel. The postulated functions of framing plans, cognitive trouble-shooting, self-monitoring, dealing with the unexpected and changing one's mind, are sufficient grounds for linking conscious awareness with the sketch. A theory about how sketches support such conscious functions in visual invention is discussed later in this chapter.

7.2.5 Consciousness and the Brain.

The discovery of several types of brain damage which cause separable deficits of awareness has made the nature of consciousness an almost unavoidable topic in neuropsychology (see Marcel & Bisiach, 1988 Weiskrantz, 1986 1988, Milner & Rugg 1992, Farah & Ratcliff 1994, Gazzaniga, 1988, Bisiach 1992). For example it has been demonstrated that damage to primary visual cortex can
leave intact the ability to respond instinctively to the damaged part of the visual field without awareness of vision (Weiskrantz 1986). A loss of conscious recognition of objects or faces may leave intact an ability to recognise without awareness (covert recognition) (Young, 1994). Similarly cases of amnesiac caused by brain damage showing retention in implicit memory tasks of the "forgotten" information are well documented (Schacter, 1992). Also in unilateral visual neglect patients can show proof of semantic processing of parts of a drawing they will not admit to seeing, even when they can trace round the outline of the drawing with a finger (Bisiach 1992, Wallace, M.A. 1994, Marshall & Halligan 1988).

The accumulation of evidence that selective damage to specialised modules of the brain can divorce normally integral conscious and unconscious processing functions locally demonstrates two important points.

1. Specific neural architecture is necessary for conscious awareness. Although this architecture is probably distributed over many processing functions and parts of the brain, this does not preclude the likelihood that there is a separate "conscious mechanism" as Posner (1980a) has pointed out. It suggests (to me at least) that there are specific (if yet undiscovered) cognitive resources, (functional processes and representational systems) which when allocated and switched on ("triggered" Shallice 1972) make otherwise unconscious processes conscious. What that may mean is discussed later.

2. The extreme handicap suffered by patients with limited loss of awareness of a cognitive capacity they still possess (to some degree at least) demonstrates the functional importance of consciousness for effective cognition. This point cannot be taken for granted, because it is sometimes argued that consciousness is a kind of illusion which does not correspond to any specific function of the brain (Dennett, 1991) or that it is an epiphenomenal consequence of the need to integrate many parallel processes (Allport, 1988).

There is also a growing interest in explaining consciousness in neurophysiological terms (e.g. Edelman 1989, Crick & Koch 1992). Edelman's theory, despite his excellent title is very hard to follow. It is based on his "Theory of Neuronal Group Selection" which explains learning and perceptual categorisation by assuming an internal reward system which causes a Darwinian type selection of behaviourally successful synaptic connections. He claims (with much neuroanatomical detail) that conscious awareness has the biological function within the brain of relating an internal state maintenance system to a perceptual and response system. Crucial to consciousness, he suggests, are the re-entrant connections between processing modules. Crick and Koch (1990) have suggested that visual awareness may have a function integrating the many processing
functions, thought to occur in separate parallel modules, and that it may be related to some recently discovered synchronous "gamma oscillations". Crick (1994) has argued that it is essential to study consciousness at a neural level even before we have a clear concept of what it is for. Most cognitive scientists would probably say this approach is premature and defend the functionalist viewpoint that like other mental processes consciousness (if it is to be treated seriously) must be modelled at a higher level of representation and process.

Although their theories are quite different, both Edelman and Crick agree that consciousness is a separable process which uses theoretically identifiable neural resources. Posner has indicated that he agrees with this (Posner 1980a, Posner & Rothbart 1992). The weight of informed opinion now seems to be in disagreement with Neisser's (1976) claim that no such separate processing capacity exists and his rather unfair misrepresentation of the functionalist approach - "many current models of cognition treat consciousness as if it were a particular stage of processing in a mechanical flow of information. Because I am sure that these models are wrong, it has seemed important to develop an alternative interpretation of the data on which they are based".

An eloquent plea that only a greater understanding of the brain can resolve this issue has been made by Francis Crick (1994).

7.2.6 Consciousness and Attention

The concepts of attention and of consciousness are clearly connected but are not equivalent. In many texts the term attention has completely replaced the word consciousness which is never used. The distinction must be made between two usages of the term "attention". It can mean the process of directing, "orienting" or selecting, limited parts of the available information for more detailed or further processing. This process may be either automatic (and unconscious) or voluntary and consciously controlled. However the contents of this selective process, the second meaning of attention, (sometimes called focused attention) appears to be correlated with the contents of conscious awareness. Theoretically a distinction still exists in that "focused attention" might imply merely selective processing conscious or unconscious. However there are now firmer reasons for believing that selective attention is indeed a mechanism for directing conscious awareness (Julesz 1985, Posner 1980b, Posner et al. 1994) and that it is associated in the brain with the anterior attention network but not the posterior one. Posner and Rothbart (1992) discuss the evidence for this. These include PET studies which show that the anterior attention network receives increased regional blood flow in tasks believed to require conscious attention. Interesting in this respect is the
involvement in the anterior system of a part of the prefrontal cortex - the **anterior cingulate gyrus** since this is very close to the **anterior cingulate sulcus** a part of the brain which is believed to be necessary for **volition** ("free will") (Crick, 1994 page 267).

If Posner and Rothbart's intuition about the connection between attention and awareness is correct then theories of attention are either theories about consciousness or the control of consciousness. Thus in Treisman's "**integrative**" theory of attention, focal attention sequentially integrates at each occupied location the object attributes output by many parallel "pre-attentive" processes. The distinction, between "pre-attentive" and "post-attentive" processing, is virtually synonymous with a distinction between "pre-conscious" and "conscious" processing (Julesz, 1985), although the word "conscious" is not in Treisman's vocabulary (see Chapter 6).

Posner and Rothbart (1992) conclude that "the study of attention is **fundamental to understanding consciousness**". Thus if it is correct, as argued in Chapter 6, that sketches support selective attention then they are also likely to be supporting conscious awareness.

### 7.2.7 Consciousness and Working Memory

Once it is accepted that the contents of selective attention are also the contents of conscious awareness then theories of working memory (such as the Baddeley & Hitch model discussed in 5.3.6) must also be linked to consciousness. Indeed most of the evidence for the different components of Working memory in this model come from studies of the amount of interference that results from doing two tasks at once. These are also referred to in the literature as studies of **divided attention**. This does not mean that all the resources of working memory are used in consciousness. However most theorists would probably agree that consciousness requires a short term memory capacity and that this is almost certainly supplied by components of working memory. Both percepts and images are usually treated as conscious experiences although many of the processes that generate them are unconscious (discussed in Chapter 5 & by Marcel 1983b). The resources that they share almost certainly include the two components of visuospatial working memory discussed by Logie (1995). It is not clear whether visuospatial memory is also used by unconscious processes, a point that surely needs more investigation.

Another link between the Baddeley & Hitch model of working memory and consciousness is shown by the role posited for the **central executive** in voluntary attention. This will be discussed below.
Baars (1988) has published a theory of consciousness which is related to theories of working memory. He calls his central idea "the global work space" with which he identifies the contents of consciousness. This is a special sort of memory which acts as a central information exchange for correlating and integrating information from many sources. It is connected to many specialist processing systems which are unconscious and operate in parallel. These can either cooperate with, or compete for, access to the global workspace. He considers that the role of attention is to provide control mechanisms for access to this work space (thus controlling the contents of consciousness). He believes that some but not all of the contents of short term memory are conscious. A crucial distinguishing component of Baar's theory is the idea that consciousness is analogous to a cerebral information broadcasting system which makes selected output from specialised automatic processors generally available for other processors (such as the language system) to use. Only selected and meaningfully integrated information is made generally available in this way. According to the theory, such co-operative shared information has a special role in learning to solve new problems and in responding to novel events. Coincidentally, the global work space theory is consistent with the hybrid image theory of sketch perception (Chapter 5) if it is assumed that the gate which selects sensory or memory derived image components is also a gate between unconscious and conscious representations. Then the gate in my postulated triple buffer system is a gate between unconscious working memory and Baar's "global work space". The Cognitive Monitor theory of 7.2.9 is also broadly in agreement with Baar's theory except that it places a greater emphasis than Barr does on the sequential and simultaneous recording of cognitive events for control and back-tracking.

7.2.8 Control Theories of Consciousness

Tim Shallice (1972, 1978, 1988) has suggested a computational model (based on cybernetic analogies) which postulates that consciousness is a supervisory system which can activate and set the goals for modular automatic action systems. Shallice's "action systems" are postulated to be specialised processing modules which connect perceptual systems to "effector" systems. The latter are automatic neurally implemented programs for performing either motor actions or, more importantly for this chapter - internal processing, and which can run in parallel. Action systems which are also parallel and automatic, activate the effector systems, providing them with specific goals and parameters to modulate their operation. Action systems are compared to hierarchically organised computer programs with a main action system activating multiple sub-subservient action
systems, which in turn may control multiple effector systems. Conflicts between action systems are handled by an automatic "contention scheduling" process by which the most strongly activated action system inhibits others and thus becomes dominant. Thus only one action system (and by implication one goal) is strongly activated at a time avoiding the problems posed by many goals which may conflict. (I assume that the "dominant action system" is similar to the concept of "focal" attention referred to by Treisman, Posner and others, though Shallice does not make this clear). Shallice distinguishes two types of input to an action system. A "specific input" which is necessary for its on-going operation and a "selector input", consisting of an activation input which can initiate it, and an input which sets its goals. Consciousness, according to Shallice's model, provides the "selector input" for the dominant action system. It includes a short term memory system to store the "selector inputs". Brilliant as Shallice's model is, I feel it is incomplete in at least two respects.

1. It surely is not sufficient to "set an action system a goal". One needs to specify also how the goal is represented and to explain how the results of an action system's operation are compared with the goal and parameters in the stored selector input. Some kind of comparison process is needed which can be used to modify an action system's parameters in order to correct and re-run an action which has not achieved its goal. Shallice (1972, 1988) is rather vague on this point although it is implicit in his model.

2. Shallice's model concentrates on the output aspect of cognition. As Shallice himself admits (1972) it does little to account for perceptual awareness. Shallice provides some tantalising hints that visual awareness (in Hilgard's "passive mode") is related to an action system which "maximises perceptual input from a part of the world determined by its selector input, combining both motor movement (e.g. eye movements) and resetting of the perceptual system". I assume that by this Shallice means that perceptual awareness is the result of a perceptual "action system" which requires continuous monitoring and revision, using stored selector inputs. The "Cheshire Cat effect" (Chapter 5) would be an example of "contention scheduling" between rival perceptual "action systems" from the two eyes with the supervisory system's selector input determining which ocular representation is dominant and conscious.

In a classic paper Norman and Shallice (1980) have developed the Shallice model to explain the difference between "willed" and automatic behaviour. This will be discussed in the next section. The concept of an action system is related to the concept of memory schemas (see Rumelhart & Ortony, 1977 and Chapter 4). Its application to sketching is also discussed later.
Johnson-Laird's (1983, 1988) computational theory of consciousness is along similar lines to Shallice's but more generalised and with a greater emphasis upon self-awareness. Johnson-Laird compares the functions of consciousness to those of the operating system of a digital computer with a massive number of semi-independent dedicated processors which operate in parallel. Turing's proof that any computation that can be performed by parallel processors, can in principle also be performed by a serial computer is one of the bases of the Functionalists' belief that mental behaviour can be usefully modelled with algorithms that run on machines quite unlike the human brain in many respects. Johnson-Laird is a functionalist. The brain is massively parallel in most of its information processing mechanisms. The chief advantage of this is speed. Johnson-Laird argues that there is a division in the brain between a high level operating system and a large number of dedicated, hierarchically structured parallel processors. (It seems to me that Johnson-Laird's parallel processors are virtually identical to Shallice's "action systems". Following current modelling fashion they could also be thought of as self-adjusting neural networks. They are also similar to Fodor's (1983) automatic cognitive modules). Once triggered and given input data, goals and parameters, these parallel systems run automatically and are unconscious. Only the results of their computations are accessible to consciousness. Despite its many advantages parallel processing generates special problems of supervision and control. "Deadly embrace" can result if two systems are mutually dependent upon each other's output. The timing of the output from several parallel systems needed by another system may have to be carefully co-ordinated. Sequential processing is of course forced when the output from one "action system" is needed as input for another. As in Shallice's model, a system of multiple, parallel hierarchical action systems, is assumed to need an operating or supervisory system (which is identified with consciousness) which can coordinate the timing of action systems, has access to their results (but not their processes), has access to a model of intentions for setting goals and can "monitor" other systems. In his own words "These advantages (of parallel processing) depend upon a hierarchy of control governed by an operating system that can monitor the other processors, and on limiting its interactions to the exchange of information so that it is unable to gain access to the internal operations of other processors. Our introspective conceptions are accordingly very different from their unconscious counterparts; we tend to force intrinsically parallel notions into a serial strait-jacket." Probably Johnson-Laird, Shallice and Baars would all agree that consciousness is related to selective attention and that it depends upon limited capacity sequential processes and a special purpose short term memory.
The trouble with both the control theories of consciousness, as presented so far, is that they do not pin-point exactly how the processes of conscious awareness differ from unconscious processes. Why could not the operating systems described by Shallice or Johnson-Laird (as in a machine operating system) be, at a higher level, just as automatic and unconscious as the systems they supervise? Why do we have to be aware of the operating system and what in any case does that mean? The answer, according to Johnson-Laird lies in the capacity for self-awareness "which underlies the unique subjective experience of consciousness" (1983, page 470). He explains that the human (or primate) operating system differs from that of any existing machines in that it uses for its control and monitoring operations a model of the self which includes (by implication) sub-models of the mental processes that it models. This model, Johnson-Laird states must be recursively embedded allowing, in principle, infinitely higher levels of self-reflection (as I tried to illustrate in section 7.2.2). In practice, he suggests, the degree of self-reflexive recursion is limited by the capacity of short-term memory. However, as other theorists have pointed out, the conclusion that self-awareness is sufficient or necessary to characterise conscious awareness is very debatable. It is easy to agree that a model of one's self is a necessary component of human thought and also that models of individual cognitive processes are necessary components of conscious control. However it is not clear, to me at least, why a model of the self needs to be either explicit or conscious. Could not the posited model of the self, or self awareness, be simply the combined influence of the immense store of buried episodic memories of past experiences that can be drawn upon in fragments when needed? More to the point, why should the use of such a model in the monitoring and control of high level actions distinguish conscious from unconscious mental events? Important though it is, I am not happy with the choice of self-awareness as the key distinguishing characteristic of consciousness. I have therefore chosen to discuss separately another candidate characteristic of consciousness which I consider to be especially relevant to sketch function.

7.2.9 Consciousness as a Cognitive Monitor.

The word "monitor" appears repeatedly in theoretical discussions of consciousness (e.g. Johnson-Laird, 1983, Marcel, Bisiach, Umlita in Marcel & Bisiach, 1988, Edelman 1989) and for good reason. However, it is rare to find a clear non-tautological explanation as to what monitoring in this context means. In the above control theories and others (e.g. Umlita, 1988), conscious processes differ from unconscious ones in that they are "monitored". Presumably this implies that, (in contrast to automatic processes which run to completion without
7. Metacognition and Conscious Control

interruption and in a predetermined way), the output from a monitored process is continuously recorded and compared with a similarly represented, desired or anticipated goal by some hypothetical higher level "monitoring" process. Moreover it must be possible for the monitoring process to guide the monitored process by continuously adjusting its input parameters, using feedback from its output. Both Umilta (1998) and Bisiach (1988) stress the importance, in conscious control, of the capacity to interrupt a process once started. However a feedback and control model of monitoring is not sufficient to distinguish "conscious" from "unconscious" since many automatic or semi-automatic processes (unattended steering of a car for example) imply such control. A self-guided robot such as "Shakey" (Boden 1987) can make a plan for moving from one room to another and then correct its behaviour if it encounters unexpected obstacles. Yet no one would claim it was conscious. However if we are to avoid homunculus theories of a mind "monitoring" the mind A.I. models of monitoring may provide useful metaphors.

For example, Sussman's early (1975) A.I. program HACKER is often quoted in relation to conscious control (Marcel 1988, Shallice 1988, Oatley 1988). HACKER was intended to model skill acquisition and could learn to solve brick stacking tasks with very few ground rules. It had a rudimentary ability to "acquire skills", readjusting its operating rules by re-writing parts of its own code in response to task failure. Cognitive theoreticians point also to its two modes of running sub-routines - a "normal" mode and a "careful" mode which is used after its mistakes. In "careful" mode a detailed episodic record is kept in a special memory system known as "chrontext" on the state of its task world and on its actions and their effects. In careful mode chrontext can then be used to monitor, line by line, the success or otherwise of revised program rules. Shallice (1988) comments that "If human episodic memory had a similar evolutionary function then the basic input to episodic memory would need to be the schemas (action systems) selected in contention scheduling (action conflict resolution) with their bound variables [and maybe, in addition, some state-of-the-world information]". Shallice's further explanation of this important idea is not very clear. I therefore propose to develop it briefly in order to clarify the nature of "cognitive monitoring" and to show why I believe it is especially relevant to sketch function.

I suggest that a special purpose dedicated short term episodic memory for cognitive actions is a crucial, perhaps the crucial, distinguishing component of conscious mental events. It behaves as a sequential "cognitive event recorder" which combines the properties of memory and process controller. I will call it the "cognitive monitor". It is similar to and influenced by Shallice's "Attentional Supervisor", Johnson-Laird's "Operating System" and the Baddeley and Hitch
"Central Executive" (as described by Baddeley 1986, 1990, Hitch 1980). At the implementation level, I would expect it to be associated with Posner and Rothbart's "Anterior Attentional Network", especially with the anterior cingulate gyrus & sulcus which as already mentioned is a region of the cortex linked to voluntary control by neuropsychologists. I would also expect the cognitive monitor to have strong connections (in the brain) to the hippocampus which is known to be necessary for the storage of longer term episodic memories. I assume that the cognitive monitor provides very short term, active records of cognitive events which are sometimes but not always encoded passively (via the hippocampus) as longer term episodic memory fragments. I have only reluctantly chosen yet another name for such a conscious control system aware that there are already too many. However whilst acknowledging the debt this proposed "cognitive monitor" owes to all the above authors my inferences about its function, described below are expressed differently. I hope but cannot be sure, that the authors mentioned above would agree that this is roughly what their theories imply. If not then the following represents yet another personal view.

The cognitive monitor supervises an action schema (Shallice, 1988) which is a stored representation of a plan for performing a mental task. Following Johnson-Laird (1983) it is assumed that the schema is a complex hierarchy of possible routines and sub-routines many of which can be activated in parallel. Many possible paths are possible through the schema and its operation and output depends upon the parameters provided to the routines at different levels in the hierarchy. A highly learned and much used path through this hierarchy of neural processes will run automatically without supervision to its final outcome with subroutine triggers and parameters supplied by default. Such a process is like a train running through a complex railway system in which all the signals and points have been set in advance. It is very fast and efficient because it operates without any interruptions or checks. On the other hand its output is stereotyped depending only upon the initial input and starting parameters. Because no record is kept of its intermediate inputs and results it operates unconsciously. According to this theory conscious awareness depends upon access to a record made by a dedicated recording system which must be simultaneously activated when an action is performed. When we are not conscious of our mental actions this is simply because we have no short term memory of its intermediate goals and results.

The cognitive monitor has two posited roles which normally operate in tandem, a passive recording role and an active processing role. In its passive role it records for later use what is going on whilst an action schema is being activated. Its records must include :-
1. Neurally implemented codes which signify which cognitive routines (Shallice's action systems) are being triggered and when (in what sequence). I assume (as does Shallice) that routines which result in overt motor acts and routines with purely internal processing functions are equivalent as far as the cognitive recorder is concerned. The collected records of these codes describes one of many possible paths through the hierarchy of processing routines in an action schema.

2. The input parameters and data which modulate and quantify how a particular routine works. (How far and in what direction to move an arm for example or where and how large to project a mental image)

3. The stage by stage results of triggered routines either in terms of overt movements or internal information. The format in which action results are recorded is an open question and is probably mode dependent. I suggest that for visual processing the results would be recorded as sequences of percepts, mental images or percept-image hybrids which as described in Chapter 5 are posited to share a common format. That is they will be represented as sequential 2½D spatial maps linked to hierarchic descriptive labels. Thus when consciously moving a pencil whilst drawing a difficult line, we are briefly recording a cognitive "video film" of the hand and the pencil as it moves. Evidence that the brain does record and sometimes store for long periods such temporal (moving) sequences of images is provided by reports made by brain surgery patients of vivid imagery experiences of past events when certain parts of their exposed cortex are stimulated (Penfield and Perot, 1963). This briefly stored cognitive "video" sequence is also used, I suggest, for integrating the output from successive saccades (Chapter 4) and is the representation underlying visual awareness. (We are not aware of single saccadic "snapshots" but of the results of integrated sequences).

I will call a stored sequence of such triple records related to a conscious action a "cognitive record". A tenet of the theory proposed here, is that sequential sketching acts support the generation of cognitive records. As generally agreed by the authors I have quoted, I also assume that in making such Cognitive records, the monitor does not have access to most of the details of the many routines and sub-routines of a particular action schema. However it does have access to the triggering conditions and the partial results output by many of the sub-processes within an action schema. It should be noticed that this model suggests an extra function for the gate which allows, for example, only one of two rivalrous images (the Cheshire Cat effect or my postulated percept-image gate - Chapter 5) to be conscious. The gate is a mechanism for choosing between alternative perceptual "results" (unconscious conclusions) for entry on the cognitive record.
In its active role the proposed **cognitive monitor** must have the following processing attributes:

4. It must have access to a stored representation of goals related to the action schema currently monitored. This would include a general model of intended outcome plus numerous sub-goals attached to specified sub-processes. Moreover the goals must represented in the same format as the results. If the results have been represented as mental images, so also must be the goals. If the results have been recorded as verbal propositions, so also must be the goals (e.g. "I am trying to answer a question about ----"). It seems likely that one way of representing schematic intentions, is to use as a prototype a stored mean of successful past cognitive records, but with alternative parameters for varying sub-goals. The idea that mental images of an intended outcome can be used as goals was suggested by James, W. (1890) and has been developed by Baars (1988,1993) as "goal-images". The model described in 5.3.7.2 for matching percepts and images can also be used to explain how perceived (or imagined) results in visual problem solving might be matched with an unconscious "goal-image" before entry to the conscious **cognitive recorder**. Evidence that mental images can become unconscious as a task becomes over-learned is provided by Pani (1982), using a percept-image matching task. Pani also showed that the mental image can become conscious again if the task difficulty is suddenly increased.

5. It must have access to a model or schema of the action about to be monitored. This is a concession to Johnson-Laird's "model of the self" in that it assumes in effect that when we tackle a mental task we have a model with default parameters (as in "frames", Minsky 1975) based on past experience of the structure of the processing tree itself. The model is used as a prototype for selecting alternative sub-processes at processing nodes and for setting sub-goals.

6. The cognitive monitor must include a mechanism for comparing the stage by stage results of an action schema with the corresponding goals and for comparing the processing path with the **action model**. The basis for such a comparison would vary greatly with the type of action and might involve eliciting other outside procedures such as image inspection or memory retrieval. The results of the comparison process must lead to a decision about the degree to which a specified goal has been or is being achieved and whether or not to deviate from the **action model**.

7. It must have a mechanism for **interrupting** action schemas (see Umilta, 1988)) and for using the results of the above comparison process to do one or several of the following, (either **after** an interruption or continuously whilst the action schema is active).
a. change the operating parameters of a pre-selected sub-process
b. switch to an alternative sub-process amongst those available at the nearest node (branch point) in the schema hierarchy
c. Reset one or more of the goals or sub-goals of the action schema.

I am trying here to suggest a process which goes beyond the type of automatic back propagation learning that can be simulated with neural net computations for example. The purpose of the cognitive monitor (or supervisory attention or central executive) is to modify the behaviour of a highly practised set of processes, in conditions when it breaks down due to unexpected input data or unfamiliar goals. In these cases the intervention of an external higher level processor is necessary.

Although I have for logical clarity separated the recording (passive) and control (active) functions of the cognitive monitor it is possible that its neural implementation in the brain could combine memory and processing in the same structures (Crick, 1994). Thus the records of input triggers might also be used to activate the processes they represent. In which case the analogy with Sussman's HACKER re-writing parts of its own code is appropriate. Note that it would be neither efficient nor necessary to store fully explicit mental images. Only the image's source location in memory and its generating parameters would be necessary in the cognitive record, since, by definition, the cognitive monitor has access to the visuospatial components of working memory used to house images and the image generation processes. (Logie (1995) has suggested, from recent data on dual task performance, that the central executive is involved in image generation and manipulation.). However there is one reason for assuming that the recording and control functions of the monitor, even if they have a shared format, can be separated. It would offer an explanation for the Hilgard distinction between passive and active consciousness or between awareness and volition. With passive awareness (as in unmotivated visual awareness) the monitor would be exercising its record functions, (making selected perceptual constructs available for other monitored action schemas such as speech) without using its control functions to interfere with the processing path. The monitor could also make cognitive records to a varying depth in the tree of processing possibilities, allowing independently varying degrees of automaticity, awareness and voluntary control. This is a consequence of the model. Whether in reality there is such a thing as passive awareness or whether it is will be found that apparently involuntary awareness conceals (weakly recorded) control functions, is in my opinion, an unresolved empirical question.
Clearly there are many disadvantages to processing with an activated Cognitive Monitor when implementing an action (or thought) schema. By any description what it would have to do is complicated and must be "expensive" in processing facilities and memory capacity. This explains the limited capacity of consciousness and is one of several reasons for the limited span of focal attention. Moreover the recording and process "re-setting" functions of the monitor would take time so a consciously monitored "run" through a schema hierarchy would be slower than an unconscious run though a preset path with the monitor switched off, even if there was no parallel to sequential conversion (as in controlled versus automatic visual search). Also because the monitored process must learn its way by trial and error through the many possible paths in the schema the monitored operation will be erratic and less certain in its outcome. For these reasons, as soon as a particular monitored path through the schema is regularly repeated with successful results, the path and its parameters are stored, to be activated habitually as required with the monitor switched off. The penalty for this is that such automatic processes cannot adjust to novel input or to unexpected goals. The constituent sub-processes cannot be re-arranged or re-written to learn new or inventive procedures. Because there has been no "official" episodic record of the procedural stages the intermediate results and parameters of an automatic process are not available for other controlled processes such as verbal reports, mental judgments or problem solving, to use as input. Many mental processes, such as those involved in the early stages of vision and those used for long term memory retrieval, have no access to the cognitive monitor. Only the end results of their paths through their habitual action schemas are conscious. However some of the partial results and sub-processes of mental action schema acquired by and available to artists can with difficulty be monitored. They are open to metacognition.

The above explanation attempts to summarise what I hope is a fair inference about the implied meaning of "monitor" "central operator" etc, synthesizing and extending a little the views of some principal theorists, in a way with which I hope they would not disagree. The above model attempts to relate theories of consciousness as a control system with theories of conscious awareness. Thus I believe it to be consistent with Marcel's (1983b) discussion of perceptual experience as a process of top-down "recovery" of the results of automatic, unconscious visual representational systems which are not themselves available to consciousness. This "recovery" proceeds in the opposite direction to what would be expected by any theory which posits that perceptual experience and the immediate results of visual processing are identical (Marcel calls this the "identity theory"). According to Marcel (1983) what we are "aware" of when we see an object or
scene are not the results of the visual "action schemas" (using Shallice's vocabulary), which structure the stimulus and access memory for meanings and identities, but a synthesis constructed by recovery after these processes have occurred. I am simply suggesting that records from the proposed cognitive monitor are candidate structures available for this purpose. Marcel's experiments suggest that the last items to be recorded - high-level meaning and object identity are the first to be recovered, and the results of first processes (local shape and colour) are the last to be used (see Chapter 4).

7.3 Automatic and Voluntary processes.

7.3.1 Criteria for Automaticity.

Before discussing what use such cognitive monitoring is to inventive thought and how it is supported by the sketch, it is necessary to consider further the distinction between automatic and voluntary processes.

Originally the conceptual distinction was made in relation to theories of attention. Posner and Snyder (1975) proposed that the following criteria be used to distinguish an automatic process from one that reflects the operation of conscious attention:-(1) it should not lead to conscious awareness (a concealed tautology ?) (2) it should occur without intention (3) it should not interfere with other concurrent processes. Experimental confirmation of the distinction was provided by Neely (1977) in a study of the priming effects of a category name followed by a word which was either related or unrelated to the prime category (as "bird"-"robin") and which belonged to an expected or an unexpected category (as "bird"-"window"). The results suggested that semantically similar words automatically activate long term memory and that false expectations cause conscious interference with the automatic process. The evidence for automatic memory access is relevant to sketch theory for two reasons:

1. Theories of recognition suggest that sketch object parts access memory automatically, without control and can therefore have unintended or unexpected visual retrieval results. This suggests a mechanism by which accidental sketch stimuli can have serendipity consequences for inventive thought (Chapters 2 & 4). Uncontrolled mental processes can have constructive consequences (7.4.2).

2. Although an automatic process may not interfere with another one its results can. Thus, when task irrelevant information is present whilst sketching, it may cause automatic retrieval interference which cannot be avoided by conscious attentional filtering (6.3). Clearly we need criteria for distinguishing sketch stimuli which are likely to cause automatic interference from those that can stimulate visual thought.
7.3.2 Learning and Automaticity

However more recent work has shown that the above distinction is not clear cut. For example the Stroop interference effect (6.3) was believed to demonstrate the effects of automatic memory access to word meaning. There is now evidence that focal attention can also influence the effect (Kahneman & Treisman, 1984) and it appears that there are many degrees of automaticity. Some processes are always automatic and others acquire varying degrees of automaticity by practice and learning. This is predicted by the cognitive monitor model. The monitor is assumed to record and control an action schema by acting at many branching nodes of a hierarchic tree of processes and sub-processes. The degree of automaticity would be related to the level within the hierarchy at which the monitor has no further access. Thus a completely automatic process would be initiated at the main root of the tree and only the final results recorded. A partly automatic process would result if monitoring ceased at certain processing levels or was forbidden at particular nodes. Any degree of partial automaticity is then possible with some sub-processes and sections of the processing path running automatically and others controlled. The consensus view is probably that prolonged practice causes a process to become automatic to varying degrees and that automaticity is far from being an all or nothing attribute. Acquired automaticity has the advantage that it allows a learned process to be performed concurrently with others. However studies of divided attention suggest that there is always some residual interference when two highly practised tasks are performed concurrently (Baddeley, 1986 Logie, 1995). Difficult, novel tasks on the other hand require a limited "central capacity" and are more likely to interfere with other concurrent processes (Baddeley, 1990 page 123).

Thus whilst an experienced driver can simultaneously talk and drive, it is much harder when one is first learning to drive. Moreover as soon as a novel event occurs, such as a child suddenly running across the road, then even the experienced driver would probably stop talking and concentrate on a (possibly novel) procedure to avoid an accident. Shiffrin and Schneider (Schneider & Shiffrin 1977, Shiffrin & Schneider 1977) studied the effects of learning on the speed of visual search. Subjects were asked to memorise 1-4 letters (the "targets") and then had to decide as quickly as possible whether a member of the learned set was present in a presented stimulus array of 1-4 letters. For example, in one series of experiments the targets were consonants from the first half of the alphabet and the background items (the "distractors") were consonants in the second half. At first the time to detect a target letter increased linearly with the number of distractors suggesting a controlled sequential search. Both response time and accuracy were affected by the
number of targets and the number of distractors. However Shiffrin and Schneider then gave their subjects continuous practice. After 1,500 trials they had become very fast and accurate and there was no effect of either the number of targets or the number of distractors, showing an acquired parallel search mechanism. However if at this point the targets and distractors were switched then the effects of the learned automaticity was very detrimental and performance was worse than when they had just started. Even after 2,100 further trials they were still not up to the level of performance they had reached when they had switched.

Within the limited process of visual search Shiffrin and Schneider's results demonstrate clearly both the advantages and the disadvantages of acquired automaticity. The effects of practice result in a fast and efficient process that operates equally smoothly with large or small target and distractor sets. In contrast the controlled sequential search seems slow and inefficient. However as soon as the task context is changed the inflexibility of the acquired automatic process becomes a positive hindrance to learning the new context. In contrast the controlled processing used when learning the task is flexible and relatively unaffected by the altered conditions of the task.

Clearly the mental processes used by artists and designers for difficult visual invention and problem solving are of a different order of complexity from those used in simple visual search. Nevertheless it is likely that successful artists acquire through experience habitual (and effective) thought processes which characterise their individual style but which also possess some of the disadvantages possessed even by partial automaticity for handling novel situations or learning new approaches to difficult problems.

7.3.3 Habits and the Control of Action

As already mentioned Norman and Shallice (1980) have developed the Shallice (1972, 1978) model of consciousness and used it to explain the difference between willed and automatic behaviour. In the model the output from sensory or perceptual systems passes first to a "trigger data base" which in turn activates the "source schemas" which are similar to the "action schemas" described in section 7.2.9. The function of the trigger data base is to represent and to store the combination of trigger conditions (sensory inputs and inputs and parameters for other internal processes) necessary to activate a given schema. Multiple sets of trigger conditions are assumed, corresponding to multiple action schemas which can be activated sequentially or in parallel. In the model, decisions about when two or more actions can be safely performed, in parallel or co-operatively, and when they cannot be because they are in conflict, is handled automatically by a
"contention scheduling" system. This processor assigns priorities and sees that conflicting, lower priority actions are temporarily inhibited. This system is described in terms of "horizontal threads". Superimposed on the contention scheduling and action schemas as a "vertical thread" is a "Supervisory Attentional System or (SAS)" which represents the influences of motivation and attention and can alter the "activation values" of the action schemas. Unlike the proposed Cognitive Monitor, the SAS does not have a recording function (although see Shallice 1988 quoted above). The authors explain that there are two levels of control with the higher level Supervisory System biasing the selection of actions by the contention scheduling system. "The result is three modes of the control of performance: automatic, contention scheduling without deliberate direction, and deliberate conscious control. Will becomes the application of attentional resources to the control of action". (Norman & Shallice, 1980). The model is important because it is the first to offer mechanisms (although rather general ones) to account for the complex mixture of automatic and consciously attended actions that occur in everyday life. Completely automatic behaviour occurs, Norman & Shallice suggest when we perform a well learned action, such as brushing away an insect, without awareness of either the action or its initiation. However they suggest it is also possible to be aware of performing an action "without paying active, directed attention to it". This they claim corresponds to situations in which the contention scheduling system handles the control of action without the involvement of the Supervisory Attentional mechanism. Such actions, containing awareness without conscious control William James, (1890) called "ideo-motor acts". There are surely many Ideo-Motor acts occurring when a skilled artist sketches.

Norman and Shallice apply their model to a number of phenomena. One is the ability of highly trained people to perform two apparently demanding tasks at once. For example highly skilled pianists have been shown to be able to sight-read and play music at the same time as shadowing spoken speech (Allport et Al. 1972). The Norman and Shallice model explains this learnt ability as the running of two practised schemata by the low level contention scheduling system, with little load on the Supervisory Attentional System. Another is absent-minded "slips of action" (well known to me personally) which have been documented and analysed by Reason (1979). Such slips are labelled as "capture errors". The Supervisory System, occupied by a different task, fails to over-ride the contention-scheduling system which triggers a schema appropriate to the perceptual input but inconsistent with the long term plan, which is presumably the province of the Supervisory System. Action slips are mentioned here briefly because it is worth considering the likelihood that action slips of thought without obvious overt consequences but
damaging to creativity, may frequently occur unnoticed during the planning stages of a visual task. Clearly the danger is greatest when two difficult visual tasks, or two aspects of the same task are competing for the capacity of the SAS (or Cognitive Monitor) at the same time. As Norman and Shallice point out, with slips of overt behaviour one's friends or colleagues will act as surrogate Supervisory Systems pointing out the error. However private action slips of thought when for example the wrong part of an image is automatically processed whilst contemplating a visual idea may be harder to detect. This reflects back to the theme of Chapter 6 which documents the evidence that irrelevant attributes may be automatically processed in an over complex representation of a visual task.

7.3.4 The Central Executive and Random Generation.

In Chapter 5, the Baddeley and Hitch model of working memory was outlined briefly to discuss the role of its sub-components in visual perception and mental imagery. The super-ordinate component of the model is the central executive which despite its importance is not well understood. In Baddeley's own words it has tended to become a vague explanation for a rag-bag of mental processes that seemed to share the requirement for some general super-ordinate non-modal working memory capacity. In a fascinating discussion (Baddeley, 1986, 1990) has suggested that the central executive may be equated with the Supervisory Attentional System and "perhaps with consciousness". One of his reasons is that it seems to explain some old but hitherto unexplained experimental data on how people generate random sequences of numbers and letters. This apparently trivial task is actually quite difficult. The experiments he describes (1986, pages 229-234) have a bearing I believe on the processes used in visual invention and how they might be supported by visual media. In a series of experiments Baddeley asked his subject to generate random sequences of letters or numbers in a variety of conditions. In a typical experiment 12 subjects were asked to generate sequences of a hundred random letters at four rates. Three different measures of randomness showed the same results. There was a neat linear relationship between the degree of randomness and hence the amount of information generated and time allowed for a response. At the time Baddeley was puzzled- "The results were all that I might have hoped for and yet how should one explain them?"

In terms of the Norman and Shallice model an explanation is as follows. The supervisory System is attempting to control a set of memory retrieval processes. As already discussed these tend to be automatic. They are likely to produce sequences which are strongly associated in memory such as alphabetic sequences and common acronyms. "Indeed the very process of retrieval is likely to
strenthen such associations producing an increasingly stereotyped output. to prevent this the subject must perpetually intervene to break up developing patterns and to favour weak schemata, a task which requires him to monitor the occurrence of prior responses and to attempt to maintain some kind of long term strategy."

Implicit in the Baddeley (1986) discussion of the common attributes of the Supervisory Attentional System and his Central Executive is the suggestion that it (the Central Executive) must represent both control functions, being able to activate and interrupt other mental processes and a general purpose short term memory. However he does not spell out exactly what must be stored to control automaticity, as I have tried to do, in outlining the functions of the Cognitive Monitor. For example it is implicit in his account of random generation, which get harder after about eighteen items, that a monitoring memory for sequences already generated must be available for inspection before the next item is selected. This would be contained in the cognitive record proposed in section 7.2.9.

Such mental monitoring is also implied, I suggest, when an artist is seeking a new solution to a problem of visual structure and must make a conscious effort to avoid adequate but uninteresting solutions, relying on highly practised but stereotyped thought patterns.

Having reasoned this way Baddeley went on to show that tasks known to be associated with the central executive, such as speeded card sorting and memory for chess positions interfered with random number generation in ways which suggest that they compete for the same limited capacity resource.

7.3.5 The Frontal lobes and the Dysexecutive Syndrome.

Other evidence that the central executive is involved in volition and is necessary for the control of inflexible and stereotyped behaviour comes from the study of the disabilities of patients with damaged frontal lobes. Evidence that a part of the frontal lobes (the cingulate gyrus) is involved in controlled attention and in volition has already been mentioned. These recent studies are consistent with numerous earlier studies of the variable and rather mysterious deficits showed by patients with less well localised lesions in the frontal lobes. One particular class of deficits Baddeley has labelled as the Dysexecutive Syndrome, because the symptoms suggest a deficit in the Central Executive. Such patients tend to be deficient in judgement and initiative and have great difficulty tackling novel problems which would be relatively easy to other patients (Luria, 1969). Shallice (1982) has shown that they perform very badly on tasks which require forward planning. A characteristic of the syndrome is a lack of flexibility and the tendency to persist with a pattern of behaviour or thought after it is been shown to be
inappropriate. In the **Wisconsin Card Sorting** test, the subject is given a pack of cards containing various numbers of different types of symbols in various colours. The task is to sort the cards according to a given dimension and then when that has been accomplished to sort again on a second dimension and so forth. Frontal patients find this extremely difficult. In general they often have difficulty both in initiating new activity and in discontinuing an activity once started. This produces the apparent paradox that on the one hand they are easily distracted and on the other hand show unnatural perseverance. This would be explained if the Supervisory System (Central executive) is not functioning properly. The patient cannot then interrupt or redirect an activity controlled by the contention scheduling system and has difficulty starting a new one. According to the Cognitive Monitor hypothesis the patient would also be missing the cognitive records necessary to back-track in order to correct mistakes (and is unaware of faults in the active schema). Using a battery of tests, including random generation on subjects of two different age groups, (including as subjects his own relatives) Baddeley (1986) provides evidence that the Central Executive is a component of working memory that is especially liable to decline in performance with increasing age.

In summary, there is evidence that a dedicated processor with limited capacity, probably involving the frontal lobes or anterior cingulate gyrus is necessary for high level planning and for monitoring and re-learning well practised but partly automatic actions and thought processes. An obvious deduction is that original invention and problem solving in the visual arts will put heavy loads on such a cognitive supervisor in addition to loading the visuospatial and visuo-descriptive memory systems discussed in Chapter 5. However we must not underestimate the power and creative complexity of highly practised but unconscious processes in the visual arts. As an artist or designer becomes older the capacity of his or her central executive may decline (according to Baddeley) but the store of unconscious visual thought patterns will become increasingly larger, more interesting and more appropriate to an individual’s style and skill as experience enriches the long term store of tacit knowledge. One must therefore expect to find a cognitive balance in the working practice of experienced artists or designers between the conscious monitoring and interruption of habitual patterns of thought and the fast and efficient triggering of automatic, but stylistically interesting, visual routines (Ullman, 1984) acquired from long experience within a specialised domain.
7.4 Consciousness and Creativity.

7.4.1 Hofstadter's Notion of "Sphexishness"

Part of the inspiration for this chapter comes from one of Douglas Hofstadter's entertaining essays entitled "On the Seeming Paradox of Mechanizing Creativity" (Hofstadter 1985a). Until now I have avoided discussing "creativity" as an aspect (or function) of sketching because the term seems to cover too many different aspects of the human mind. I have preferred instead to refer more specifically to "invention" and visual "problem solving". However Hofstadter's discussion of the role of consciousness in creativity is so relevant to the role of metacognition and the central executive in sketching that it is worth briefly considering the literature on creativity in a wider context. Usage here follows Bruner's (1979) definition of creativity as "An act that produces effective surprise". This is a succinct statement of a consensus view that creative thought must fulfil the two conditions of (1) being useful (effective) in the sense that it fulfils a need, solves a problem or answers a previously unanswered question and (2) contains an element of newness or unexpectedness (e.g. Sternberg, R. 1988). In her book "The Creative Mind" Margaret Boden (1992) distinguishes two senses of creativity. In the psychological sense ("P-Creativity") creative ideas or discoveries need only be novel to the individual mind that produces them and it really doesn’t matter how many other people may have had the same idea already. In the historical sense ("H-Creativity") creativity applies to ideas or inventions which are "fundamentally novel with respect to the whole of human history". In the context of the type of private sketching that is the subject of this study, it is probable that P-Creativity matters more to a sketch user than H-Creativity. This does not imply that H-Creativity is not a very desirable consequence of the mental processes supported by the sketch. When sketching is used to support the users cognition in the early stages of visual invention it is likely that an artist is more concerned with her or his own private thoughts than with how new an idea is in the H-Creative sense. However the nature of sketch indeterminacies is such that, in principle, H-creativity may occur simultaneously with P-Creativity. For example, Kandinsky's "discovery" of abstraction in painting (see the quotation in 7.4.3), which could have been evoked by a sketch, might well be classified, with the benefit of hindsight, as H-Creativity, since Kandinsky can claim to be one of the inventors of abstraction in art. However it is clear, from his own words, that he saw the insight, initially at least, as only relevant to his own painting and hence as P-Creativity.

Whilst Hofstadter agrees that "Creativity is part of the very fabric of all human thought, rather than some esoteric rare, exceptional and fluky by-product of
7. Metacognition and Conscious Control

the ability to think", (Hofstadter 1985) (Boden's P-Creativity) he also considers that it is quite wrong to treat creativity as something (e.g. "the divine spark") that is beyond intelligence. The crucial component he believes is "the ability to watch oneself as one deals with the world, to perceive in one's own activities a pattern, and to be able to do so at many levels of abstraction". Although expressed in the concealed "homunculus" terminology that I have tried so hard to avoid, this is very close to saying that the Shallice- Baddeley Central Executive (my Cognitive Monitor) is the cognitive basis of creativity.

Hofstadter's comic term "Sphexishness" is a metaphor based on the behaviour of the wasp "Sphex". This insect shows a complex, sinister and apparently purposeful predatory behaviour. "It lays its eggs in the body of a paralysed cricket which it stores in its burrow for the grubs to feed on when they hatch. Its routine is to leave the cricket on the threshold, inspect the burrow, then come back to bring the cricket inside." However "if the cricket is removed a few inches while the wasp is inside, once again it will move the cricket up to the threshold and re-enter the burrow for a final check. The wasp never thinks of pulling the cricket straight in. On one occasion this procedure was repeated forty times, with the same result." Hofstadter lists twelve different examples of a continuum from the most sphexish to the least. It starts with a "stuck record" and ends with "Styles in art which become dated and routinized to the point of no longer being creative. This happens to every style, but at the moment of its happening, there are always some people who are breaking out of the rut and creating totally new styles." All human beings, thinks Hofstadter have an ability to spot "unanticipated patterns" of unanticipated types in unanticipated places and times. Whenever they "get into some kind of 'loop' they quickly sense it". The ability to break out of such ruts he implies is a function of consciousness or metacognition - the ability to record and control ones own thought processes.

7.4.2 Unconscious Knowledge and Creativity.

In contrast, the importance of unconscious or "tacit" knowledge for visual invention must not be underestimated. It is obvious that the unconsciously stored mental processes resulting from a life-time's practice are not only an advantage, they are a necessity, for most useful visual creativity. If this were not so artists would suffer from a double disadvantage as they grow older. Not only (as Baddeley tells us) would the monitoring capacity of their central executives be in decline, but also they will be accumulating ever more "Sphexish ruts" to bar their way to new invention. Although mathematicians and some poets may produce their
7. Metacognition and Conscious Control

best work when young, we still have to explain for example the inventiveness of Titian, Monet, Cezanne and Picasso well into old age.

Just as one is grateful that one does not need to think how to move one's pen each time one scrawls a signature, nor explain why the result is so easily recognised, so surely a concert pianist can concentrate on monitoring her or his interpretation of a Mozart concerto, knowing that the fingers movements will automatically follow their musically imagined goals. Whilst painters and designers may know that a certain colour relationship "works", they can rarely make the decision process conscious or explain why a minuscule colour difference can destroy such a relationship. Hofstadter tacitly recognises this when he admits that what he calls "mental ruts" are also unavoidable attributes of the inventive mind. "Each of us - even the Mozarts among us - exhibits a "cognitive style" that in essence defines the ruts we are permanently caught in... Far from being a tragic flaw, this is what makes us interesting to each other. We celebrate individual styles."

Psychologists commonly classify long term memory into three kinds (Norman and Rumelhart, 1975). **Semantic** memory represents facts or beliefs about the world which are not tied to specific times and places. In the visual arts, knowledge of the tradition within which one works, of techniques and materials and of the work of other artists, are examples of semantic knowledge important to most practitioners. **Episodic** memory is the stored memory of personally experienced past events. One form of episodic memory, important to visual creativity, is the memory of one's own past inventive acts, their circumstances, their objectives and their degree of success. (I have already suggested that a crucial component of conscious awareness is a short term episodic memory for thought processes, but that this is not necessarily passed to long term memory). The third type of memory, **Procedural** memory is the memory about how to do things, and is normally associated with skills (such as riding a bicycle) which can only be acquired by repeated practice. All three types of memory refer to a large capacity long term storage that is unconscious. In order to make fragments of such information available for use, it must be "retrieved". Retrieval processes are themselves unconscious, but at least in the case of semantic and episodic memory the information "recalled" as a result of such retrieval must be conscious to be useful. Visual recognition is a form of semantic memory retrieval, which according to Hypothesis 2 (Chapter 4) is an important sketch function. Procedural memories on the other hand can be both retrieved and used without conscious control. A question pertinent to sketching is "Is tacit knowledge in the visual arts of the procedural kind and if so, can it be usefully made conscious?". When procedural
knowledge refers to explicit motor skills such as juggling or riding a bicycle, its legitimacy as a separate type of memory seems well founded. However when applied to the pattern of thoughts underlying the act of drawing, for example its meaning is less clear. Skilful draughtsmen appear to be using unconscious procedural knowledge in the way that they draw, since as with hand-writing, a common style can usually be recognised in drawings by the same individual (Rawson 1969). At the same time those who sketch are surely aware of each mark as they make it and are consciously in control of the important variables of the sketch. In the act of drawing, Semantic Memory, ("what does this look like?") Episodic Memory ("what did I draw yesterday?") and Procedural Memory ("why is an artist's drawing style so recognisable"), seem to be intimately combined in complex and poorly understood ways. A similar conclusion can be drawn about the invisible mental processes underlying visual invention.

We have seen that automatic mental processes, in contrast to voluntary ones have the advantages of speed and (providing they are not in conflict) of being able to operate co-operatively in parallel. What Hofstadter seems to have temporarily forgotten is that, when highly learned, specialised skills become automatic, they may operate in parallel at a lower non-creative level in the processing hierarchy of visual thought. By doing so they become valuable tools, which release precious but limited resources of the conscious monitor to be concentrated at a different level (usually higher) of processing at which conscious control is needed to support creativity. Thus, it is insufficient to ask "how does the mind break out of its ruts?". One must ask in addition "at what level of mental processing are automatic habits a handicap and at what level are such habits essential subservient skills". For example in the education of a Fashion designer the ability to draw the human figure, in a rather specialised way, is regarded as an essential skill. When such a skill becomes highly learned it "finger-prints" the individual style of a designer's sketches (Illustrations VIII, IX). However far from being a "rut" which hinders mental invention, the automaticity of such drawing styles releases cognitive attention where it is needed - to the form and fabric of the garments. It is at this level where habitual thought may be a handicap. Thus conscious control (in the form of concentrated attention) at one processing level, is dependant upon effective automaticity at others, or it will become diluted, as when we attempt to perform two unlearned tasks at once.

7.4.3 Conscious and Unconscious Thought Combined.

We need to understand how the act of sketching can support the "antisphexishness" of cognitive invention, without hindering valuable supporting
automaticity. The literature on creativity is full of anecdotal support for a role for unconscious thought in creative invention. Usually these relate to novel insights, or unexpected answers to difficult problems, suddenly coming to mind without any apparent conscious effort. Amongst the most widely quoted are Kekule's reported idea for the benzene ring suddenly occurring after dreaming (or day-dreaming) about a snake biting its own tail, the words of "Xanadu" coming to Coleridge in an opium induced reverie and Poincaré suddenly thinking of a fundamental mathematical series while boarding a bus (Vernon, P.E. 1970, Sternberg, R. 1988, Boden 1992). Most accounts of such sudden intuitions suggest that they are the result of unconscious processing which occurs after extensive conscious mental effort (on the same problem) has been temporarily abandoned. Poincaré in a widely quoted paradigm, distinguishes four phases of creativity (Vernon, P.E. 1970, Boden 1992). A initial process of preparation involves conscious effort to solve a problem by familiar methods. A second phase of incubation occurs unconsciously whilst the conscious mind is elsewhere. It may last minutes, days or even months. Poincaré suggests that during this phase ideas, patterns and analogies are being unconsciously combined with a freedom denied to conscious rational thought. This then results in a sudden illumination, in which the results of these unconscious processes become unexpectedly conscious. Finally a process of conscious verification is necessary to explore the new idea and build on its implications.

It is possible that the influence of Poincaré's model has exaggerated the role of unconscious processes. It may be that during the phase of "incubation" many conscious processes occur which are later forgotten when the sudden "illumination" occurs. Nevertheless there are so many analogous accounts, by both artists and scientists, of their inventive thought processes, that Poincaré's model deserves to taken seriously. However, at least one widely used example - Mozart's supposed account in a letter of how musical ideas come suddenly into his head - is probably a forgery (Boden 1992). Ghiselin (1952) quotes the American poet Amy Lowell - "An idea will come into my head for no apparent reason; "The Bronze Horses, for instance. I registered horses as a good subject for a poem; and, having so registered them, I consciously thought no more about the matter. But what I had really done was drop my subject into the subconscious, much as one drops a letter into the mailbox. Six months later the words of the poem began to come into my head, the poem - to use my private vocabulary - was 'there'". Besides a faith in the power of unconscious thought, what stands out in both Poincaré's and Amy Lowell's accounts is their metacognition, their high level of awareness of their own creative strategies.
Does the Poincaré paradigm of constrained conscious effort followed by unconscious and less constrained thought apply also to the visual arts? Arnheim (1962), in his introduction to his account as to how Picasso created "Guernica" suggests that it does. However he also emphasises the role played by a restructuring of the problem situation or what the Gestalt psychologists termed "productive thinking". If unconscious automaticity is necessary to facilitate the free "combination of ideas", then it is also necessary to consciously struggle for a clear image of the "goal situation". "Sudden flashes of discovery" may be partly the result of unconscious processing but "the evidence seems to indicate that these happy solutions do not appear unless intensive conscious wrestling with the problem precedes them. Rather than assume that special creative powers dwell below the threshold of consciousness, we might point to the premature freezing of orientations and connections that may occur in conscious thought.. " (Arnheim 1962).

Suddenly we have come a full circle, with the plausible but paradoxical implication by Arnheim, that conscious effort is needed to elicit unconscious processes, one of the effects of which is to avoid the very "sphexishness" which according to all the evidence is the product of unconscious habits. A possible resolution of this paradox lies in the fact that although the procedures used by unconscious processes may be automatic and "sphexish", the results that they turn up may be unexpected and novel. A candidate for such unconscious serendipity, suggested by Johnson-Laird (1988), is memory retrieval. Although retrieval has been thought to involve both voluntary strategies and automatic processes, the process which connects a retrieval cue with the stored information it accesses, is automatic. My own theory is that conscious efforts at retrieval may constrain the results of automatic and perhaps parallel memory search procedures, in ways which lower the chances of accidental serendipity combinations of retrieval products. Thus when attention is displaced from the retrieval process, novel results are more likely. A possible mechanism for such serendipity results from unexpected memory retrieval, is that the automatic memory search process "runs on" after a "find" in contrast to a conscious sequential search which stops when a required item is found (Shiffrin & Schneider, 1977). In a classic well-replicated study of automatic memory search, Sternberg, S. (1966) obtained memory search results which have never been well understood. Briefly, he found that the time to detect whether a target digit was a member of a previously memorised set of digits, had a neat linear relationship to the size of the memorised set, suggesting a process of sequential search. The surprising finding was that, unlike a search consciously attended to, the search times for found and not found target digits were identical,
suggesting that the search process continued to the end of the memorised set even after the target had been found. In computer programming terms it was as if the memory search loop had no "If found then exit" clause. Whatever the explanation for this finding, if a similar unconscious "running-on" occurred in more open-ended retrieval searches it might partly account for the otherwise mysterious unattended "Eureka" effect. (Sternberg, S.'s explanation is controversial. See Baddeley, 1990 for a discussion).

Automatic mental image generation may also have serendipity effects. According to the theory of hybrid-image formation outlined in 5.3.7, the preliminary sketch percept-memory image matching process is assumed to occur unconsciously. A sketch percept only becomes conscious after the percept-image matching and gating process has occurred. The conscious results of such automatic image processing may be another source of "antisphinxishness" in the sketch.

Sudden visual insight may be elicited by new arrangements of familiar representations. Thus James Watson (1968) "suddenly" (his words) saw how purine and pyrimidine bases could be paired, by playing with cardboard models of these molecular bases of cellular inheritance on a table top, (leading to the famous double helix structure of D.N.A). Similarly, as an example from the visual arts, Kandinsky described how in 1908 an accidental experience led to his sudden "discovery" of abstract art. "I was returning, immersed in thought, from my sketching when on opening the studio door, I was suddenly confronted by a picture of indescribable and incandescent loveliness. Bewildered I stopped, staring at it. The painting lacked all subject, depicted no identifiable object and was entirely composed of bright colour patches. Finally I approached closer and only then recognizes it for what it really was - my own painting standing on its side on the easel... One thing became clear to me - that objectiveness, the depiction of objects, needed no place in my paintings, and was indeed harmful to them" (Read, 1959).

7.5 Conclusions

7.5.1 Sketching as a Cognitive Monitor

There is a growing consensus that conscious attention is necessary for learning new tasks, for coping with novelty and for breaking the mould of habitual thought patterns. Drawing on the work of Norman & Shallice, Baddeley, Johnson-Laird, Baars and Hofstadter (amongst others) I suggest that the brain implements an executive "Cognitive Monitor" which records and controls (by interruption and re-direction) a pathway through a stored hierarchy of procedural "action" routines used for visual thought (7.2.9). With constant repetition this pathway becomes optimal and is hence repeated. As it is thus learned it becomes "automatic" -
meaning that once started it runs without conscious control and no longer uses the
cognitive monitor. The advantages of this are that it runs faster and can operate in
parallel with other routines, releasing the cognitive monitor for other less familiar
problems which require new, unlearned processing pathways. If Hofstadter is
correct, (I argue he is partly correct), creativity must put a heavy load on the
cognitive monitor to break out of "sphexish" mental ruts. In the visual arts
creativity is, partly at least, a form of problem solving in which well practised
ways of perceiving and imagining visual relationships no longer work. The mind is
searching for a solution to a visual problem in which either the solution or the
problem or both have not been encountered before. Since sketches are, by
definition, used to support visual invention, one way by which we would expect
them to do this, is by extending the limited capacity of the Cognitive Monitor (or
"Central Executive" or "Supervisory Attentional System" or "Global Work
Space").

This hypothesis is complicated by persuasive, if circumstantial, evidence
that unconscious mental processing is also an important component of creative
discovery. There is no reason to believe that the visual arts are an exception to this
rule as Kandinsky has testified (7.4.3). Having observed the growth of creativity in
many art students over thirty years, I agree with Arnheim (7.4.3) that the evidence
suggests that when an unconsciously generated insight into a difficult visual
problem occurs, (Boden's P-Creativity), it normally does so only after a prolonged
conscious effort to solve the problem has previously occurred. The problem then is
to understand how sketches and sketch attributes can simultaneously support the
conscious "anti-sphexing" of inventive thought while at the same time eliciting
(like Cozens' blots and Kandinsky's mis-oriented painting) unconscious (and so
presumably automatic) processes to achieve the same end. My solution to this
apparent paradox is to suggest that there are two different kinds of automaticity
involved which operate at different levels of the problem solving hierarchy. At a
high level, the level at which problem solving strategies are chosen, and the
(possibly lower) level at which the decisions taken define the degree of
inventiveness, acquired automaticity due to practice may be a handicap and need
the services of the Cognitive Monitor for metacognition. However at another level
we can infer the existence of acquired sub-ordinate slave processes which can
access stored knowledge automatically and which included essential stylistic skills,
which must operate unconsciously in order to free the metacognitive resources of
the monitor for its high level task. It is argued that un-monitored memory retrieval
and unconsciously generated imagery may operate more freely than when the
cognitive monitor is operating, producing a richer source of visual memory
permutations. Thus we may explain the occurrence of sudden insight, when not consciously thinking about the problem concerned, by supposing that the slave retrieval and imagery systems continue to operate after the conscious monitor is switched off (e.g. as in sleep or mounting a bus) or is directed elsewhere.

How do sketch attributes support such a balancing act between conscious and unconscious thought? Consider the demands that novel visual tasks place on conscious monitoring and metacognition. It is likely that the cognitive monitor (or central executive) evolved primarily to facilitate the learning of appropriate responses and action schemas when novel percepts and events lead to the breakdown of existing automated schemas. According to this hypothesis, the monitor was designed by natural selection to interrupt and "debug" (Mandler, G. 1985) unconscious behavioural patterns which are insufficiently flexible to cope with a novel situation (Baars 1988). Whilst the storage and control capacity of the conscious monitoring system is sufficient for this purpose, it is not sufficient for the culturally ancient, but biologically recent, tasks of planning and imagining possible future objects and events. With experience, artists and designers acquire a battery of well practised thought processes which operate most efficiently below the level of conscious awareness. With novel visual problems, such automaticity can be a handicap. For example Langer and Imber (1979) have demonstrated that error detection becomes quite poor when skills become automatic. As Baars (1988) points out (in tacit agreement with Hofstadter) new problems require new processes or new pathways through learned processes which must become "de-automatized" and conscious for this to be possible (7.2.9).

One way by which the capacity of such conscious monitoring may be enhanced, is by the forced slowing down of an otherwise automatic and unconscious stream of imagery and descriptive propositions, by the sequential process of overt drawing. Images and propositions become more available for conscious inspection as a result of trial and error hybrid-image formation described in 5.3.7. Those ambiguous perceptual parts of a sketch which trigger automatic memory retrieval mechanisms (4.1), can also support the cognitive records (7.2.9) so necessary for back-tacking and error correction. Not every one may agree with Rawson (1969) that the temporal sequence of drawing is visible in finished sketches. However there can surely be no doubt that in the act of drawing a sequence of marks is temporarily available as a sequence of reminders to the artist, of very recent images and statements by the "inner voice". In order to be able to usefully "change one's mind" a sequential record of immediately past thoughts must be available. just as E. M. Forster is supposed to have remarked "How can I know what I think until I see what I write?" so an artist might say "how can I know
what I imagine until I see what I sketch?". Tulving's theory (1983) of episodic memory retrieval by "ecphory" may be relevant here. Each visible sketching act in context is perhaps a cue to a "micro-episode" in a visual thought pattern which retrieves and makes conscious a corresponding mental episode by combining with its buried memory trace (4.3.9). There are a number of ways in which the marks made in rough paper sketches are tagged with such "micro-episodic" markers available for metacognition. For example,

- A contour line may be drawn very faintly at first, then repeated with increasing emphasis and contrast, as its indeterminacy is reduced. A set of such lines is an example of a "forwards" uncertainty reduction record.

- Erasures are often incomplete (e.g. with an India-rubber), leaving faint traces of earlier lines. Unlike the crossing out, imperfect erasures invite tolerance and multiplicity in image generation. A set of such erasures might be termed a "backwards" uncertainty reduction record. (Illustration XXXIV)

- An error or change of mind may be indicated not by erasure, but by a conventionally understood "crossing out" with zigzag lines (Illustration XIIb). Unlike erasure such a transparent crossing out records both an initial thought and an unequivocal change of mind.

- Written notes and labels on sketches can sometimes have metacognitive functions corresponding to any of the recursively embedded levels of self-awareness discussed in 7.2.2. For example the designer Charles James, C. wrote across one of his sketches "Forget all you know and learn something every day" (Coleman 1983).

Despite these attributes, it has to be admitted that a single sketch on paper records only poorly and for a short term the sequence of mental acts that the cognitive monitor needs. However sketches also typically occur in multiple sets on the pages of sketch books. Such a set is structured like a branching tree with vague "root" sketches leading to "trunk" and "branches" of the tree preserving different alternatives, many of which are later abandoned. In a conversation with Christian Zervos, Picasso revealed his interest in visual metacognition. "It would be very interesting to record photographically, not only the stages of a painting but its metamorphoses. One would see perhaps by what course a mind finds its way towards the crystallization of its dream" (Ghiselin, 1952). In 1954 this suggestion was taken further, when Clouzot made his remarkable filmed record of Picasso's sequential thought processes, using the techniques of painting on glass and time lapse photography. The opportunity to use information technology to improve upon the usually rather haphazard episodic record of the sketch book is immense and will be discussed later. Here we are only concerned with understanding the need
for such a record and the way in which sketching might amplify the mind's own record, which must be kept in order to reconstruct over-learned mental processes and to find more effective pathways through the tangled hierarchy of possible images and concepts.

If sketches assist the conscious monitoring and decomposition of thought patterns that have become automatic, then what are we to make of the extensive claims that unattended mental processes are important to creativity? (7.4.3). An obvious conclusion would be that sketching only supports Poincaré's Phase 1 of creativity (conscious problem analysis) and Phase 4 (conscious development of the implications of an insight). Phases 2 and 3, which apparently occur when the mind is not consciously occupied with the problem, appear to use automatic processes to achieve a highly non-automatic result. I have here proposed two candidate processes which may produce fruitful insights during unattended processing, namely automatic memory search and automatic image formation. In ways that we do not yet understand, conscious control may limit rather than extend the novelty of the results of such processes. A role for untidy, accidental attributes of sketches in assisting such delayed unattended insights, is suggested by Amy Lowell's metaphor of posting a subject into the subconscious "much as one drops a letter into the mailbox". Indeterminate or accidental sketch stimuli can arise whilst helplessly tackling an unresolved visual problem. Such stimuli may act as "letters" to poorly understood unconscious retrieval and imagery processes which work more efficiently when they are not interrupted or controlled by conscious monitoring. However Amy Lowell's (or Picasso's) knowledge that they are "posting such letters" is pure metacognition.

The sketching functions outlined in chapters 3-6 all suppose that sketch attributes amplify the capacity to perform specified, well practised visuo-mental tasks using successive refinement, within the limits of descriptive constraints. These tasks are all aspects of the process of thinking about an unresolved visual problem, of imagining a non-existent object. In this chapter, I have tried to suggest mechanisms by which sketching is also used to amplify the capacity to adapt and alter at a higher level, the way such strategies of thought are used. In an open-ended "dialogue with the self" (Goldschmidt, 1991 Scrivener & Clark, S. 1992) sketches amplify the capacity of artists to think about thought itself, and to invent new ways of using the conscious and unconscious processing patterns that sketches also support.

These tentative conclusions are summarised as the fifth hypothesis about sketch function.
7.5.2 Hypothesis Five.

Sequences of sketching acts support the conscious awareness of one’s own cognition. This assists creativity, providing voluntary control over highly practised mental processes which can otherwise become stereotyped. Unforeseen percepts from untidy or accidental stimuli can elicit unconscious processes which break the mould of habitual thought, whilst a temporal record of recent ideas makes it easier to change one’s mind at appropriate stages.
8. TOWARDS A UNIFIED THEORY.
"He who has a hundred miles to walk should reckon ninety as half the journey"
Japanese Proverb.

8.1 Do Five Hypotheses make a Theory?

This study was conceived as an attempt to improve the design of computer sketching systems for artists. However the task immediately raised questions inextricably linked to the nature of visual cognition. Scholars such as Ernst Gombrich, Rudolf Arnheim and Philip Rawson have shown that sketches, although apparently impoverished representations, are (in the hands of the skilled at least) cultural inventions of great subtlety. Since the time of Leonardo da Vinci, artists have realised that such sparse or untidy representations can greatly amplify the power to imagine what does not yet exist. However it is only in the last few years that psychology and neuroscience have begun to provide the evidence on which a theory as to why Leonardo's intuition that "confused things rouse the mind to new inventions" is true. We now have evidence that when the brain is used to imagine new visual objects, it can only simulate visual experience by "borrowing" the cerebral machinery for perceiving and completing unexpected objects in the real world. The descending connections in the visual cortex enable the brain to inspect and manipulate into new structures a stored museum of visual images that would otherwise be concealed from consciousness. It is for this reason that inventive imagination can be amplified by generating artificially ambiguous or unexpected stimuli which act as substitutes for the type of real world surprises the visual system has evolved to handle.

Until the last two decades, the technology necessary for a realistic but cheap and easily manipulated imagery medium did not exist. Therefore it has never been clear historically that a more detailed, accurate and complex representation of imagined appearance would not have served the visualising function better than the paper sketch. As indicated in Chapter 1 the possible advantages of traditional sketch attributes are hard to disentangle from disadvantages imposed by the physical constraints of the medium. By removing many of these constraints, new image handling technology forces system designers to re-examine the practice of sketching in order to decide what representational attributes the ideal visualising medium should have.

In Chapter 2.3 it was suggested that the invention of photography, far from making sketch indeterminacies redundant, actually helped to improve their status. Attributes used hitherto only privately to amplify the imagination of an artist became, in Fine Art at least, public property to enhance the "beholder's share". In
All the five hypothetical functions of the sketch are ways of amplifying visual thought and invention. However they are interdependent in other ways that are not yet clear.
one sense figurative art has returned to its roots, since the suggestive sparseness of modern sketches seems also to have been a property of the earliest known figurative images, 35,000 years ago (White, 1989).

So far, (as shown in Diagram I) the cognitive functions of the sketch have been described, separately in Chapters 3 - 7, as five hypotheses about the way that sketch attributes can amplify visual thought. The secondary functions, mentioned in Chapter 1, of recording and communicating visual ideas, are outside the scope of this study. (The possibility that future machine sketching systems might advantageously provide separate, but linked, support for these secondary functions is a subject for future research.) However this separate treatment of the five functions is incomplete for two reasons.

Firstly, cognitive functions have the alarming tendency to multiply when closely attended to. Future workers may well decide that the classification of Diagram I is too simple because it confuses distinct functions that deserve to be treated separately. Thus further consideration of Function 1 (mental translation of representational type) may easily show that it is really two functions. Depictive to descriptive and descriptive to depictive "translations" are perhaps processes so different from each other that they deserve to be treated independently and not as two sides of the same coin. In Chapter 3 I may have given the unintended impression that these two processes were the "forward" and "reverse" directions of a single symmetrical process. This cannot be correct. For example to translate a descriptive proposition into a depictive image, added information, retrieved from memory, is usually necessary. On the other hand descriptive propositions can be extracted from a spatial depiction by mental scanning and inspection or by other "visual routines" (Ullman, 1984). Both processes imply one to many mappings, but because the first is synthetic and the other analytic the translation process cannot be symmetrical. Function 4, "Selection", now seems so obvious that perhaps it ought logically to have been the first function to be discussed. Selectivity was recognised very early on as characteristic attribute of sketches but was originally perceived only as a slave function of imagery support. As Chapter 6 shows, the selective attention function has many sub-components. Moreover, because I was discussing sketch attributes rather than sketch media, I did not do justice to the idea that well designed sketching systems may play a role in supporting divided attention. To what extent, for example, do sketching systems facilitate or hinder the dual task of attending to an imagined image and attending to a machine sketching interface (or, for that matter, a palette and paints)? Is this another case of a function multiplying? Finally the monitoring function, the "discovery" of which has its origins in the claim by Rawson (1969) that drawings
made by artists record a temporal sequence of drawing acts, leads (in Chapter 7) to the perplexing subject of consciousness, and is about to give birth to a new function - Learning. Unfortunately the acid test for this - "Do sketches support learning in ways other than by cognitive monitoring and metacognition?" cannot be applied without further research. So for pragmatic reasons learning, probably an important function of sketching, is treated here as a slave consequence of the cognitive monitoring function. The sub-structure of the five main functions is discussed in Chapters 3-7. However this sub-structure is fluid and must be regarded as provisional. The multiplicity of hypotheses is an embarrassment, making it difficult to merge them into a simple simplifying theory. It is even possible that the search for such a unified theory will prove to be misguided. Minksy (1985) has claimed that attempts by psychologists and philosophers to find a few simple principles to unite the many functions of the mind are due to "physics envy" and is bound to fail because the brain consists of a collection of loosely collected processors which have evolved to do many different things which have little in common. This is a rather extremist viewpoint and it is to be hoped that a useful theory of sketch function will eventually unite these disparate hypotheses by showing that they are consequences of some unifying and related principles about how visual media support the mind.

The second reason why Diagram I is unsatisfying is that it suggests that the cognitive functions of the sketch are discrete and can be considered independently. Sketch functions are interdependent in complex ways. In this study I confess that these interdependencies are only dimly understood "as through a glass darkly". In this chapter some rival candidates for a "unifying principle" of sketch function are examined.

8.2 Mental Type translation as a Connecting Link.

One possible connecting link between sketch function is the representational translation process described in Chapter 3 as Hypothesis 1. The two complementary types of information processing, "descriptive to depictive" and "depictive to descriptive" translation are to varying extents constituent functions of the other four functions. It was convenient to discuss type translation in an early separate Chapter because I wanted to develop the existing rather crude binary distinctions between description and depiction into what, I believe, is a more realistic continuum, with several quantitatively varying dimensions and many intermediate types. It was also important to introduce early the concept of the descriptively labelled, spatially depictive map as an intermediate data type fundamental to sketch function. Only gradually did it dawn on me that this data structure has beautiful and elegant processing advantages that the primate brain
must have discovered by natural selection long before humans invented its cultural equivalent. These seem to have been largely ignored in theories of perception and memory. The suggestion that such an intermediate data type might "catalyse" depictive to descriptive translation lowering some processing "effort" barrier in the way that enzymes form intermediate compounds with their substrates to lower the "activation energy" necessary to initiate a chemical reaction is intended only as an analogy. By this analogy, cognitive work load or processing "effort" is analogous to chemical activation energy whilst the instability or fluidity of a visual idea is analogous to the instability or energy level of a chemical compound. Having made this analogy it might be fruitful to speculate further and to suggest that Function 2 (perceptual retrieval) and Function 3 (supporting imagery) are subservient functions of the translation function. According to this conception, the purpose of retrieving descriptively encoded information about objects and their parts is to generate an improved depictive image of the sketch subject matter. The purpose of generating mental images in spatial register with visual components of the sketch, is to extract and make explicit new descriptive information, which is only implicit in the spatial image. Depictive to descriptive translation occurs in different forms and contexts in the sketch. According to Function 2 it may occur automatically at two levels. The early stages of visual processing in which descriptive structure is added "bottom-up" to the image are depictive to descriptive translations. That is, implicit information is made explicit and encoded. Whether the later stages such as "recognition by parts" is aptly named a theory of depictive to descriptive translation, depends upon whether information enrichment by retrieval "counts" as translation. I think it can. After all when we use a dictionary to "translate" a single word we are also accessing a paragraph of alternative meanings. "Translation" is not used here in the same sense as "deciphering". Mental image generation as described in the Kosslyn theory is descriptive to depictive translation occurring in two stages (both one to many mappings) -

1. A stored categorical description of spatial structure is translated to a coordinate representation
2. The coordinate representation is mapped at varying resolutions onto an internal (cerebral) spatial array.

According to Chapter 5, both these processes are supported by the formation of a hybrid image in which a mental image is superimposed upon the physical marks of the sketching surface. Written notes, especially when tagged to spatial locations, assist translation processes by acting as memory look up keys for image generation (Paivio, 1971a).
DIAGRAM IV

Is translation of representational type a unifying link between the cognitive functions of the sketch with functions 2, 3, 4 and 5 subservient to function 1?
A primary purpose of the sketch (implemented by hybrid image formation) is to discover new depictive and descriptive relationships by overt and covert movements of attention (using eye-movements and covert scanning). Thus the discussion of selective attention in Chapter 6 is also concerned with translations of representational type. Evidence was discussed that serial movements of an attentional "window" or "zoom lens" are used for conscious information extraction such as counting, curve tracing and visuospatial comparisons. This type of controlled inspection and comparison is probably the most important way in which sketches support depictive to descriptive information extraction. I have discussed Kosslyn's evidence that image generation occurs in automatic sequential stages and his suggestion that in one form it can occur by covert scanning using a small attentional window like an "inner pencil" (Kosslyn et al. 1988, Kosslyn & Shin 1994). This is the complementary translation process in action, descriptive to depictive translations being tried out on an internal "sketch-pad". Thus it is reasonable to assume that many of the functions of which support sketch attribute selection also support representational type translation.

Finally, type translation is implicit in many of the operations discussed in Chapter 7. Consider how necessary the proposed Cognitive Monitor would be for depictive to descriptive translation by image scanning. The monitor would record temporarily, the scanpath, (scan directions, distances, field "apertures", sub-goals and attention results) and store the incremental results of sequential processing (counting for example). (Experimental evidence that unconscious perceptual scanpaths are indeed stored is provided by Noton and Stark, (1971). Not only would this record be intrinsic to some types of information extraction, it would also be necessary for error correction and mental experiment. For example if an attempted scan path fails to meet an expected goal in one direction the "cognitive record" can be used to back-track and to correct the direction or to reset the goal. The distinction between voluntary and automatic processes is also central to an understanding of descriptive to depictive translation. The widely supported belief that creative invention has alternating voluntary (conscious) and automatic (unconscious) stages (7,4,3) may be seen as a collaboration between monitored and un-monitored processes for the purposes of information translation. Thus when monitored attention to a depictive element fails to meet a goal, the cognitive monitor initiates retrieval and comparison processes, which run automatically and which, as Baars (1988) suggests, may be inhibited by consciously attended sub-goals.

It is therefore tempting to suggest a "unified" theory in which functions 2,3,4 and 5 are subservient to function 1, mental type translation as in Diagram
Are all the hypothetical functions of the sketch ways of supporting indeterminacy or uncertainty in the early stages of visual invention?
IV. However this conclusion is premature. Although type translation may run like a horizontal thread throughout the other "vertical" functions it is surely also true that sketches mediate many forms of mental information processing which are purely depictive - manipulating spatial relationships for example, or purely descriptive such as using imagined language to manipulate categories and propositions.

In thus increasing the emphasis upon representational translation, it must be remembered that two types of descriptive to depictive and depictive to descriptive translation are involved. One is the purely visual translation process discussed above, involving the hypothetical categorical and perhaps coordinate visual information implicit in theories of recognition (Chapter 4 section 4.3). The other is the translation from verbal description to visual imagery (Paivio 1971a). This translation is implicit in many of the written notes found on sketches (Illustration XXIII). As Denis and Cocude (1989) have shown, people can generate and successfully scan spatial images from purely verbal descriptions. Following Kosslyn (1994) it is assumed here that this process uses "association memory" and is mediated by the same stored descriptive and semi-descriptive representations that are used for recognition. Such a process would include the translation from "categorical" (left cortex) to "coordinate" (right cortex) representations of spatial structure (Kosslyn el Al., 1989, Kosslyn et Al. 1992).

It is clear also from the written notes on sketches and the evidence for verbal thought whilst sketching (Goldschmidt 1991, Scrivener & Clark, S. 1992), that verbal working memory is also used in sketching. In the next section (8.3) I therefore include the "articulatory loop" and the "phonemic store" from the Baddeley and Hitch model of working memory (Baddeley, 1986) as resources supported by the sketch.

8.3 Is Visual Indeterminacy a Common Principle?

The role played by visual indeterminacy or vagueness may signpost another direction in which a common principle for sketch functions can be sought (see Diagram V).

Each of the functions imply a degree of indeterminacy, as noted below.

Function 1

Indeterminacies support mental translation. Pentimenti, for example, are one of the means by which the mind’s internal descriptive representation of shape can be mapped to its multiple depictive alternatives.

Function 2

Skillfully used, visual indeterminacies can promote perceptual retrieval. The subtle indeterminacy of incomplete contour fragments provides, I have suggested,
improved conscious access to unconsciously stored structural descriptions. Complete contours being too narrowly specific would actually limit rather than assist this retrieval process (Illustration X). This is necessary to interrupt inappropriate automatic cognition and to trigger and support the "surprise" monitoring machinery of the brain's control system.

**Function 3**

The indeterminacy of incomplete drawings or colour studies is necessary to leave room on the perceptual part of a sketch for its superimposed imagery. Thus empty spaces or randomly textured areas need not be accidental but fulfil valuable cognitive functions.

**Function 4**

Selectively representing some visual attributes only, whilst leaving others indeterminate, supports the capacity for focused thought. Such segregated visual attributes, (contours without surface detail, colour without drawing etc), can simultaneously leave room for imagery (Function 3). Such indeterminacy is implicit in the need to leave some decisions for a later stage of design, preserving multiple options for the unrepresented attributes.

**Function 5**

In Chapter 7 it is suggested that sketch actions support the sequential recording and subsequent correction of cognitive acts. Here sketch indeterminacies can act as branch points for experimental back-tracking and error correction. They assist the brain to monitor its own progress and to find novel pathways through a complex tree of possible decisions. For example a contour category is often represented in a sketch by a number of different but roughly parallel lines varying in emphasis (contrast). This provides a visual "audit trail" which can be used by the monitor to return to a decision point and select a different alternative and to then consider the consequences for a new imagined structure (Illustration XXXIV).

For all five functions, the most useful indeterminacies will be those which provide an element of appropriate surprise. An understanding of what exactly "appropriate" means here is one of the tacit skills of visual thought. However indeterminacy is only one of the functional attributes of the sketch and is perhaps only incidentally a common factor. The view supported here is that indeterminacy is simply one of the consequences of more fundamental underlying principles such as the one discussed next - support for visual consciousness.
The functions of the sketch as support for visual consciousness. Is this a unifying principle?
8. Towards a Unified Theory.

8.4 Sketch Function and Visual Consciousness.

8.4.1 Is conscious control a unifying principle?

Another possibility, and it is only speculative, is that a unified theory of sketch function might be based on the role of the sketch in metacognition and conscious control as discussed in Chapter 7. Thus the primary function of the sketch is to amplify the limited capacity of the brain to make tacit visual and semantic knowledge available for conscious manipulation into novel structures. In other words are all the five functions of the sketch forms of support for visual consciousness? (Diagram VI)

Thus it is possible that Function 1 (Chapter 3) is not only to "catalyse" mental translation but also to make over-learned and automatic types of descriptive/depictive translation conscious and available for voluntary control. Although the processes of memory retrieval itself do not become conscious, Function 2 might be related to the generation of visual surprise which increases the awareness of the results of perceptual retrieval (the difference between recall and recognition). Similarly, the hybrid images of Function 3 may improve the conscious awareness and control of mental imagery. Function 4 is support for conscious selective attention - Posner's anterior attention network (Posner & Rothbart 1992). Function 5, as discussed in Chapter 7, is support for the awareness and mental monitoring of the sequence of cognitive acts during visual thought.

There are at least two sorts of tacit knowledge possessed unconsciously by artists and designers. The first is stored knowledge about the nature of art or design. Some of this is general or cultural and shared with others, but much is private and personal and cannot be easily shared. Such knowledge is acquired by experience, is different for each individual and certainly includes an enormously large but not easily tapped, reservoir of remembered images. It has been shown experimentally, for example, that if subjects are shown 2,500 colour slides, for only ten seconds each, they can still recognise 90% of them several days later! (Standing et Al. 1970). This knowledge is tacit not only because it is not easy to retrieve, but also because much of the information it contains is only encoded in memory implicitly. However the task of consciously recalling on demand, specified examples from this vast reservoir of visual information, is much harder. This depends, in part, upon the difference between recall and recognition (see Baddeley 1990). It requires cognitive effort and special skills to make such tacit visual knowledge explicitly available. If such knowledge can be represented explicitly and made available to conscious covert attention (the "mind's eye"), it can be used more freely for inventive manipulation into new structures, in new contexts. A second type of tacit knowledge, perhaps even more important to visual
invention, is a huge "cerebral data-base" of automatic mental processes (Shallice's "action systems"). Although some of these are innate, it is the many visual skills acquired by practice and experiment which characterise an individual artist's style. Throughout their careers, artists and designers acquire new cognitive processes related to perceiving, recognising, imagining, drawing, composing and inventing. As these processes or sub-processes are successfully repeated they become increasingly automatic and unconscious. There is strong evidence for the belief that visual thought and action are dependant upon the use of such acquired routines, which the brain implements efficiently and in parallel without awareness (7.3.2). The goal contexts which trigger such automatic processes can themselves become automatic to various degrees (Baars, 1988). However as soon as the normal conditions for triggering such unconscious mental habits include unexpected, novel or ambiguous stimuli or when the goal image is itself novel, then automaticity can become a handicap and even, in some situations, become a danger (Reason 1979).

As Langer and Imber (1979) have shown with letter sorting tasks, when a mental process becomes unconscious as the result of practice, subjects can no longer report the number of steps in the task and are unable to monitor their own effectiveness. In order to cope with unexpected stimuli, I argue that humans and other mammals have evolved the Cognitive Monitor which records and assesses the intermediate results from a complex hierarchy of cognitive processes, and can redirect them before they progress too far in an undesired direction. Thus, when necessary (e.g. when there is a percept - goal-image mismatch), the monitor can interrupt a faulty process, back-track and change parameters or re-compose the processing sequence. The evidence suggests that it is novel, experimental thought processes which most need such selective support, because of their reliance on limited capacity resources. It is conjectured that all the five functions discussed in Chapters 3-7 are different aspects of the same global function, namely external support for a sophisticated visual monitoring system. Thus we can suppose that percept-memory hybrid images are mapped into a limited capacity spatial array because this supports novel types of depictive to descriptive translation by recording and monitoring the movements of a spatially selective scanner (Posner's anterior attention system). The perceptual retrieval system maps object shape and colour into a limited capacity visual memory, (distinct from the spatial map (Logie 1995)), because there is a demand to make novel, un-expected but monitored, links between spatial structure and object detail. A new reason for the limited capacity of conscious working memory suggests itself. The capacity of spatial and object memory cannot exceed the instantaneous "frame" capacity of the cognitive recorder which must record sequences of visual images and their goals. It appears that not
all automatic processes can be closely monitored. Therefore there is a need for a mechanism which can suppress unwanted slave processes and initiate others. It makes sense to suppose that the conscious monitoring system evolved primarily to provide a mechanism for learning new procedures for handling unexpected or incomprehensible percepts and events in the immediately present environment. Conscious processing control of visual memory has evolved to be activated and supported by a continual stream of unexpected and un-comprehended (hence ambiguous) perceptual stimuli from the natural world. As the cultural need to imagine non-existent objects and future events increases, the danger of automated internalised action "thought slips" occurring also increases. Internal percept-like simulations of imagined objects cannot produce the kind of ambiguity and surprise necessary to elicit closely monitored processing (5.2.5). Thus like the proposed mechanisms for hybrid-imagery (5.4), the Cognitive Monitor (or "Central Executive"), is imperfectly adapted by evolution for monitoring and reconstructing over-habituated processes applied to imaginary objects. Thus a cultural invention is needed to provide artificially, the ambiguous and surprising visual stimuli which trigger a consciously monitored processing mode (Baars 1988, page 145-8 discusses an analogous theory in relation to speech processing). Note that two apparently contradictory functions of the Leonardo type indeterminacies are now beginning to converge. If Hypothesis Two (outlined in Chapter 4.1) is correct, then an indeterminate or incomplete sketch contour would be expected to elicit automatic perceptual retrieval mechanisms which access a greater variety of visual memories than would an unequivocal one. As was argued in section 4.3.10 "The evidence is consistent with the view that confused and untidy sketch attributes stimulate creativity because they elicit the automatic recognition matching processes in a set of nested search loops which activates many more components of long term visual memory than would either a complete and unambiguous image or a very sparse one." Moreover it was further argued in Chapter 6 that the absence of some types of irrelevant visual attributes was necessary to inhibit otherwise unstoppable processing distractions. It is proposed (as a consequence of hypothesis 5, section 7.1) that the Leonardo indeterminacies also activate consciously controlled "anti-sphexish" thought processes. Automatic ("intuitive") retrieval and conscious processing are not contradictory functions. As suggested in 7.4.3, automatic memory retrieval is a necessary slave process of consciously controlled visual thought. When visual surprise occurs in the natural environment, the twin processes of automatic retrieval and of voluntary image - percept matching can collaborate. Moreover the two processes can run in parallel which would not be possible if both were consciously controlled. When, over a longer time period,
A model of sketch support for the cognitive control of working memory resources. The "inner artist" is intended only as a metaphor.
accidental sketch "surprises" are processed by the cognitive monitor, they can be simultaneously "posted" as uninhibited retrieval cues to unconscious memory (to repeat Amy Lowell's metaphor).

8.4.2 Resources of Working Memory for Visual Thought

It is worth considering how consistent the above proposal is with the theories of resources available for visual thought as proposed, for example, in the Baddeley and Hitch model of working memory (Baddeley 1986). Diagram VII attempts to outline how the components of this model can be related both to the hybrid image model of sketch function (Diagram III) and to the Cognitive Monitoring function discussed in Chapter 7.

As described in Chapter 5 the Baddeley and Hitch model proposes a "Central Executive" resource which can control and monitor a number of separate, "servant", short term stores and associated processes. Baddeley (1986) provides evidence linking the Central Executive to focal attention and to volition. The Baddeley Central Executive also bears several similarities to the Supervisory Attentional System (SAS) of the Norman and Shallice (1980) model of volition. In Chapter 7 (7.2.9) I proposed a additional function, that of recording, for the purposes of control and correction, a short-term record of cognitive actions. I call this revised and extended, combined Central Executive and SAS, the "Cognitive Monitor".

Baddeley-Hitch model posits two visuospatial working memory components, the Visuospatial Sketch Pad (a spatial memory system which supports movements of attention and can be interfered with by visual tracking tasks) and a short-term Visual Store store used for object shape and colour processing (Logie & Marchetti 1991, Logie 1995). These two parts of the Baddeley model map quite well onto the depictive and descriptive components of the "hybrid image" model of sketch function discussed in Chapter 5 (5.3.7) and illustrated in Diagram III. In Diagram VII these components are simplified into the ellipses labelled "Hybrid Spatial Image" and "Hybrid Visual Description", although I have had to compress the triple buffer structure of these into single components. Logie (1995) might not agree that the more passive visual store was "descriptive" but that difference would only be one of definition. In fact he is rather vague about the representational format of these two components. I assume that my "descriptive" store must include coordinate based definitions of object structure, but that if, as Logie (1995) maintains, representations cannot be scanned or spatially transformed in this resource, then they cannot be "depictive" as defined
in Chapter 3. Logie associates movement with the spatial working memory component in two ways -

1. It can be used for the mental representation of imagined movement and movement planning.

2. It supports movements of attention allowing high resolution processing of small sequentially represented parts of a neurally implemented spatial medium.

Logie (1995) also discusses evidence that the Central Executive is necessary for image generation, refreshment and transformation. However it is surely (on the basis of the evidence he quotes) still an open question whether image generation and transformation processes should be considered as part of the Central Executive or merely as additional slave processes which can be controlled and monitored by the Central Executive. In this model (Diagram VII) the latter hypothesis is adopted.

As proposed in the Kosslyn model of mental imagery (Kosslyn, 1980, 1984) (and my hybrid image model of Diagram III) the spatial components of an image are mapped onto a medium with an array format which supports movements of "focal attention", (Kosslyn's (1994) "window of attention"). This is assumed to be physiologically implemented in those parts of topographically mapped visual cortex which are now known to be shared between perception and imagery (Kosslyn et al. 1993). Besides providing a mechanism for binding spatial locations to object description (Treisman & Gelade, 1980), movements of attention have two other functions which, in the model, are controlled and monitored by the Cognitive Monitor:

2. Movements of attention are one of the mechanisms by which new descriptive information may be extracted from a depictive image. A well investigated example is that of curve tracing (Jolicoeur et al. 1986).

3. As Kosslyn (1994) has suggested, covert movements of focal attention can also be used to mentally image a new shape or structure without recourse to memory, as if drawing with an imaginary pencil.

Also included in Diagram VII are the two Verbal components of the Baddeley and Hitch model, a passive "Phonemic Store" or "Phonological Cache" (Logie 1995) and the "Articulatory Loop" which refreshes by rehearsal the contents of the Phonemic store with a type of sub-vocal speech. Because there is evidence that these two stores are used in written as well a spoken language, (e.g. Baddeley & Lewis 1981), they are relevant to the function of written notes found on sketches and to the implicit use of verbal thought whilst sketching. Just as Freud proposed that mental images are a mechanism for making unconscious visual
memory conscious, so it is also likely that "inner speech" is a mechanism for consciously manipulating descriptive categories and propositions.

During visual thought, as Diagram VII outlines, the Cognitive Monitor is assumed to be able to activate and to monitor the processes of memory retrieval, image generation and manipulation and movements of focal attention on the Hybrid Spatial Image. Whilst doing this is keeps a continuous record of processing results in the form of depictive images linked to, or labelled by, descriptive information. Movements of attention linked to memory access can thus result in depictive to descriptive translation. Image generation linked to memory access and to Kosslyn's "categorical to coordinate" translation (Kosslyn et al. 1992) is one route by which the Cognitive Monitor may "supervise" descriptive to depictive translation.

During verbal thought, control of the Articulatory loop and the monitoring of briefly stored words in the Phonemic store are assumed to be supported by the written notes on sketches and perhaps also by the act of writing such notes. Depictive to verbal translation is assumed to be mediated by visual descriptive information which can be translated and made conscious by an "associative" memory system in Long term memory (see last two paragraphs of 8.3). Not shown in the diagram is the immense reservoir of stored Semantic Knowledge which such verbal-visual translation processes can tap. A telling point made by Jackendoff (1990) explains why some verbal thought will always accompany the sketch. The contents of verbal memory and the articulatory loop can easily be communicated and recorded by direct translation into overt speech and writing. Unfortunately says Jackendoff we have no "television monitor in the middle of our foreheads" for similarly translating our mental imagery into external physical imagery. Language is necessary, not only for communicating our thoughts to others, but also for clarifying our thoughts to ourselves (Jackendoff 1983).

8.4.3 Support for an "Inner Artist".

According to this tentative model, sketch attributes support the Cognitive Monitor both directly, by assisting the recording of cognitive actions or indirectly, by supporting the servant processes that the monitor must control. The evidence that the internal processes used for imagining and planning future objects and events depend upon those neural processes within the brain, which are biologically better adapted for perceiving and acting, is still growing. For example PET studies by Parsons et al. (1995) have recently shown that when subjects solve a visual discrimination problem of the type usually believed to require mental rotation (Shepard & Cooper, L. 1982), differentially increased blood flow occurs in exactly those regions of the motor cortex and cerebellum which are activated by both the
imagined or the actual movements of a hand. It appears that mental rotation uses a neural resource that is also used when one physically rotates or plans to rotate a real object. For several years, Baddeley and his colleagues (1986) have coined the terms "Inner Voice" and "Inner Ear" for the articulatory loop and the phonemic store, to capture such an apparent equivalence between covert and overt verbal behaviour. It is equally tempting to think of the covert movements of attention as movements of an "Inner Eye" and the neural medium which support such movements as an "Inner Sketch Book". Now the Parsons et Al. PET studies (1995) suggest that we may extend the metaphor to include an "Inner Hand" capable of generating and manipulating imagined objects. (It is perhaps taking the metaphor too far, to suggest that it also wields an "Inner Paintbrush" loaded with imagined colours and operates an "Inner Video Projector" running moving sequences from a visual library). At first glance, there appears to be a problem with applying such a metaphor to the imagined drawing capability of the "Inner Eye" (focal attention). Real drawing results from movements of the hand and pencil not the eye. However further thought suggests that even here the metaphor is not inappropriate. When we draw, the pencil point follows the focused movements of our eyes over the surface of the paper as they trace an imagined object. If our eyes emitted "drawing rays" they might leave a trace resembling a sketch, as the classic studies of eye movements by Yarbus (1967), have shown. It is interesting to speculate that physical movements of a pencil may support the invisible cognitive recording of imagined movements of attention for later thought.

The manipulation and perception of physical media and images is performed by the publicly observable "Outer Artist" with real eyes, real hands, ears, a voice and a real sketch book and pencil. But to understand the sketching acts of this "outer artist", one must infer the existence also of a metaphorical "Inner Artist" which controls and monitors those internal neural processes and representations which precede and accompany the inventive images and actions of the "Outer artist", during visual thought. The Cognitive Monitor, as described in Chapter 7, fills this role.

The use of the term "Inner Artist" as a functional metaphor for the Central Executive or Cognitive Monitor (when used in the visual arts) cannot of course be taken literally. It too obviously resembles an "homunculus" theory. There can be no recursive mind within the mind. The Cognitive Monitor is postulated to be a neurally implemented recording and control mechanism which can be activated automatically and to varying degrees by unexpected or confusing data from the sense organs or from acts of memory retrieval. Because the Cognitive Monitor is itself controlled automatically by the neural environment (data from stored
memories) and the physical environment (data from the sense organs) we are not conscious of the processes which activate it. To make it produce unexpected and inventive ideas, to make it seem like an "Inner Artist" external stimuli are necessary. This is why designers and artists sketch. The metaphor is used here to emphasise the fact that the brain is using, for imagining and planning the future, cerebral resources and processes which were designed by natural selection for perceiving and responding to the immediate present.

Genuine creativity combines appropriate forward planning with novelty or surprise. The cognitive mechanisms available to the brains of artists for mentally manipulating non-existing future objects in inventive, unexpected ways did not evolve for that purpose. They evolved rather for coping with surprise, indeterminacy and incomprehension which results when stimuli and events in the immediate environment cannot be matched with stored memories. In this situation highly practised, automatic cognitive processes fail and new processing structures must be learned by internally monitoring and controlling the parameters of more flexible but experimental action sequences. The processing systems available for such conscious monitoring have limited capacity. To ensure that the information on which such systems act is consistent, the brain provides a hybrid image constructed from repeated attempts to match current stimuli to stored experience. Thus the environment itself provides stimulus support for the mental manipulation of stored processes. This is a kind of biological improvisation, in which an existing cognitive mechanism becomes adapted to a new culturally derived function. Sketches are cultural inventions which "capture" the use of this percept-memory monitoring system from competing internal processes, providing artificial substitutes for the environmental surprises that normally elicit its operation.

No less an authority on vision than H.B. Barlow (1981) has remarked that the homunculus or "mind's eye" "has such a grip on our ability to conceptualise what is going on" when we mentally inspect an image "that I think it will be better to ask precisely what he is doing before declaring that he is redundant". The cognitive monitor which I claim is necessary for visual invention has many of the characteristics of a mind within the mind. This "Inner Artist" is apparently controlling an "inner voice" and an "inner sketch book". The latter resembles an erasable sketch-pad which the monitor can either scan by moving a variable sized attentional "Inner Eye" (focal attention), or "paint on", generating pictures by collaging together fragments of images cut from a vast but private visual library. Kosslyn (1994) has suggested that mental images can also be created by mentally moving the "window of attention" over the "visual buffer" (the inner sketch-book). In this respect at least the Inner Artist is unlike the real or "outer" artist who needs
devices such as pencils to record the movements made by their eyes. The inner voice can likewise be monitored by the "inner ear".

However despite all appearances, the theory is not a recursive "homunculus" theory because the proposed Cognitive Monitor is activated and controlled automatically, in response to stimulation from the internal and external environment. We are not aware of the details of the cognitive monitor's processes, only of its results in the cognitive record. Nor are we aware of the many processing systems that are competing for the monitor's services. Thus Posner's Posterior Attentional System (6.5) is an automatic system which can "capture", without voluntary control, a region of the visual field for the Cognitive Monitor. Despite the complexity of its operation the "inner artist" has at its core an automatic process. Only by modelling this in detail can the homunculus paradox be avoided.

Whether or not an attempt to find common principles underlying all the functions of the sketch will prove successful, conscious visual experience is clearly a primary concern of the visual arts. It is not surprising that visualising methods in the arts should exploit mechanisms used by the brain to provide conscious control of unconscious visual knowledge. It is also inevitable that such a theory of the sketch should be incomplete, because the nature of conscious awareness is so poorly understood. However it is also clear that there is a great deal of highly relevant evidence and theory about the brain, vision and memory, that has until now been little used in the design of new systems and media for artists.

8.5 Conclusions

The human brain provides an enormously adaptable but sometimes faulty capacity to use stored information to perceive, recognise and respond to visual objects, scenes and events in the real environment. It is at first rather puzzling that this impressive information processing power is, relatively, so poorly developed for using stored information to manipulate and act upon imaginary or future objects and events. A plausible explanation for this, suggested by converging evidence, is that the cultural need to visualise future objects is too recent to have allowed much biological evolution of the brain. The cultural need to imagine must therefore "borrow" neural resources that evolved for handling the endless surprises, ambiguity and complexity, with which stimuli from the immediate environment assaulted the senses our hominid ancestors. It is likely, for example, that neural resources, such as area V1, used to represent topographically, information from the eyes, have a short duration, "fast fade" performance, because they evolved to handle the rapidly changing, short duration images sent from the retina to the brain by the saccadic movements of the eyes. This resource is less well adapted to
support the vaguer and less easily refreshed images generated from memory (Kosslyn 1994). For this reason mental images need perceptual support to assist their maintenance and regeneration.

Like the textured and stained walls of neolithic caves, or the glowing embers of the fire into which Dickens’ hero David Copperfield would stare, rough sketches provide artificially, stimuli which can substitute for some of the variety and surprise of the real world. In so doing they assist the brain to adapt resources intended by natural selection to be used for perceiving and acting to the cultural purposes of imagining and planning.
9. SKETCHING AND COGNITION: SOME IMPLICATIONS.

"Variations on a theme are the crux of creativity" (Hofstadter 1985b)

"Perhaps we can now see what is required of the computer: it must be (lovely word) idiosyncratic, and either behave like a human brain, or allow more interaction and participation with the user - more democracy, more fuzziness" (Brian Reffin-Smith, 1984)

9.1 The Uses of a Theory of the Sketch

The criteria implicitly used by most psychologists for assessing the usefulness of a theory, have been summarised by Gordon (1989) as follows

"1. Theories should offer economical accounts of a range of facts. A theory is of not much use if a description of it is as long as that required to describe the relevant phenomena

2. Theories should attempt to explain phenomena, or at least suggest causal links between them

3. Theories should be testable. They should be stated in such a manner that deductions can be derived and tested empirically"

In Chapter 8 an attempt was made to apply the first two of these criteria to the five hypotheses about sketch function to see if indeed the hypotheses were sufficient grounds on which to base a "theory" of the sketch. The uncertain conclusions were that the outlines of a theory may be starting to emerge from a description of the ways in which the sketch supports working memory and the conscious control of thought patterns which need working memory.

It is the third criterion which concerns this chapter. How can the theory be tested?, or rather, following Popper (1959), has the theory been framed in such a way that it is easy to see what empirical tests could be done to invalidate it? In particular, since the stated aim is to advance the design of sketching systems for artists, how can the theory be used by system designers? The theory or at least its constituent hypotheses does make predictions which could be tested by controlled experiments in the laboratory by professional psychologists. It also contains hypotheses about cognition which make testable predictions. For example the triple buffer, hybrid-image theory of perception developed in Chapters 4 and 5 predicts the conditions under which the perception of impoverished stimuli will cause subjects to report hybrid visual experience. But the purpose of these theoretical suggestions was not to produce a new theory of cognition, but simply to make sense of an otherwise confusing experimental literature such as, for example, the apparent conflict between results which show percept-image interference and those which show reinforcement. The psychology of visual experience is as much in
need of better theories to explain the mass of data that already exist as it is of yet more experimental data. The literature on cognition is littered with descriptions of experimental results which, even after many years, have never received a clear explanation. For example, it would be difficult to think of a better test experiment for the triple buffer hybrid image theory of perception than the Bruner and Postman experiments performed forty years ago, on the perception of incongruously coloured playing cards (5.3.7.4). The results reported are exactly as predicted by the hybrid-image theory. It seems to me that the standard explanation of such "perceptual compromises" in terms of "unconscious inference" or "expectancy" completely fail to explain why a hybrid experience should be reported, or how it could be represented in the brain. Unfortunately I cannot prove that this and other experimental results that support the theory proposed here, were not used in framing it. In fact many were noticed only after the model was framed so I cannot see why they cannot be treated as "retrospective" experimental tests of the hybrid image theory. Other functions have also passed such "retrospective tests". For example only after I had framed the "inner artist" cognitive monitoring function of the sketch (Chapter 7) did I spot a paper by Annette Karmiloff-Smith (1990) on inventiveness in children's' drawings. She found that, below a certain age, children could only draw "funny pretend" objects by omitting parts of familiar objects they already knew how to draw. Older children however, were able to plan and execute drawings of imaginary people or houses with unusual shapes and with parts from other objects combined inventively. Karmiloff-Smith concluded that the difference was due to the fact that the younger children drew automatically without being able to monitor how they were drawing. The older children on the other hand have a "more explicit representation of their skill and so are able to control and alter their productions in ways which the younger children cannot". Applied to adults these experiments support the theory that skill monitoring is an important component of inventiveness. One purpose of this study is to examine how existing theories of cognition can be developed (for they have had to be developed) to make testable predictions, not only about cognition, but also about the effective design of new machine systems. I claim, for example, that although speculative, the model for hybrid-image formation (5.3.7) does make testable predictions about how visualising systems may encourage or discourage inventive imagery.

Thus the usefulness of the theory depends upon what predictions it makes about the relative effectiveness of different existing sketch attributes, systems and media and upon what predictions it makes about desirable improvements to such systems. Since there is no certain way of either measuring or defining inventiveness in the visual arts, perhaps the best way of experimentally testing
these predictions is to use them to design several different sketching systems and to simply ask artists to assess their relative usefulness. This clearly would be a very long second project. Meanwhile the theory of cognition relevant to the sketch will be advancing and changing. Here there is only scope to illustrate how the five related hypotheses which form the core of this study might be applied.

Four uses of the theory spring to mind:-

(1) To assess the effectiveness of traditional sketch attributes and media to support visual thought and invention. For example, do the hypotheses provide some of the criteria necessary to distinguish, from amongst those sketch attributes which result, sometimes accidentally, from the properties of a medium, the desirable from the undesirable attributes?

(2) To assess the cognitive effectiveness of recent machine systems to support the visualising needs of artists in the early stages of visual composition. Do existing systems, as is sometimes claimed, eliminate the need to sketch altogether, allowing a designer to pass straight from a concept to its detailed and realistic representation? In a least five respects, according to the theory, the following statement is ill-founded.

"Vivid computer graphics replace architects' pencil sketches and handcrafted models with realistic moving images" (Cover of the Scientific American February 1991)

(3) To provide criteria and a philosophy for designing, improving and assessing new sketching media. This is necessary to exploit the growing capacity of computer image processing technology in ways which enhance without interference the visual capacity of the human mind.

(4) To provide clearer and more explicit guide-lines to students of art or design about the cognitive purposes of sketches and how to exploit them. There is now convergent evidence that metacognition of one's own thought processes is an important component of learning. It is not the skill of drawing that matters in the sketch as much as the skills of visual thought which the sketch attributes support.

9.2 Assessing Traditional Media

So far, the properties possessed by sketches in traditional media have been treated in terms of the evidence they provide for the thought processes used by artists for visual invention. Only when there is a plausible theory about why such thought processes need sketch support, is it possible to assess how well traditional media provide such support. There is now a partial answer to the question raised in Chapter 1 about how one can distinguish functional from incidental attributes of sketches. The medium of the sketch, marks made on paper by a variety of pigments and hand held marking devices, is very old. It has hardly advanced in
five hundred years. Conceptually, attributes which are meaningless and purely a consequence of the medium - the sketch "hardware" in contemporary jargon - can be distinguished from sketch "software", those attributes which carry a culturally derived meaning, independent of the medium. As already discussed, the sketch medium, although it is certainly not "the message" (McCluhan, 1969), has an important role in generating subtle surprises and ambiguities, which may be promoted serendipitously to carry part of the message. Part of the skill of sketching lies in knowing intuitively when such medium based message carrying effects are desirable and when they are not. Thus the wobbles and accidental directions of a manually controlled pencil line, the smudges of partially erased charcoal, accidental blots of ink and the vagaries of watercolour as it runs or soaks into the many different types of paper available, can all be used to assist invention by one or more of the cognitive mechanisms I have discussed. However despite the richness and subtlety of cognitive function that this ancient medium performs, it is surely far from ideal. In fact when compared to the technologies used routinely for storing, manipulating and disseminating linguistic and symbolic information or for storing, manipulating and viewing very depictive images, the paper sketch appears very primitive. An alien visitor from outer space would surely think the continued use of paper and pencil for visualising future objects an astonishing anachronism. Today (1996) the sketch-book still has an advantage, in that it is cheaper and more easily portable than most other visualising media. However these practical advantages will soon disappear as the image quality, performance to cost ratio and cost per stored sketch of small, portable, battery driven image handling systems improve. Compared to modern electronic media, paper and pigment is passive in that the marks made on it are much harder to modify and control. The sketch-book provides only unstructured storage of an image. Compared to an electronic "paint" system paper, even with skilled manual support, provides a feeble and inflexible medium for manipulating the stimulus components of the sketch hybrid image. Some of the deficiencies of the paper sketch in performing its role as support for the mind are discussed briefly below.

Sketch attributes support a one to many descriptive - depictive translation (Chapter 3) by the use of semi-determinate sign systems. It has been argued that these access non-verbal, unconsciously descriptive, visual memory and that superimposed written notes elicit the "inner voice" and assist mental imagery generation. The use of semi-depictive signs to represent visually tolerant categories is fraught with difficulty, although in skilled hands "pentimenti" and lines representing contour fragments can be used to great effect. The pentimenti used so successfully by Toulouse-Lautrec for example to represent a rearing horse
(Illustration XXIX) would never do to represent the larger tolerance needed by Ingres to represent a pointing arm (Illustration XXVIII). Although often necessary, written notes on sketches have the disadvantage that they cannot be easily moved or made temporarily invisible. Their interference with depictive components is inevitable. This is a consequence of the fact that unlike marks made on machine media, those made on a paper sketch have no sub-surface processing support.

Unfortunately the mass technology of representation has become polarised in contemporary culture, towards the very descriptive (the printed word) or the very depictive (television and photography). Yet it has been known since the time of Aristotle (Yates, 1966) that human thought occurs at the interface between these two widely separated modes of representation, where as I have tried to show, intermediate types of representation are needed. Words (concrete ones at least) achieve most of their meaning by the mental images they elicit for the "inner eye" and pictures are useful to us to the degree to which we can extract from them descriptive propositions for the "inner voice". Despite the fact that the need to translate between verbal and pictorial information systems is central to our culture, research into alternative non-polarised systems which support this mental dialogue, has been sadly neglected. To support such an internal dialogue when visualising imaginary objects, artists are too often forced to improvise their own language and medium. To do this it is usually easier to use primitive but cheap and familiar media than to grapple with a complex technology which, though powerful, is often unsympathetic. The design of existing machine systems for artists still seems to take for granted the linguistic-pictorial polarity which controls its market.

In the hands of the skilled, the subtleties of line, the variety of cross-hatching and the unexpected range of colour and form provided by pencil, crayon, ink and paint have provided an enormous range of techniques for serendipitously accessing hidden visual knowledge and for inventive perception (Chapter 4). However, using traditional media it is difficult, and sometimes impossible, to manipulate meaningfully the position, size, orientation, shape or colour of sketch components. This is not only because erasure is difficult (though it may be) but because the medium does not support the necessary back-up copy of the visible components and there is no necessary record of the hierarchic structure of the represented scene, its objects, parts and attributes. In order to flexibly move, re-scale, repeat or modify any component of an image a concealed back-up, representation of that component is necessary.

Perhaps it is because the visible components of the paper sketch are relatively inflexible that, (as suggested in Chapter 5) they are used, consciously or unconsciously, as the stabilising skeleton of a hybrid image, in which the users
mind supplies the transient labile parts (as superimposed imagery) for rapid
manipulation. Is this too a rather less than ideal arrangement, or are there reasons
why an attempted machine improvement must fail? It is argued in Chapter 5, that
an important function of the sketch is to create appropriate spaces for superimposed
mental imagery. It is possible that future visualising systems, will attempt to
eliminate the vague uncertainties of the user's own imagery by providing detailed
machine generated substitutes from some vast, but specialised, pictorial data base
(Gray 1992). Even if this is not the intention, disciples of the Gibson-Hagen theory
of picture perception, which denies a central role for mental structures (4.2.2), are
unlikely to support a visualising system which deliberately encourages such mental
components (Smets 1992). At present the likelihood that a machine system could
ever rival the imagery generating capacity of the mind seems remote. Moreover
although it may be possible, it will surely never be desirable to "replace"
artificially the unseen imagery of the human mind since, this is the crucial and
unique source of individual experience - mysterious and unspecifiable but the life
blood of the visual arts. However it certainly does not follow that an improved
sketching system should not provide an ancillary image generating system. This
would be based upon a pictorial data-base, partly public and partly privately
collected, and controlled by each artist-user. There is nothing in the imagery
support function of the sketch which suggests that the only useful source of the
image components is the user's long term memory. There are clearly many ways
that the imagery support function might be improved, provided this function is well
understood.

It was proposed in Chapter 6 that sketches support selective attention, by
separating and temporarily isolating, different parts and attributes of the object
represented. Thus a sequence of different sketches facilitate without distraction,
different partial aspects of a visual task. Traditional sketches perform this function
well for some kinds of selection but less well for others. Thus the contour line is a
superb 30,000 year old invention, which facilitates the mental manipulation of
form without the distraction of surface reflectance, chromaticity and texture.
However other selective attributes are harder to control. A precise and
concentrated attention to colour subtleties is actually hindered by the system of
mixing paints on a palette since attention must be divided between the palette -
where colours are seen in the wrong context - and their mental image in the desired
context. Mental interference is here to be expected from the fact that the colour
image of the task and the colour percept of the palette compete for the same
resources in working memory (Logie, 1995). Passive media can provide little
support for the selective representation of attributes such as spatial frequencies that
need visual skill and cognitive effort to extract (6.8.2). Integral dimensions, such as colour value and saturation are by definition impossible to attend to selectively, although there are good reasons why artists might wish to (6.3). An active image processing medium could provide image filters to support these and other kinds of visual selectivity discussed in Chapter 6.

Another implication of the selectivity function as described in Chapter 6 is that it would be an advantage if the segregation of selected visual attributes were quickly and easily reversible. It was pointed out that whilst it may be advantageous to separate, for example, colour from form or global from local structure for selective thought, it is equally important to be able to re-combine them to explore their interaction in an integrated whole. For this purpose also the passivity of traditional media is a disadvantage.

The passivity of traditional media is an even greater handicap in the fifth function - metacognitive monitoring - discussed in Chapter 7. However, before improvements can be designed, it is important to understand how, according to the theory, sketches support inventive metacognition. The very act of drawing or colouring a perceived stimulus may facilitate the conscious control of visualising processes, which might otherwise have become automatic by over-learning. Artists' paper sketches often preserve the faint, only partly erased tracks of earlier mark making (Illustration XXXIV). Understanding how such attributes may serendipitously support the cerebral resources of the "Inner Artist" is essential for improved system design. Despite the claims of Alan Rawson (1969), it is far from clear that variations of mark-making (emphasis and weight etc) in artists sketches provide the temporal record, within a single sketch, necessary to amplify the capacity of the "inner artist's" own record for cognitive back-tracking and correction. Immediately after he sketched the "Nude with four left arms" (Illustration XXVIII) Ingres would have known the order in which the different alternatives were tried and hence the direction in which his "mind's eye" was moving. However days later he might easily have forgotten and despite the claims of Alan Rawson (1969), I can see nothing in the sketch which preserves that important information. It is probably for this (unconscious) reason that most documented studies of working practice reveal sequences of many unfinished sketches for a single project rather than many alterations on a few (1.7.1, 1.7.2). The opportunities for improvement are numerous. There aren't that many successful "nudes with four left arms".


There are now many computer based machine systems specifically intended to assist artists or designers to manipulate and to visualise an as yet non-existent
object. These can be roughly divided onto two classes, vector based solid modelling systems and raster based computer "paint" systems.

Vector based graphics systems represent images as two or three dimensional coordinates, structured as lines, edges, surfaces (e.g. as polygon nets) and volumes. These can be mapped and displayed at any scale, orientation or position on a raster display, or drawn on a pen-plotter. Solid modelling versions use perspective computation to allow the user to inspect the two dimensional projected views of stored three dimensional models of possible objects. Arrays of normalised coordinates are used to represent the three-dimensional structure of possible objects. Rigid shapes and other structures with a numerically defined geometry are much easier to represent in this way than the complex surfaces of flexible materials or the organic forms of natural objects, such as trees or the human body. For this reason solid modelling systems have to date found their main visualising applications in engineering structures, architecture, car styling and product design where the techniques used to represent three dimensional structures numerically have been mastered and experimental visualisation by building models or prototypes is either too time consuming or expensive. However research into solid modelling is advancing quickly. For example research in Japan and the U.S.A has shown that it is possible, in principle, to use materials science to simulate numerically the different folds and drapes of garment materials. So true solid modelling systems for the Fashion design industry may be available in the future. Current systems for fabric garment design are not truly three-dimensional, but solid modelling systems for shoe design have been in use for at least 15 years (e.g. Gerber systems).

Visualising systems for architecture, for example, allow the user to inspect the simulated three-dimensional appearance of a projected building from any selected viewpoint. Simulated movement of the viewpoint within the interior may provide visual insight impossible using other types of representation, such as scale models. The object of most of such systems is to reproduce, with the greatest possible realism, the array of luminances and chromaticities which would be presented to the eye of an observer standing in a particular position under specified lighting conditions. Simulating the lighting and reflectance of all surfaces visible from a particular viewpoint is, naturally, an extremely complicated processing task. Current graphic techniques use either "ray tracing" or "radiosity" algorithms. The former calculates the chromaticity and intensity of an imaginary ray of light as it strikes and is reflected from each visible point from the surface on its calculated path to the eye. The latter calculates the light intensity of each surface in a scene as a function of the intensity of all other surfaces. It is remarkable how well some of
these algorithms can perform (Illustration XXXVIII). Great computing capacity is needed and many variables, with unknown degrees of randomness must affect the result. In a natural scene, illuminated by diffuse light from the sun, every surface scatters varying proportions of selected wavelengths on to other surfaces in directions which in turn depend upon their variable specularity and texture.

From the view-point of simulated realism, one of the problems with conventional solid modelling systems is that they depend upon fixed view-point monocular perspective to calculate a series of static view points of the numerically represented models. Such views are un-realistic (as are photographs) whenever they are viewed, as they normally are, from the wrong angle or distance and with the normal eye and head movements. "Virtual Reality" systems, although still immature attempt to overcome this particular obstacle to realism by generating twin stereoscopic images assigned to miniature monitors worn over the head and eyes. By including sensors of gaze direction, controlled by the head, and hand positions, it is possible to generate the experience of movement in a simulated environment and of manipulating with hand movements, virtual representations of three-dimensional objects with a hand held input device. Clearly such systems will have immediate applications in games and in skill training systems such a flight simulators. Although currently too crude to replace existing modelling systems their visualising applications in the arts are being seriously researched (Smets 1992, Gray 1992).

There is no doubt that advanced solid modelling systems and perhaps their virtual reality offspring, will find an increasing number of applications in the near future (Greenburg 1991, Gray 1992). Clearly existing systems can fulfil useful functions in the visualisation and marketing of a design in its later refinement stages. The fact that they are "false indices" (C.S. Pierce, see section 3.3) may even give them a marketing advantage, Pseudo-photographs as used, for example, in the collaging of imaginary objects with real photographs of the environment in which they are intended to be seen. However as I argue in the next section, such systems will need to be drastically re-designed if they are to improve upon the cognitive functions of the paper sketch. For the objective of such visualising systems is to generate the most the detailed and convincing simulation of the "ambient light flow" which would be projected on the user's retina by an imagined object or scene as if it actually existed. This has little to do with the cognitive functions of the sketch. How does one make a "virtually real" object vague and incomplete? It is possible to imagine future three-dimensional visualising systems which respond to the movements of the user's head and hands, but which are also fuzzy and incomplete representations. However as long as the designers of such
systems interpret "reality" as meaning "very detailed and complete visual representation as close to the light pattern reflected by real objects as possible" then "indeterminate virtual reality" seems a contradiction in terms. The obvious application for such very realistic interactive imagery is in training, as in flight simulation systems or in creating virtual environments for engineering development. Detailed, non-selective realism is appropriate to the late stages of design and prototype illustration, but not to the early inventive stages. There is little evidence that such systems are intended to assist descriptive to depictive translation, for example, and it is clear that their data structures have not so far been designed to allow links between depictive and verbal constraints, or to address the problem of meaning with, for example, semantic nets or frames (Chapter 4). Yet, with present knowledge, the opportunity clearly exists for them to do so. Some of the selective functions of Chapter 6 are already possible, (if not convenient) with existing systems (e.g. "wire-frame" displays (outline form) without colour). However both the designers of such systems and their skilled users would probably be astonished by the suggestion that untidy incompletion or spatial fuzziness would assist the imagination, or that detailed realism can be a hindrance to the spatial imagery of imagined alternatives with which a prematurely over-specified image may compete (Chapter 5).

It must be remembered that the concept of "projected" imagery, and the idea that pictures support it, is regarded as "mischievous nonsense" by Gibson (1979 Chapter 15) and, it can be assumed, by his "virtual reality" followers. In short, detailed solid modelling systems have great potential (as they are improved) for refining and illustrating the later stages of those three dimensional designs which have already reached a degree of certainty and precision. They also have a potential for constraining, or at least by-passing, the user's inventive imagination. If they are used, as is sometimes advocated, to side-step the cognitive functions of rough sketches then there is (it seems to me) a danger that the very detailed "pseudo-index icons" that they generate will seduce their designer users and the companies who employ them, into believing that their products are more inventive or more appropriate than they really are.

Solid modelling and virtual reality visualising systems require a detailed numeric model from which to derive their computed images. It will therefore always be tempting (i.e. cheaper) to design by modifying an existing stored model than to conceive and input a radically new one. This is the approach of the design stylist rather than the innovator. Yet it is claimed that with such systems, architects (for example) can put aside their sketch-books and "work through the entire preliminary design process - sketching and exploring aesthetic alternatives, refining
Sketching and Cognition: Implications.

their designs and generating realistic images for analysis" (Greenburg, 1991). For this to be true future systems will have to be radically re-designed. There is nothing in the proposed theory which suggests that the functions performed by traditional sketch attributes cannot be performed better by modern technology. Indeed it is very unlikely that future machine systems will be designed to simulate all the media bound attributes of paper sketches. What is contested is the proposition that very realistic machine generated representation will make the cognitive functions themselves redundant; that for example it will no longer be necessary to use mental imagery, because the machine generated alternative is so much clearer. This implies that "virtual reality" will obviate visual imagery by facilitating the easy manipulation of the designed object itself (the "virtual" object) as oil paint can for a painter or clay for a potter. I believe this proposition to be based on a false conception of the visual arts. Although it is true that many painters manipulate their ideas directly in paint upon the canvas (Chapter 1) it is also true that paintings have become themselves substrates for cognitive imagery. This way of working, without sketching, seems to have become popular historically as finished works of art, augmenting the "viewer's share", acquired many of the privileged attributes of sketches (Chapter 2).

There is clearly a danger in all systems which use a model to automate some of the processes which determine the appearance of an imaginary object. The viewers share can become discounted and the user's inventiveness thereby distorted or constrained. Certainly this danger can be avoided, but only if it is well understood by those who design the system.

Vector based modelling systems have a theoretical advantage over raster based systems in that the numeric model on which they are based can be structured, (in principle at least) to reflect much of the mental structure so essential for visual thought. Thus scenes, objects, parts, surfaces, edges, colour dimensions and textures can, in principle, be represented separately and displayed or not displayed as required. This distinction between stored visual data and visible image is also reflected in the Kosslyn model of mental imagery (5.2.2) in which unconscious "deep" descriptive representations of object structure are used to generate conscious images in a spatially depictive medium. The advantages are many. A single normalised representation can be used repeatedly in different positions. Attributes of interest (e.g. colour, detail, parts) can be varied independently within a constant skeletal structure. Especially important, is the fact that a sub-surface representation is an essential prerequisite for the machine representation of the non-visible cognitive components of the sketch. This however does not mean that such representation will be easy (8.2.4).
A great disadvantage of existing solid modelling systems, for sketching purposes, is that the internal model is tedious to input. Manipulations of a view dependent appearance, (the purpose of the sketch as defined in Chapter 1), are indirect. The user manipulates the model and the machine generates the display. However this may be incompatible with a sketch's purpose in which appearance is imagined first, and then later the structure necessary to generate such appearance. In the early stages of design or fine art the images used to support visual thought, must be not only flexible and tolerant, but in many cases view dependent. Thus a means is needed to manipulate quickly and easily perceptual views which can support the users internal imagery. This is likely to be too vague initially to be used to generate a machine model. Moreover as already mentioned, the visual complexity of only a few types of designed objects are within the scope of current solid modelling technology.

Freehand sketching is easier and less constrained in raster based machine "paint" systems. Most computer graphic systems represent the displayed image as a "bit-map" - a list of binary coded numbers corresponding to the colours of the "pixels" or grid units into which the screen is divided. Vector based systems provide only indirect access to this bit-map but interactive raster based systems for artists provide a library of functions based on the direct manipulation by the user of the bit-map controlling the screen image. A powerful but very expensive system, using specialised hardware - Quantel's "Paint-box" (1980) has been influential as a design prototype for numerous other derivative but less expensive raster graphics systems which run on desk-top computers. I myself have introduced many hundreds of undergraduate students of fine art and design to such systems, particularly Noble Campion's "Cameo Paint" and "Cameo Image" software. In doing so I have been forced to reflect upon the advantages and disadvantages of using machine systems for sketching, and upon some of the extraordinary implications for art education. Two of the tentative conclusions I have come to are that

1. Future machine assisted drawing and colouring systems have the potential to perform some of the functions of the sketch much more effectively than traditional media. In the near future, inexpensive portable machines - perhaps with light weight touch sensitive screens and connected to digital cameras - will inevitably compete with the popularity of the sketch-book, though perhaps never completely replace it.

2. That for complex reasons, partly theoretical and partly commercial, the design philosophy used in most existing machine "paint" systems hinders their use as inventive sketching systems. Moreover they are not appropriately designed for
use in education. Their influence upon the teaching of visual skills and the growth of creativity in the inexperienced could even be harmful.

One reason for the first conclusion is the active flexibility of digitally stored images which can be copied, manipulated and recombined ("collaged") much more easily than their paper parents. In this sense the electronic medium which can store, process and generate digitally encoded images provides an as yet unrealised opportunity for a close co-operation with the neural medium of the user's brain, which also encodes, stores and processes descriptive and depictive images dynamically. Unlike an image on paper, images or parts of images, represented as bit-maps, can be copied to disk at intermittent stages of a sketching task, allowing easier back-tracking and recording than is possible with paper (function 5).

The digital representation of colour also allows many colour handling processes not possible with paper and paint but which are potentially useful enhancements of sketch function. Colours are represented by numeric triads which control the intensity of the red, green and blue phospors on the screen of the monitor. In some systems every pixel in the image has a separate R,G,B value in a table. More typically the screen pixels are mapped indirectly to R,G,B values in a numbered "colour table". A colour can be re-used exactly simply by pointing with a stylus to part of an existing image. (The inconvenience of re-mixing a colour exactly in oil-paint is illustrated by Bonnard's habit of mixing his colours on a sheet of glass, which was preserved under water over-night for re-use). Changing the R,G,B values of a colour table number, changes the colour of all the pixels in the image with that number. This allows a provisionally sketched colour to be adjusted in context, a great advance over the divided attention caused by the colour mixed out of context on a palette. (Unfortunately in their eagerness to simulate traditional media most "paint" system designers have chosen to use a pop-up imitation palette for adjusting colour by selection from a chart or even (Quantel's paint-box) simulating pigment mixing). The desk-top software "Photo Shop" (written for the Apple Mac) is intended for the manipulation of digitised photographs. It allows the specification of colour tolerances but these are not used to generate indeterminacy in the image but to compensate for indeterminacy. Thus tidy "flood-fill" operations are possible up to a boundary of all colours within a certain range. Clearly the algorithms used to process colour tolerances could be adapted for the inverse purpose, i.e. to allow the many to one mapping implicit in specifying colour categories (function one) and for applying colour with some random variability (functions 3 and 5). However they were not designed for this purpose.
Another reason for conclusion (1) is the capacity to manipulate and collage digitised photographs, combining the potential advantages of "contingent" representations with cognitively generated ones (see Chapter 3). Since Delacroix and the invention of photography (Scharf 1968), artists and designers have used photographs as source material when, or if, it suited their purpose, although this practice was strongly criticised in some circles (Wilenski, 1945). More recently artists such as Richard Hamilton and Andy Warhol have been interested in exploring the representational nature of photographs in the "no-man's land" between painting and photography. This is a subject matter of cultural importance. However until 1980 (Quantel's Paint-Box) artists who wished to explore the potential of combining the index qualities of the photograph with the expressive icon qualities of paint had to use the slow technology of the photographic enlarger and paper collage. Hamilton who states that he "sketches with photographs" was converted to the new medium in 1986 when he was invited with other artists to record a session with Quantel's paint-box for the BBC. ("Painting with Light" BBC 2 1986). He has now invested in a Quantel paint-box system which he uses regularly in his studio. However, the cultural questions raised by this way of working in which the "evidential" qualities of an index is manipulated as if it was an icon have hardly been answered (3.3).

Manual drawing and colouring with a raster paint system, using "mask rasters" which are intended to simulate "brushes", is cleaner and tidier than on paper with no "muddying" of the colour by those that are overlaid. Many other operations such as "flood-fill" are also much tidier than their paper counterparts, or at least they encourage neatness since the slightest (one pixel gap) in a drawn contour causes disastrous colour leaks. In order to understand the numerous drawing and colouring functions provided by such computer "paint" systems, it is helpful to remember that Quantel's influential system (and similar products) was originally designed for television graphics and to be used by mature designers who had learned their craft using traditional media - pencil, pen and inks, poster-paints, stencils, air-brushes and cut and pasted photographs etc. It is likely that their products had to be presented in a format where order, neatness and clarity were expected. Perhaps also the style of animation graphics was influential, where for technical reasons neatly enclosed shapes filled exactly with flat colour are necessary. Thus the design philosophy was to provide an electronically implemented tool-box which would simulate, as closely as possible, the traditional tools (pens, brushes, air-brushes, stencils, cameras) and media with which the designer users would be already familiar. The logic of the functions provided by this, and other paint systems, can also only be understood if it is assumed that they
are intended to simulate, enhance and accelerate the generation of those attributes of visual art believed, by the system designers, to be desirable and to avoid those seen as not desirable. Thus no attempt is made to simulate the untidy attributes of paper sketches, such as wobbly lines, colour application which flows erratically over a boundary, random but subtle variation of colour and texture, incomplete contours, smudges and blots and the imperfect erasure of an India-rubber. Yet, as argued here repeatedly, these attributes can have important cognitive functions. However simply imitating such media based attributes, begs the question raised by the "tool-box" simulation approach. Why should such a flexible and powerful machine medium be designed to imitate a much more limited one? One is reminded of tank commanders at the beginning of the second world war who continued to use the tactics and command vocabulary of equestrian cavalry, or of those Chinese Han dynasty pots made to imitate bronze (Honey, 1945). I will refer to this design philosophy as the "Imitation Bronze" approach. Computer based image handling systems are not "paint systems", they do not have any "brushes", they do not use "cutting and pasting" and they cannot mix colours. They do offer a new medium with colour and form handling possibilities which are unthinkable in any other medium.

The Imitation Bronze approach may be comforting to an older generation of artist users, but it may also result in software which fails to exploit the capacity of a new medium as fully as a younger and perhaps less prejudiced generation of artist users deserve.

The usefulness of raster based image processing to designers, has been advanced by the marketing of specialised systems adapted to specific kinds of design. One such system is "MicroDesign" produced by the Microdynamics Corporation of the U.S.A. for the Fashion design industry. Although it has no solid modelling capabilities, this system allows the realistic mapping of different fabric textures and colours onto a digitised photograph of a modelled garment, convincingly imitating the shadows and distortions of the patterns caused by the folds of the fabric (Illustration XXXVI). The manufacturers correctly claim in their brochure that "MicroDesign enables designers to modify sketched or photographic images and to alter design concepts by changing colours, patterns and textures. Designs can be transported electronically for use in desktop publishing. Completed designs can be used in catalogues, leaflets or advertisements. Text can be added to any picture making MicroDesign an effective sales tool". However their claim that it is more than this is controversial. "MicroDesign is a tool for unlimited creativity. Designers can produce and develop concepts quickly with many variations in colour, texture and dimensions. Microdesign is as easy to use as an artist's sketch
pad but can communicate ideas faster and more vividly than ever before". The rather vague implication behind the advocated use of both design specific solid modelling systems and design specific raster systems, is either that sketching is faster and easier with a machine system or that, with the improved visualising capacity of pseudo-photographs, rough sketching is an unnecessary stage of the design process which can be advantageously skipped, passing straight from the early descriptive concept to the manipulation of very realistic pseudo-photographs or virtual models.

Raster based machine systems allow the free manipulation of very detailed digitised photographs and other images with perfect erasure, very tidy colouring of complex shapes (provided they are neatly outlined) and magnified "touching up" of small detail. Therefore they have a natural application for illustrating the late stages of a design and for design styling, where existing designs in the form of digitised sketches, photographs or even video tapes can be modified. However they are not well designed to provide the subtle unexpected variation and controlled confusion which assists invention. It is of course possible to sketch with a machine system and to generate manually many (but not all) of the functional attributes of sketches. However the system designs fail in most respects to exploit the machine potential to improve support for sketch functions as much as they should.

Recently government sponsored research (in which I have participated), has investigated the use of computing in the Fashion design industry in the U.K, continental Europe and the U.S. It was found (Makirinne-Crofts et Al., 1992) that computer systems intended for the Fashion industry (such as Microdynamics’ "MicroDesign"), although in widespread use have not replaced the use of paper sketches and sketch-books by designers. Typically the machine system is operated by a trained specialist, who is not a designer, to produce illustrations and brochures for marketing purposes. When rough sketches are needed in machine readable format, they are "scanned" into machine memory, as unstructured bit-maps, from their paper medium. Most designers probably feel intuitively, that the advantages of digitised machine imagery are still out-weighted by the convenience of the sketch-book in the inventive early stages of design.

Besides its cheapness and portability, paper has several other advantages as a sketching medium over existing machine systems. Only very expensive systems can rival the subtlety, spatial resolution and chromatic texture of paper. Many Fashion designers like to spread many sketches over a large wall or table top for easy comparison (e.g. Muir 1984). Existing technology does not facilitate this. Although the much prophesied "paperless revolution" has taken longer to arrive than expected, such hardware problems are probably temporary. The next
generation of parallel processing micro technology may use the heralded flat "plasma" monitors which may allow designers to compare and discuss wall sized, multiple image, electronic displays of design ideas from a single portable sketching system. The competitive spread of global markets for design ideas combined with the rise of global information networks may in the near future, make it almost obligatory for designers to keep sketch stages of their design ideas, in machine readable form. Real progress in providing improved support for the sketch, will be made less with improving the hardware than with re-thinking and re-designing the software, to provide appropriate support for the hidden processes of visual thought.

I conclude that although the mark making freedom essential for any appearance based sketching system is only possible with raster based systems, the lack of any meaningful stored representation of the sketch structure and semantics makes it impossible for them to support the five cognitive functions of the sketch any better than paper media. Moreover because machine displayed images fail to generate the spontaneous, accidental subtlety, and surprise of paper media, they are in some respects less supportive of the "inner artist" (function 5) than traditional media. This does not mean that there are no reasons for sketching with machine media. There are many, as it is still desirable to manipulate electronically digitised sketches for illustration and communication. But because a purely raster based system typically discards much of the implicit structure represented in the act of sketching, much of the system potential for assisting invention, can not be exploited. A scanned-in bit-map from a paper sketch is just as useful for these non-inventive functions.

A new type of sketching system is needed, which combines the surface image manipulating freedom of the raster "paint" system with the descriptive structuring of visual ideas, that is only possible with a rich but flexible sub-surface data structure. This should be linked at a later stage of research to a design specific, user built, data base. Experience already available from linguistic "expert systems", which simulate some aspects of creativity may well be applicable to such visual systems. That it is possible to apply the ideas developed for task specific expert systems, based on propositional rules and descriptive data, to inventive drawing systems with a highly personal style, has been shown by Cohen with his "Aaron" series of programs (Cohen 1974, 1982, Boden 1992). "Aaron" is not a sketching system and it is not an interactive medium. The infinite variety of drawings it produces, as a result of Cohen's own image making rules, are much too personal in style to be useful to designers or other artists (Illustration XXXVII). Nevertheless the software techniques Cohen has invented to generate various types of drawing with surprising variety, are surely applicable to some of
the functions of a task specific "one to many" sketching system. Cohen states that the software he has developed over the last twenty years is intended to "investigate the nature of representation" (private communication 1991). I assume that most artists have a similar goal when they sketch but they are rarely prepared to make their unconscious rules of invention so explicit. Non-interactive expert drawing systems may provide part of the answer to the question "What components of visual thought (if any) can be automated?". They leave completely open the answer to the question "What components of visual thought is it desirable to automate?". Although the methods developed for expert systems may be useful for machine sketching systems, I believe that an attempt to design a machine which can simulate human visual invention would be misconceived. Much more interesting, is the possibility of using the methods of artificial intelligence to amplify, by harmonious cooperation with an artist user, the human brain's own capacity for visual processing.

9.4 Some Suggestions for the Design of Improved Systems

9.4.1 Sketching as a Mind-Machine Synergism.

It is beyond the scope of this study to analyse all the implications of cognition for the design of new sketching systems. Here I propose only to outline briefly some ideas for future research. The most promising approach, I suggest, is to research the design of a Mind-Machine Synergism in which a model of the mental representations and processes of visual thought is used to design matching representations (data structures) and processes (machine algorithms) within the machine. The term "synergism" has been chosen to express the hope that the performance in terms of appropriate visual invention of mind and machine cooperating on a visual task will be greater than the sum of the mind's contribution and the machine's contribution, because the system has been designed to amplify but not to replace the user's thoughts. This is more than can be expected from existing raster based "paint" systems, the output of which can be no more than the sum of contributions from the machine and the user's brain. This is a consequence of the "partitioned approach" to C.A.D. system design, dubbed by Negroponte the "idiot slave" approach (Negroponte, 1972). An improved approach would be to find ways to integrate the best of the "tool-box" algorithms with some of the techniques developed for "expert systems". However the purposes of such a "mind-machine synergism" would be quite different from either of these methodologies.
9.4.2 Machine Representation of Visual Structure

Sketch attributes work by supporting private representations of imagined scenes and objects within the user's mind. Apparently vague, they possess physical attributes which support and interact with the highly structured mental representations of their user. Neither the solid modelling systems nor the raster graphics systems, described above, can provide much support for the user's cognition because they lack suitable data structures and interfacing procedures.

Wherever the evidence suggests that a cognitive task depends upon limited mental capacity that needs sketch support, a non-interfering interfacing mechanism will be needed which allows the machine to represent, (and hence support but not substitute for), a particular thought process. The well designed system would lighten the user's workload and amplify mental capacity without distorting or constraining the user's goals. The raster graphics "tool-box" approach allows little representation of the user's mental structure of the object represented. Whilst the user "sees", in the mind's eye, a complex, but often vague structured hierarchy of imagined scenes, objects, parts, surfaces, edges, colours, lines and suggestive marks, the raster system only "sees" (i.e. represents) lists of codes representing coloured pixels. An improved system would combine the surface mark making flexibility of the raster "paint" system with the image structuring possibilities of the modelling system. A machine representation of structure, combined with a much richer mind-machine data sharing mechanism, is an essential step towards machine participation in the flexible manipulation of the image structure as the user actually sees it (see Function 4, Chapter 6). This implies a complex image data structure and user interface, but is not incompatible with the rapidly rising capacity of contemporary hardware. Machine support for all five of the sketch functions is handicapped by the lack of such an internal structure, available to but separable from, the representation of the visible sketch. Within (initially) a task specific domain (e.g. a Fashion sketching system for one kind of garment) it should also be possible to represent in machine data some of the implicit structural schemata which have been shown to be important in sketch recognition (4.3.8). The implementation of this possibility needs to be researched. One approach would be to design structures on the lines of Minsky's (1975) "frames" with slots for default parameters. Palmer (1975c) has also suggested possible structures to model mental schemata of object appearance. The descriptive to depictive translation function suggests that in addition it may be fruitful to research the possibility of adding task or object specific semantic nets such as those which have been used for quasi-intelligent verbal processing systems. For example the "ACME" and "WordNet" systems use a symbolic net of associated concepts to model analogical thought.

(Holyoake and Thagard, 1989). Further, but I believe rewarding, complications to
the machine data structure will be needed to represent the many types of sketch
indeterminacy. For example, before it is closely defined, an object's contour may
need to be represented not by one line but by many, within a specified tolerance.
Thus a vector representation of multiple contour lines would have internal pointers
to either a "shape tolerant" machine "pentimenti" generating procedure or to user
generated alternatives. Lines would be classified according to visual meaning.
Edges for example could be assigned to shapes and surfaces in ways which allow
alternative interpretations, (e.g to represent the fact that contour lines switch their
affinities if a perceived figure-ground relationship changes). These in turn might be
linked to multiple labels. Similarly, vector data structures representing contour
lines must allow for both visible and invisible (cognitive or imagined) contour
fragments. If these are represented internally as parts of the shape definition then
shape colouring and movement operations are possible without losing the
perceptual advantages of imagined boundaries.

A few possibilities for representing colour may serve to illustrate the
general point. Luckily much of the research needed for descriptive to depictive
colour translation already exists in the literature. Thus the raster machine
representation of RGB values can be mapped to the C.I.E. (Commission
Internationale de l'Eclairage) system of colour specification and then in turn to the
Psychophysical variables for the C.I.E. Standard Observer (Dominant wavelength,
Luminance and Chromatic purity (Le Grand 1968, Chapter 8)) which in turn have
been mapped (more controversially) to appearance notations systems such as
Munsell or the Natural Colour System (Hunt, 1977). Unfortunately none of the
well calibrated systems are very familiar to artists or designers, who often have
their own vaguer terminology. However adding user control of colour terminology
and user control of colour tolerances with machine assisted mapping to less tolerant
colour dimensions (e.g Munsell Hue, Value and Chroma) would be comparatively
easy and allow sketch spatial components to be represented internally categorically
with a specified range. There are several ways in which such a "vague" colour
could be visibly expressed in a particular region of the visible part of the hybrid
sketch image. One method would be to use entries to colour tables, which map in
turn to a selected portion of colour appearance space (in either 1, 2 or 3
dimensions) and then to the corresponding screen RGB values. The use of very
large colour tables would be a possible mechanism for representing and
manipulating in situ spatially separate regions with descriptive colour tolerances
(colour vagueness). Thus even if they had the same RGB values, separate colour
regions would have different table entries for independent in situ manipulation.
Attached to the RGB values of each table entry would a tolerance factor calculated from the user defined colour description.

By using such large colour tables, at least three ways of mapping a selected descriptive colour category to its many depictive instances can be envisaged.

1. Upon request the machine could generate from the colour description a series of partly random, spatially variegated pixel structures within a shape with soft or hard edges, and any of several alternative contours (defined by concealed pentimenti). To some extent this might imitate naturally coloured surfaces, which when digitised by a camera or scanner always have a scattered pixel variation. Pressing a key could provide endless variations of the surface colour, texture and edges as defined by the sub-surface tolerances. Note that the sub-surface representation of pentimenti offers a mechanism for optionally allowing an applied colour to overflow a visible boundary as it does on paper.

2. An alternative command would allow the user to sample, either sequentially or at random, any number of example colours within a specified range or which are implied by an attached colour description. Thus sample colour "instances" within a required tolerance can be assessed as possible alternatives before a more specific decision is taken. I regard the method of adjusting a coloured area *in situ* without the distraction of having to select from a "pop-up" colour menu, a requirement which follows from the selective attention function of the sketch and the limited capacity of visuospatial working memory.

3. Thirdly, the specified colour description may not displayed at all (it must be removable), with only an optional label used to indicate that a colour description has been assigned. The purpose of this option would be to allow space for imagined colour (see Watkins & Schiano 1982) In a true mind - machine partnership the machine software must be designed sometimes to work without user assistance (when the user's attention is needed elsewhere), sometimes collaboratively with user interaction and sometimes to withdraw, to allow non-interfering space for the user's imagery.

The data structures needed for implementing such a descriptive to depictive mapping of colour will be complex and need careful thought. For example the system must also allow the user to attach multiple colour descriptions to a chosen region, with an easy mechanism for switching from one description to the next. It is likely that a user will often be unable to decide whether a given part should, for example, be "pale green" or "pale blue". After attaching both descriptive labels to the part or region, the user after experimenting with the above descriptive mapping system, might then "prefer" three possible blues and three possible greens, eliminating several other colours, but wish to preserve some other possibilities
within the descriptive ranges. Thus a mechanism for priority listing of colour possibilities and for marking colours for later appraisal would be needed.

I have chosen colour as an example because it is relatively easy to see how existing machine colour structures might be altered and enriched to allow greater support for Functions 1 to 4. Support for Function 5 implies that each time a colour instance is considered it is recorded in a way which allows later backtracking to review and edit colour thoughts. The example also shows how the need to preserve descriptive tolerances in a way which supports successive refinement, multiple priorities and mind-changing, implies surprisingly complex and carefully designed machine data structures.

In principle, similarly improved descriptive and depictive data structures, with appropriate translation algorithms, should be applicable to the representation of spatial structure. Clearly there will be many difficulties in designing such data structures and even more in designing user interface techniques which do not interfere with the need to sketch quickly. Perhaps it will be considered rather perverse to propose that a visible image on the screen that is apparently sparse, and impoverished in its surface representation should need a data structure, richer and more complex in some respects than a very detailed and precise image. However the potential advantage of such a sub-surface representation is that it defines not one image but many, with a correspondingly expanded power to assist visual thought. The sub-surface structure of the visible sketch must be easy to enter and manipulate. It would need to have many default values linked, in some systems perhaps, to design specific semantic nets. It is particularly important that this sub-surface structure - a descriptively labelled depictive map, linked to a temporal record of sketching acts - should be easily available to the user for editing and monitoring. Thus a "sketch" would have two different mechanisms of user interaction, preferably but not necessarily on separate screens. I suggest, as signposts to further research, that at least four different but richly interconnected types of information structure would be necessary, accessible through a multimodal interface.

9.4.3 The Visualising Support System.

The data structure (and its associated library of manipulative algorithms), which is intended to assist the artist user to imagine the appearance of new visual ideas, is here termed the "Visualising Support System". It is used primarily to support Functions 2 and 3. This data structure corresponds to the most common type of sketch, where the user needs assistance to visualise an idea as it would been seen from a single view point. Because the sketching system must allow the
representation of indeterminate attributes with assigned of tolerances and descriptive categories there will be many possible "surface" representations of a particular sketch subject matter, even when limited to one view-point. Individual sampled instances of such sketch appearance, (the 2½D hybrid image) is here termed an "Appearance Instance". A user controlled display would support the user's super-imposed imagery and provide a fast mark-manipulating interface. Its function (as with paper sketches) is to generate visual stimuli which support and amplify the perceived and imagined components of the user's depictive imagery. It is displayed upon a surface upon which the user can project and manipulate internally generated images of imaginary objects. All surface image making activities would be displayed in a way designed to support this mental task with as little attentional distraction (pop-up menus etc) as possible. Nothing on the screen or within the periphery of the visual field, should interfere with the sketch imagery function or distract attention from the users global or local imagery task.

One of the most important components of this module to research, will be the collaborative, "mind-machine" processes which up-date this "surface" representation. As its visible components are up-dated by manual drawing and colouring, a mechanism is needed to simultaneously edit the implicit structure, attribute tolerances and descriptive notes in the "sub-surface" representation of the sketch, the Cognitive Structure Map, which are only displayed when needed. This will be one of the tasks of the multi-modal interface described below.

9.4.4 The Cognitive Structure Map

This is a machine implemented "engineering diagram" of a user's visual concept, used to manipulate and generate multiple "Appearance Instances" of the user's visual imagery. A second interactively controlled screen would display, (and provide interactive user control of), the machine's representation of the descriptive spatial structure (relationships between scenes, object parts and attributes etc) and the chromatic structure (using user preferred notation systems). The structure map would probably take the form of a descriptively labelled viewer-centred map or drawing in which the user could see and manipulate descriptive and non-visible (i.e. cognitive) depictive components of the sketch. The structure of the image in terms of objects, parts, edges, attributes and colours would be displayed diagrammatically and with labels. Descriptive or depictive notes could be attached to different elements of the structure and recalled optionally by pointing to the part or attribute concerned. This would obviate the goal conflict which occurs in paper sketches between the need to add descriptive labels to the sketch and the need to avoid attentional conflict with super-imposed imagery. In addition the structure
map would provide spatially linked access to a task related data base of communal and user chosen depictive and descriptive information. The fact that so much of the information that artists and designers need when sketching, is only useful if it can be attached to meaningful parts or attributes of the sketch, is another reason why the structure map is necessary. Thus, instead of pinning a random piece of fabric to a sketch, a Fashion designer might simply attach the index or look up key to a detailed depictive-descriptive fabric data-base. The structure map would provide the designer with rapidly inter-changeable access to fabric data and allow a range (using the category representation system mentioned above) of fabrics to be mapped onto the appropriate garment part, using a MicroDesign style fabric mapping method. When the user manipulates or edits the structure map directly, rather than by editing instances in the Visualising Support System, this will sometimes result in corresponding alterations to the displayed visualising image, allowing the user to see immediately the visual consequences of the change. However other information attached to the structure map will only show on the displayed image indirectly, such as for example, when it manipulated. Thus a part previously inseparable from an object may become detached and moveable. The operation of erasure becomes more complex but also more powerful. Thus a part or region can be erased from an "appearance instance" without necessarily removing its representation in the structure map thus making it easily regenerated. Also, if the erasure is one of several alternatives, of multiple contour lines for example, then the user could alter the priority and displayed emphasis of the temporarily erased part (e.g. by making it fainter). Finally an erased part can be erased from the structure map itself but can still be recovered by back-tracking to a sketching branch node using the action monitor. Naturally a carefully thought out data management system will be needed to see that these different ways of editing a sketch structure are consistent with each other.

In some tasks, vectors corresponding to visible or invisible contours might be adjustable in meaning with respect to the figure ground structure of the object. Such structural ambiguities would need double representation. It is suggested that as with colours contours, shapes and other object attributes can be represented descriptively, with machine generated algorithms to sample depictively such descriptive attributes within user tolerances. Thus every structure map of an early stage of visual thought would generate many different sketch depictions or "Appearance Instances" of the concept. Appearance Instances could be generated at random (to create serendipity "awareness surprises" (Chapter 7)) and gradually reduced in number by successive refinement. Here I am suggesting that image handling technology can offer, within a single stored data structure, improved
control over the stages of the successive refinement that, on paper, would require many pages of a sketch book. At different stages of refinement the artist user may wish either to save a sequence of sketch structure maps (each of which can generate multiple depictive alternatives) for later refinement or she or he may wish to save particular appearance instances or better, a number of refined or severely pruned versions of the structure map. Thus the structure map is an intermediate representational type of the kind I have argued is essential to visual invention (3.5). The difference from paper sketches is that whereas on paper the translation function must be performed by the same spatial array as the perceptual and imagery functions, with undesirable side-effects (such as written notes interfering with spatial imagery), machine image processing offers the possibility of separating them and at the same time vastly improving the selectivity function discussed in Chapter 5. (Illustrations XIV and XXIII show attempts to use the same sketch to fulfil both depictive and structural functions at once.) Of course the effectiveness of this idea depends upon having an efficient mechanism for updating, in both directions, the effects of editing either of the two representations. All the "data structures" here proposed, are envisaged as program "objects", the properties of which are given by the procedures ("methods" in object oriented programming language) with which they are packaged. There are nearly always two ways, as a consequence of the "mind-machine synergism" approach, in which the Structure Map can implement the one to many mapping of descriptive structure to depictive instances. One is to automate the mapping process under user supervision. The other is to provide assisted mechanisms for users to enter sketch variations themselves. In a practical system both machine and user generated alternatives would have to be catered for. For example when pentimenti (multiple alternative contour lines) are indicated to express an indeterminate edge, a machine algorithm, which can generate these from a user's contour prototype, might be useful, provided the user has easy access to the various parameters which determines the way it performs. However there must also be a way of saving and representing economically user drawn pentimenti representing a single object edge or edge segment. This would of course produce extra interfacing demands. Whilst sketching the user must be able to effortlessly register with the structure chart when a contour line is a new edge segment and when it is an alternative to an existing one. Where the domain of a descriptive category is well defined, automated "instance" generation quite feasible and is used routinely, for example, in generating different "weights" (shape variations) for type face designs and has been used experimentally in generating shape variations for glass-ware design. Descriptively encoded colour is an obvious case for automatic instance generation.
For example, a user might enter (or pick) "grey-brown" on the structure map and attach it to the corresponding structure component (part, surface, area, ground) and is then able to view as many instances as desired on the sketch depiction (at random or systematically). The usefulness of such procedures can only be determined experimentally and will depend upon the degree of task sensitivity with which they are designed. Recently I have been experimenting with the usefulness of a curve interpolation algorithm, Overhauser’s Parabolic Blend, (Rogers & Adams 1976, pages 133-138), for predicting and representing the invisible but imagined components of fragmented contours, which as suggested in Chapter 4 may be useful in inventive retrieval. These would be visible for entry or editing on the Structure map but not usually on the Sketch Depiction. By using a separate colour table entry for such imagined contours, with the same RGB values as the background colour, such invisible contours can remain available for shape colouring and moving operations.

9.4.5 The Action Monitor: A Record of Sketching Acts.

It was argued in Chapter 7, that one of the functions of the sketch was to support the users' awareness of their own thought processes. This ability has been shown in some tasks, at least, to be necessary to assess one's own cognitive performance and is implicated in theories of creativity. It has already been argued that whilst unconsciously automatic mental processes have an important role in creativity, they can also inhibit the discovery of novel directions for visual thought. To avoid such "sphexishness" the brain supplies a limited capacity monitoring and control system which I call the cognitive monitor. I have, perhaps controversially, related this hypothetical processor to models of voluntary, as opposed to automatic, mental processes and to Baars' "global workspace" theory of consciousness (Baars 1988). It is also clear that paper sketches do a rather poor job of supporting this monitoring facility. By recording the sequence of sketching acts and allowing easy editing access to this record, machine systems could be designed to support the user's visual monitoring and awareness much more effectively than they do at present.

Many raster "paint" systems provide a "record" facility which can, on demand, preserve a file of the movements of the user's actions, which can then be replayed and interrupted at a certain point. However the usefulness of such recordings is usually limited by the lack of any meaningful structure for backtracking and the poor editing facilities offered to the user. In fact it is not usually clear what the function of such records is. (Quantel's recording system is sometimes used for animation). The use by Scrivener and Clark, S. (1992) of a
sequential record of descriptive and depictive sketching acts to investigate the nature of co-operative design, on remote work-stations using a shared interface, has been mentioned already. They made the interesting observation that an automated file structure based on pauses of a certain length, usually coincided with meaningful sketch actions. However these actions needed to be classified and related in detail to the descriptive and depictive components of the user's visual thought. What is implied by the visual monitoring function of the sketch is that the user's short term memory of what he is doing and thinking in a particular task can benefit from spatial and temporal support. ("How can I know what I imagine until I see what I sketch?"). In the schematic system proposed here, this function is managed and recorded by the Action Monitor. The Action Monitor records operations on both the Cognitive Structure Map and the Visualising Support System and would generate a hierarchic event data structure. Because the function of the action monitor is to assist artist users to back-track on their own thought process and to re-consider earlier decisions, the natural structure of the sketch monitor files is not a single sequential file, but a densely branched tree of actions linked at important (user selected nodes) to stored structure maps and depictive instances. Because, like the search trees in a chess playing program, monitor files would have the tendency to become unmanageably large, it would need a well designed, machine assisted, user controlled editor with the capacity to prune unwanted pathways, to compress a complex sequence of actions and to select which structure map nodes and appearance instances should be preserved. The nature of the depictive parts of such a record and its editing is beautifully illustrated in Clouzot's film (1954) of Picasso painting on glass. The value of such sequential records was in fact discussed by Picasso with Roland Penrose before the film was made (Penrose 1962).

9.4.6 A Multi-Modal Mind-Machine Interface

The sketching system advocated here, implicates a much richer machine representation of the spatial structure, visual attributes and descriptive information attached to sketches. Therefore a more sophisticated interface will be needed than the type usually provided with machine "paint" systems for example. The accepted interfacing wisdom for CAD systems, judging by current practice, (Shneiderman, 1987) seems to be that when the user's main task is visually controlled then the machine operating task should be also. However the hybrid-image hypothesis of Chapter 5 indicates clearly that such a single-mode approach to interfacing is misguided. Only spatially guided input which is compatible with the user's task related mental imagery should be used. When attention must be divided between
cognitive tasks, especially when one of these tasks is non-spatial, then it is easier to perform the two tasks in different modes (Brooks 1968, Baddeley and Lieberman 1980). One of the implications of Chapter 6 is that a multimodal interface would assist the user to divide attention, where necessary between visuospatial thought and the simultaneous input of descriptive labelling or structuring operations. When David Hockney, Howard Hodgkin, Richard Hamilton and other artists were experimentally sketching with Quantel's paint-box for the BBC, they seemed to have very little difficulty with simultaneously commenting verbally on what they were doing ("Painting with Light" BBC 2, 1986). This is consistent with research into divided attention tasks which suggests separate (parallel) processing modules for the "inner voice" and the "inner eye" (5.3.6). Thus cognitive theory suggests that speech recognition systems (provided they have reached the required reliability) may solve some of the divided attention problems which will certainly arise, if the desired descriptive-depictive machine interaction is to be harmoniously implemented. An analogy with driving a car suggests foot-pedals might also be useful. An implication of the fact that percepts and images use shared resources, is that the working memory mode of any interfacing task which is not supportive of the user's sketch imagery task should be in a different mode from that task. Thus visuospatially contingent colour should be adjusted with a non visuospatial interfacing mechanism such as pen pressure and angle (such pen are available) or voice recognition. Clearly a greater sharing of visual structure between mind and machine will need a richer, more complex interface. I see no escape from this conclusion.

9.4.7 Descriptive and Depictive Domain Specific Data-Bases.

Although the sketch cognitive structure map will support written notes and labels, a task specific sketching system would be much more useful if it were linked to a well structured data base of technical and client or user, task specific information. Information directly attached to the structure map would be specific to a particular visual task and the set of depictive appearance instances which it can generate. Information in the data-base on the other hand would support the descriptive and depictive information sources, which artists or designers need or draw upon, to support and stimulate their creative thought and which are potentially common to many different sketching projects. Clearly designing such a data-base is a major task, since its usefulness will depend not only upon the quality of the information (variable and visual) provided, but also upon the sensitivity of the design of its structure, its access keys and its dictionaries, to the needs of the user. It has been suggested that human concept formation is a product of the re-
organisation of human memory (Baddeley, 1990). A goal of sketching systems is to amplify inventive concept generation in the minds of artists. It follows that machine support for an artist's long term memory must also support the meaningful re-structuring of concept related information. I suggest that initially it will be most practical to implement two data-base modules a descriptive verbal one and a structured depictive store of digitised drawings, illustrations and photographs.

Typically, descriptive labels in the structure map would contain index references both to the descriptive data-base and to a design specific version of the user's favourite word processor. With extensive subject cross-referencing this would have the obvious advantage of eliminating the need to repeat notes used on many sketches and of enabling fast editing of notes attached to labelled sketch attributes and parts. One can envisage design specific versions of popular word processors and data-bases with "Structure Map" entries in their option menus.

Designing the visual data-base is an even bigger task, since pictorial information is much more demanding of machine memory. However a sketch related pictorial data-base becomes more attractive as cost/capacity performance of the appropriate technology improves. Currently laser disks (CD ROM) are available which provide fast access to an entire video film. Unfortunately they are read only devices. However historically the increase in machine capacity has always over-taken the development of appropriate soft-ware to use it (Michie & Johnston, 1984). So now is the time to consider how artists might routinely exploit not only private stores of digitised images, but also public libraries of digital photography and video available on global networks. Sketch references to private image libraries would include access to the user's previous sketches and to illustrations, photographs and edited video passages, that have been used, or might be used, to supplement memory and refresh hazy mental imagery. All sketches and sketch structure maps relevant to a single task, or a set of related tasks would be linked into an easily edited tree structure which documents the history, including the "blind alley ways", of a user's ideas. Designers refer frequently to their own earlier ideas both successful and unsuccessful. For example Fashion designers sometimes report that they keep drawers filled with used sketch-books for later reference (interviews for Makirinne-Crofts et Al., 1992). Robert Welch (1986) keeps an barn loft filled with prototypes of uncompleted designs, which he visits to stimulate his ideas for new products.

Another reason for including sketch structure maps, with each family of related sketches in the data base, is that the structure map provides the information
necessary for retrieving separately, significant sketch components and their attached descriptive information.

Because of their "index" nature, untampered photographs present a completely different problem of representation from structured sketches. Their disadvantage is that they contain no structure. Such images would be much more useful, I suggest, if descriptive structure is added to them by the machine-mind partnership. On the other hand, in so far as such cultural indices duplicate patterns of light reflected by real objects, they provide a valuable reservoir of those kinds of visual surprise that only nature can offer and which, I have suggested, is necessary to awaken inventive consciousness (7.4.3 & 8.3). The kinds of image processing necessary to make them compatible with sketches should, I believe, be properly considered to be functions of the photographic data-base or at least its interface with the sketch structure map, and not (as in machine "paint" systems) part of the user's depictive image handling function. How much of this image structuring can be assisted by machine and how much must be performed co-operatively, will depend upon progress in image analysis and "machine" vision. Machine procedures can, for example, provide filters for colour dimensions supporting selective attention to hue, saturation and lightness relationships as discussed in 6.7. Also the photographic interface should be able to provide spatial frequency filters which simulate the spatial frequency filters of the human brain, supporting spatial attention as discussed in 6.8.2. Other structuring functions such as the isolation of edges, can be assisted by machine (Marr 1982, Watt 1988) but still need the user's cognitive recognition systems to supply the "top-down" components needed for meaningful structure. Thus a mind-machine sharing algorithm is indicated. Finally the user's eyes and brain are needed, (using machine assisted masking techniques) to segment a photograph into its usable objects and parts for future collaging, re-scaling and linking to sketch structure maps. The data-base management system would sort, classify, label and outline, for user inspection, all such pictorial components. In other words the data base would have attached to it an image processing interface which assists the user to enter appropriate pictorial structure maps, linked to and compatible with, the cognitive structure maps of a relevant family of related sketches. Task related depictive information could then be accessed and manipulated by the user as meaningful sketch components, (via the Structure Map and the Data-base indexing system). It would be worth researching the possibility of using such structured photographic components as temporary semi-transparent overlays, to be re-scaled and oriented to overlay sparse depictive sketch instances, as a means of refreshing or supplementing (but never replacing) the user's mental imagery.
An outline schema for a sketching system in which mind and machine collaborate. The output from the "sketch entry system", consists of an hierarchy of files of sketching actions, sketch "structure maps" and "appearance instances". This output can be developed and refined using the studio based development system (see text).
9.4.8 Two Separable System Components

One of the reasons for sketching directly at the machine interface (as opposed to sketching on paper) is the opportunity to capture quickly the spatial, chromatic and sequential structure of ideas, in such a way that they can be interactively re-processed later on at leisure. It would not be necessary and probably not feasible to implement the whole of the above system in a single unit. Rather it makes sense to divide the system into two interdependent but optionally separable components (Diagram VIII) -

1. A light-weight, portable sketch entry system but with a large secondary memory capacity. This would support user control of the Visualising Support System, and would automatically record the raw data needed for the Cognitive Structure Map and the Action Monitor. Whilst sketching the user would be able to complement machine updating of the structure map and the action monitor files with simple non-spatial commands (tactile or verbal). When necessary the user would have to interrupt his view of the visualising screen to, view and correct structure maps and monitor files. It would need a user interface as versatile as current technology can provide with manual pressure sensitive stylus and key-board plus digital camera and perhaps a voice recognition system and "data glove". This would be the "field" unit, to be constantly on hand wherever or whenever its artist or designer user has an idea or finds a visual motif.

As the machine records and files sketching "acts" the interface would be designed to allow the easy entry of depictive "tolerances" (degrees of vagueness) for sketch attributes as they are entered (contours, shapes, positions, sizes, orientations, colour dimensions etc). In addition, as and when convenient, the user would switch to a representation of the sketch Cognitive Structure Map to edit visual structure, add alternatives, edit tolerances and insert descriptive notes and labels with meanings known to the studio system. This last process of communicating private information to the machine (quickly before it becomes unconscious) is a price which must be paid, (as when giving instructions to a human assistant), for the use of machine processing capacity to support those mental processes which are normally private. Whether or not artists and designers will feel this price is worth paying will depend, of course, on the quality and usefulness of the added imagery and descriptive structure that it enables..

2. This in turn will depend upon the power, appropriateness and adaptive flexibility of the second component, the library linked Sketch Development System. Records and files from the above portable sketch entry system would then be "developed" in the studio which would contain a the sketch Development System. It is inevitable that in the rapid sketch entry of "first thoughts"., the sketch
user will sometimes omit many of the necessary links between sketch "surface" components and the underlying descriptive structure. It is therefore essential that there will be ample facilities to up-date Cognitive Structure Maps, re-assign links to the corresponding set of sketch appearances and to add further attribute tolerances and descriptive notes, later, as and when they come to mind. The post processor would also provide interactive processing links (mind-machine synergisms) to the user's idiosyncratic, private libraries of depictive and descriptive information and (ideally) to net-worked public libraries of application specific, pictorial and verbal data. I have included here descriptive to depictive and depictive to descriptive translation systems linked to the library data-bases. The former would provide machine (post sketch entry) support for the user's mental imagery generation system. It would develop, by user controlled access to the two data-bases, the one to many mappings implicit in the users' category labels, tolerances, written notes and abbreviated references encoded in the structure map attached to each sketch. Thus a single sparse rough sketch could by post-processing or studio mind-machine development generate a large family of descent images as the user's ideas are refined. The latter (the depictive to descriptive processor) would exploit image analysis techniques to assist the user to segregate raw depictions and edited stills or sequences of video imagery into their most relevant visual attributes and components. Not only would this system support selective attention, (Function 4, Chapter 6), but it would allow user controlled manipulation and analysis of a rich source of "natural surprise" (Chapter 7.3). This processor would be designed to support (but not replace!) the user's perceptual and attentional image descriptive processes and would include reversible attribute filtering as suggested in Chapter 6.

Finally the development system would be used for re-structuring, editing and pruning the sets of structure maps, sketch instances, sketch fragments and written notes which belong to a particular sketching session or project and which have been recorded in a branching sequential tree, the Action Monitor File, by the sketch monitor. By re-playing and editing the monitor files, a temporal record of the user's thought processes whilst sketching, would become structured into a branching hierarchy of possible ideas, which will need drastic pruning as the ideas become clearer and more constrained. As mentioned above, each structure map with its descriptive tolerances will represent many sketch instances any one of which can be generated at will and specified economically by simply recording its descriptive parameters. Thus editing a sketching session would be more like editing a film than a drawing, except that the images generated, would become
progressively sharper and richer (more descriptive and more depictive) as the session progresses.

9.4.9 Outline Schema of a Mind-Machine System

Diagram VIII shows how these components might be linked in a sketching system. Throughout this study I have concentrated on the functions of sketches in the very early stages of design or visual composition. I have argued that existing computer raster "paint" and solid modelling systems are more useful for the later than the early stages of visual creation. However digital image handling technology is well suited to the combined support of both the early and the late stages of visualising in a single integrated system.

The "Portable Sketch Entry System" is intended to combine the manipulative freedom of the electronic raster "paint" system (the Visualising Support System) with a super-ordinate, vector based representation of both the visible and invisible structure of the sketch subject matter (objects, parts, attributes, surfaces, edges etc) and descriptive tolerances specifying multiple options for the sketch components (the Cognitive Structure Map). The Structure to Appearance Translator implements a one to many mapping from the structure map tolerances to the many implicit images (simple or complex, clear or confused) which can be displayed (and edited) by the visualising support system. This is envisaged as combining two methods of translation. Where feasible machine algorithms will optionally generate multiple attribute exemplars within a specified tolerance (see 9.4.2). However it must always also be possible for the user to specify and record in the structure map chosen alternatives (as is sometimes found in paper sketches Illustration XIV). The Structure to Appearance Translator is also used to segregate or filter image components in order to generate on request some of the many possible partial images which support selective attention, as discussed in Chapter 6. This process would always be reversible. Any individual image generated by the Appearance Translator is here termed an Appearance Instance. Visual parts and attributes within any Appearance Instance may be unitary, as when only one of multiple contour lines (pentimenti) are displayed, or multiple as when many pentimenti are overlaid. As any Cognitive Structure Map can, depending on the specified tolerances and options, be translated into a small or an astronomically large number of Appearance Instances, there must be a easily used facility for the user to file, for later consideration, a limited sub-set of those specific instances of special interest (as discussed for colour in 9.4.2). One of the advantages of having the structure map linked to a machine implemented appearance translator, is that each appearance instance can be encoded simply by
reference to a specific cognitive structure map and a record of the attribute parameters used by the translator to generate a specific image. Thus any image generated can be recorded easily for later "regeneration".

The sketch entry system also supports the Action Monitor which files depictive and descriptive sketching acts as described in 9.4.5. Similar encoding economies are also possible for the information monitored. Each action monitor file, which would normally correspond to a single sketching "session" related to a single task, has a one to many logical relationship with a set of structure maps. However this set of data records are related to each other and can be regenerated by a stored sequence of editing actions. However because the purpose of the monitor file is to facilitate mental back-tracking and re-thinking, an Action File would have the structure of a complex tree with multiple branches. This could easily become unwieldy and uneconomic. Therefore a user control editing system, with a dedicated interface would be provided. This would allow the editing of priority branch points (for re-considering earlier ideas etc) and the "zapping" (after security back-up) of unwanted branches of the action tree.

Thus the sketch entry system will generate and save a hierarchy of three types of file, all of which will represent sketch ideas with some uncertainty or vagueness. The action monitor file allows the review and replay of a family of cognitive structure maps and a priority set of recorded "appearance instances" (ie. sketches for visualising). However, because it represents degrees of attribute "vagueness" each Cognitive Structure Map supports many possible appearance instances which are unlikely to have been envisaged by the user during the early stages of sketch entry, when there will have been the most indeterminacies. In fact only a very small amount of indeterminacy, if repeated locally will result by permutation, in an astronomically large number of possible sketch instances (see discussion Chapter 1.5). This ensures that the system can generate those media related "perceptual surprises" which are so thin on the ground in existing machine software design. A second phase of visual contemplation in which more descriptive and depictive information can be referred to, and more detailed visual options considered, will therefore be needed. The studio based Development System would support the extensive and detailed testing of preliminary ideas. Many more of the options, implicit in unexplored branches of the action monitor files, can then be sampled and further, as yet unconsidered, visual implications, implicit in the structure map, considered and developed. Visual ideas will become refined and less vague as "branches" in a recorded sequence of the monitor action tree are pruned, and visual or descriptive options in the structure map are edited. As the user's ideas becomes clearer, so links to the studio data-bases (9.4.7) will make, initially
rather sparse cognitive structure maps, gradually richer, with attached depictive and descriptive detail. It is at this later point that the techniques of solid modelling and very realistic image handling may be applicable. Drawing upon its links to images and models in the data-base, a late stage Cognitive Structure Map might then be used to supplement, but not to replace, the user's mental imagery by generating a stream of task associated physical imagery.

9.5 Conclusions.

Research into the human brain and its astonishing capacity to support visual thought and invention, will continue to be one of the most exciting and controversial areas of research in the next century. Obviously the theory I have tried to document, is tentative and incomplete. I had hoped to implement a sketching system or at least some small part of one to illustrate my ideas. Alas the literature I felt relevant to this task has proved to be so extensive that, even after reading 1070 books and papers, I know have only skimmed the surface of the subject matter. Thus this study is intended to be used as a conceptual frame-work for others to develop and apply as they see fit. It is certain that the study of cognition, especially visual consciousness, will continue, in the next twenty years, to provide exciting and unexpected insights into the visual arts and the design of machines to assist them. Artists and designers deserve better support from information systems than they have yet received. System design based on rule-of-thumb knowledge and the Imitation Bronze approach is far too flimsy a foundation on which to base such an important task. I hope therefore that this study will influence others to look at the literature of cognitive science and to see if they agree with me that this vast, barely tapped, reservoir of ideas is not only relevant but necessary if we are ever to design better, more sympathetic systems for the visual arts.

Predicting a future software famine, Michie and Johnston (1984 page 20) quote figures to show that since 1965 the growth of computer hardware performance to cost ratio has outpaced programmer productivity by a ratio of 250,000 to 1! But there is another even more important comparison to be made, for which they quote no figures. How fast have both software and hardware capacity outpaced the growth in our knowledge and understanding of the user's own mental processing capacity? The successful application of information technology to the design of machine visual processors for artists must surely be linked to a better understanding of those unique and irreplaceable biological processors that they are designed to support. The Congress of the United States of America have apparently declared the 1990s the "Decade of the Brain" (Kosslyn and Koenig, 1992). It is to be hoped that this new initiative, applying ever more
powerful technologies and human resources to discover how the human brain works (dubbed "Cognitive Neuroscience") will also lead to a better understanding as to how machines can be used not to replace, but to enhance, mental inventiveness in all those directions which support and enrich our culture. If the theory presented here is correct, then the reason why visually inventive people need to sketch is itself a product of the brevity of our culture and the pace of technological change. Natural selection has had too little time to evolve specialised neural processes in the brain which are as well adapted for imagining future artifacts as those which evolved in our hominid ancestors for perceiving and acting on present events. Instead the modern brain has been forced to "borrow" the ancestral visual system and adapt it for the purpose of manipulating stored knowledge to invent the future. It is because of its biological origins that the visual brain needs culturally invented sensory support. According to this interpretation very rapid cultural change has led to an improvised and perhaps imperfect adaptation of the human species' most precious bodily possession. Thus these culturally "stranded" functions of the brain are very deserving and appropriate candidates for advanced technological support. But to provide this support it will not be enough to understand how the human brain works it will also be necessary to understand why it works the way it does. Without such an understanding we must frame the best theories that we can. Artists' sketches and their history provide some of the clues.
BIBLIOGRAPHY


Arnheim, Rudolf (1970)
"Visual Thinking" Faber and Faber. (London)

Arnheim, Rudolf (1974)
"Art and Visual Perception (New Version)" University of California.

Ayrton, Michael (1957)
"The Act of Drawing"

Baars, B. (1988)
"A Cognitive Theory of Consciousness"
Cambridge University Press

Bacou, Roseline (1975)
"Millet. One Hundred Drawings"
Phaidon Press Limited.

Baddeley, A.D. (1986)

Baddeley, A.D. (1990)
"Human Memory: Theory and practice"
Lawrence Erlbaum Associates, Hillsdale, New Jersey

"Working memory" In G.H.Bower (Ed) "The Psychology of Learning and Motivation" Vol. 8, Academic Press

Baddeley, A.D. and Lewis, V.J. (1981)
"Inner active processes in reading: The inner voice, the inner ear and the inner eye" In "Interactive Processes in Reading" A.M. Lesgold & C.A. Perfetti (Ed.) pp 107-29 Erlbaum: Hillsdale, N.J.


"Critical limiting factors in the design of the eye and visual cortex : The Ferrier Lecture 1980"
Proceedings of the Royal Society 212, pp 1-34.

Bartlett, F.C. (1932)
"Remembering. A Study in Experimental and Social Psychology"
Cambridge University Press 1964 (5th printing of 1932 edition)

"Strategic Study to Investigate the Opportunities and Difficulties in Fashion Design for Automated manufacture" Report of research Sponsored by The ACME Directorate, S.E.R.C. Polaris House, North Star Avenue, Swindon, SN2 1ET.

Beck, J. (1972)
"Surface Colour Perception"
"Visualisation of compound scenes"
Perception, 7, pp 129-138

Biederman, I. (1972)
"Perceiving real world scenes"
Science, 177, pp 77-80

"On the semantics of a glance at a scene"
"Perceptual Organization" Lawrence Erlbaum, New Jersey (1981)

Biederman, I. (1987)
"Recognition by Components: A theory of Human Image Understanding"

Biederman, I. & Cooper, E.E. (1991a)
"Evidence for Complete translational and reflectional invariance in visual
object priming"

"Priming Contour-deleted images: Evidence for intermediate representations in
visual object recognition"
Cognitive Psychology 23, pp 393-419

"Size Invariance in Human Shape Recognition"
Journal of Experimental Psychology: Human Perception
and Performance 18, pp 121-133.

"Priming Depth Rotated Image Objects: Evidence for Three_D Invariance"
Manuscript referred to in Biederman and Cooper, 1991a)

Bisiach, E. (1988)
"The (haunted) brain and consciousness" In Marcel, A.J. & Bisiach, E. (Ed.)

Bisiach, E. (1992)
"Understanding Consciousness: Clues from Unilateral Neglect and Related
Disorders" In Milner, A.D. & Rugg, M.D. (Ed.) "The Neuropsychology of
Consciousness" Academic Press.

"Unilateral neglect of representational space"
Cortex, 14, pp 129-133

Bisiach E., Luzzati G. and Perani, D. (1979)
"Unilateral Neglect, Representational Schema, and consciousness" Brain 102:
pp 609-618.
Black, M. (1937)
"Vagueness: An exercise in logical analysis"

Blake, R. (1989)
"A Neural Theory of Binocular Rivalry"

"Artificial Intelligence and Natural Man" (2nd ed.)
London: MIT Press.

Boden, M. A. (1992)
"The Creative Mind"
Abacus. London.

Braitenberg, V. (1985)
"Charting the Visual Cortex"
In A. Peters and E. G. Jones (Ed.)
New York: Plenum Press.

"Mental image reversal and verbal recoding: When ducks become rabbits."
Memory & Cognition, 21, pp 23-33

"Experimental evidence for scanpath eye movements during visual imagery"

Breitmeyer, B.G. and Ganz, L. (1976)
"Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing"

Brennan, S.E. (1985)
"The caricature generator"
Leonardo, 18, pp 170-178

Bronowski, J. (1973)
"The Ascent of Man"

Brooks, L.R. (1968)
"Spatial and verbal components of the act of recall"
Canadian Journal of Psychology 22, pp 349-368

Bruce, V. (1988)
"Recognising Faces"
Essays in Cognitive Psychology. Erlbaum Associates, New Jersey

Bruce, V. and Green, P.R. (1986)
"Bruner, J.S. (1957)
"On Perceptual readiness"
Psychological Review, vol 64, pp 123-152

"Bruner, J.S. (1979)
"On knowing: Essays for the left hand"

"Bruner, J.S. (1990)
"Acts of Meaning"
Harvard University Press

"Bruner, J.S. and Postman, L. (1949)
"On the perception of incongruity"

"Bruner, J.S., Postman, L. and Rodrigues, J. (1951)
"Expectation and the Perception of Color"
American Journal of Psychology, 64, pp 216-223

"Bundesen, C. & Larsen, A. (1975)
"Visual transformation of size"
Journal of Experimental Psychology: Human Perception and Performance 1, pp 214-220

"Mental transformation of size and orientation"

"Selective suppression of the magnocellular pathway during saccadic eye movements"

"Carmichael, L., Hogan, H.P., & Walter, A.A. (1932)
"An experimental study of the effect of language on the reproduction of perceived forms"
Journal of Experimental Psychology, 15, pp 73-86.

"Carpenter, P.A. & Eisenberg, P. (1978)
"Mental rotation and the frame of reference in blind and sighted individuals"
Perception and Psychophysics, 23, pp 117-124

"Cave, K.R. & Kosslyn, S.M. (1989)
"Varieties of size-specific visual selection"

"Can mental images be ambiguous ?"
Journal of Experimental Psychology: Human Perception and Performance, 11, pp 317-328

"The Mind's eye in chess" In Chase, W.G. (Ed)
"Cheeseman, J. and Merkle, P.M. (1985)

"Chevreul, M. E. (1839)
"The Principles of Harmony and Contrast of Colours and their Application to the Arts"

"Chomsky, N. (1957)
"Syntactic Structures"
The Hague Mouton.

"Chomsky, N. (1965)
"Aspects of the Theory of Syntax"
Cambridge MIT Press.

"Rules and Representations"
Columbia University Press.

"Church, R.M. & Church, K.W. (1977)

"Matter and Consciousness"
Cambridge: MIT Press.

"Clark, K. (1939)
"Leonardo da Vinci"
Cambridge University Press.

"Cogniat, Raymond (1970)
"Raoul Dufy"

"On purpose: an enquiry into possible roles of the computer in art"
Studio International 187 (962) pp 9-16.

"How to Make a Drawing". Science Colloquium Nat'l Bur. of Stds.


"Corballis, M.C. (1988)
"Recognition of disorientated shapes"
- Corballis, M.C. and Nagourney, B.A. (1978)  
  "Latency to categorize disoriented alphanumeric characters as letters or digits"  
  Canadian Journal of Psychology, 32(3) pp 186-188.

  "Why are there so many visual areas?"  

- Cowey, A. (1985)  
  "Aspects of cortical organization related to selective attention and selective impairments of visual perception: A tutorial review"  
  In Posner, M.S. & Marin, O.S. (Ed.) "Attention and Performance Vol XI"  
  Lawrence Erlbaum : Hillsdale N.J

- Cozens, Alexander, (1785)  
  "A new method of assisting the invention in drawing original composition of landscape"  

- Craik, F.I.M. (1977)  
  "Depth of processing in recall and recognition"  
  In Dornic (ed.) "Attention & Performance VI" Erlbaum.

- Craik, F.I.M. & Lockhart, R. S. (1972)  
  "Levels of processing: A framework for memory research"  
  Journal of Verbal learning and verbal behaviour", 11, pp 671-684

- Craik, K. (1943)  
  "The Nature of Explanation"  
  Cambridge University Press.

- Crawford, Alan (1983)  
  "The Yamazaki Serving Collection  
  Drawings by Robert Welch."  
  Skelton's press, Wellingborough.

- Crick, F. (1994)  
  "The Astonishing Hypothesis"  

  "The Problem of Consciousness"  

  "Are we aware of neural activity in primary visual cortex?"  

- Critchlow, Keith (1976)  
  "Islamic Patterns"  
  Thames and Hudson. London.

- Davidson, J.M. & Davidson, R.J. (Eds) (1980)  
  "The Psychology of Consciousness"  
  New York: Plenum Press.
  "Spatial frequency selectivity in cells of the macaque visual cortex"
  Vision Research, 22, pp 545-59

  "Scanning mental images generated from verbal descriptions"
  European Journal of Cognitive Psychology 1, pp 293-307

- Dennett, D.C. (1981)
  "The Nature of Images and the Introspective Trap"

- Dennett, D.C. (1991)
  "Consciousness Explained"
  Allen Lane. The Penguin Press.

- Deregowski, J.B. (1973)

- Deregowski, J.B. (1977)
  "Pictures, symbols and frames of reference"

- Diehl, G. (1958)
  "Henri Matisse"

  "Normalisation of Irrelevant Dimensions in Stimulus Comparisons"

- Donne, John (1628)

  Lawrence Erlbaum :Hillsdale N.J

- Duensing, S. and Miller, B. (1979)
  "The Cheshire Cat Effect"
  Perception 8: pp 269-273

- Duncan, J. (1985)
  "Selective attention and the organisation of visual information"

- Eco, Umberto (1976)
  "A Theory of Semiotics"
  Bloomington 1976.
"The Remembered Present"
Basic Books, New York

Egeth, H. E. (1966)
"Parallel versus serial processes in multidimensional discrimination"
Perception and Psychophysics, 1, pp 245-252

Ellis, A.W. and Young, A.W. (1988)
"Human Cognitive Neuropsychology"
Erlbaum Associates, Hillsdale, New Jersey

Ellis, W. D. (Ed) (1938)
"A Source Book of Gestalt Psychology" Routledge & Kegan Paul

Engel, E. (1956)
"The role of content in binocular resolution"
Amer. Journal of Psychology, 69, pp 87-91

Engel, E. (1958)
"Binocular fusion of dissimilar figures"
Amer Journal of Psychology 46, pp 53-57

Engel, E. (1961)
"Binocular methods in psychological research"
In Kilpatrick, Franklin, P. (Ed.) "Explorations in transactional psychology"
University Press, New York. 1961

Engel, F.L. (1971)
"Visual conspicuity, directed attention and retinal locus"
Vision Research, 11, pp 563-576

"Visual conspicuity and selective background interference in eccentric vision"
Vision Research, 14, pp 459-471

"Protocol analysis: Verbal reports as data"
MIT Press: Cambridge MA

Evans, J.B.T. (1980)

Evans, R.M. (1974)
"The Perception of Color"

"A Handbook of Cognitive Psychology"
Lawrence Erlbaum Associates.

"The neurological basis of mental imagery: A componential analysis"
Cognition 18, pp 245-272
Farah, M.J. (1985a)  
"The Psychophysical Evidence for a Shared Representational Medium for Mental Images and Percepts"  
Journal of Experimental Psychology: General Vol 114, No 1 pp 91-103

Farah, M.J. (1985b)  
"A Left Hemisphere Basis for Mental Imagery?"  
Neuropsychologia Vol 23, No 1, pp 115-118

Farah, M.J. (1988a)  
"Is visual imagery really visual? Overlooked evidence from Neuropsychology"  

Farah, M.J. (1988b)  
"The Neuropsychology of Mental Imagery: Converging evidence from Brain-damaged and Normal Subjects"  

"Mechanisms of Imagery-Perception Interaction"  

"Structure and strategy in image generation"  

Farah, M.J. and Ratcliff, G. (Eds.) (1994)  
"The Neuropsychology of High-Level Vision"  
Lawrence Erlbaum Associates.


"Electrophysiological evidence for a shared representational medium for visual images and visual percepts"  
Journal of Experimental Psychology: General, 117, pp 248-257

"Visual angle of the mind's eye before and after unilateral occipital lobectomy"  

The effects on speeded classification of implicit and explicit instructions regarding redundant dimensions  

"Levels of Equivalence in Imagery and Perception."  
Psychological Review, 87, pp 113-132.
Finke, R.A. (1985)
"Theories relating mental imagery to perception"
Psychological Bulletin, 98, pp 236-259.

"Mental imagery and the visual system"
Scientific American, vol 254, No 3 (March 1986) pp 76-83

"Principles of Mental Imagery"

Finke, R.A. (1990)
"Creative Imagery, Discoveries and Inventions in Visualization"

"Mental Imagery Acuity in the Peripheral Visual Field"
Journal of Experimental Psychology: Human Perception and Performance Vol 6, No 1, pp 126-139

"Mapping the Visual Field in Mental Imagery"
Journal of Experimental Psychology: General 110, pp 501-507

"Area and Contrast Effects Upon Perceptual and Imagery Acuity" Journal of Experimental Psychology: Human Perception and Performance Vol 7, No 4, pp 825-832

"Directional scanning of remembered visual patterns"
Journal of Experimental Psychology: Learning, Memory and Cognition, 9, pp 398-410

"Visual functions of mental imagery"

"Explorations of creative visual synthesis in visual imagery."
Memory and Cognition 16: pp 252-257.

"Reinterpreting visual patterns in mental imagery"
Cognitive Science, 13, pp 51-78

Fish, J.C. (1991)
"A Model for the Mental Representation of Sketches"

Fish, J.C. (1994)
How Sketches Work: Bibliography

"Fish, J.C. and Scrivener, S. (1990)  
"Amplifying the Mind's Eye: Sketching and Visual Cognition"  

"Fish, J.C. (1981)  
"Colour as Sensation in Art and Science"  
Leonardo, Vol 14, No 2, pp 89-98.

"Fodor, J. (1983)  
"The Modularity of Mind"  
MIT Press: Cambridge MA.

"Mapping human visual cortex with positron emission tomography"  

"Franks, J.J. & Bransford, J.D. (1971)  
"Abstraction of visual patterns"  
Journal of Experimental Psychology, 90, pp 65-74

"Freyd, J.J. (1987)  
"Dynamic mental representations"  

"Facilitation of length discrimination using real and imagined context frames"  

"Frisby, J. P. (1979)  
"Seeing. Illusion, Brain and Mind"  

"Galton, Francis (1883)  
"Inquiries into Human Faculty and its Development"  
Macmillan (and also J.M. Dent)

"The Processing of Information and Structure"  
Lawrence Erlbaum Associates, Hillsdale New Jersey.


"Integrity of stimulus dimensions in various types of information processing"  

"Gassier, Pierre (1975)  
"Les Dessins de Goya" Volume II  
How Sketches Work: Bibliography

"Brain modularity: towards a philosophy of conscious experience" In
"Consciousness in Contemporary Science"
Oxford Science Publications.

-Georgeson, M. (1979)
"Spatial Fourier Analysis and Human Vision"
In Sutherland (Ed.) (1979)

-Ghiselin, B. (1952)
"The Creative Process"
New York: Merton.

-Gibson, J. J. (1937)
"Adaptation with Negative Aftereffects"
Psychological Review, 1937, 44, pp 222-244.

-Gibson J. J. (1950)
"The Perception of the Visual World" Boston Houghton-Mifflen
Lawrence Erlbaum, London

-Gibson, J. J. (1966)
"The Senses Considered as Perceptual Systems"
Houghton-Mifflen

-Gibson, J. J. (1971)
"The Information Available in Pictures" Leonardo Vol 4, pp 27-35

-Gibson, J. J. (1979)
"The Ecological Approach to Visual Perception"
Houghton-Mifflin: Boston

"Spatial firing and visual form perception"

"Je suis le cahier : The Sketchbooks of Picasso"
Royal Academy of Arts (Exhibition Catalogue), London

-Goldscheider, Ludwig (1960)
"Rembrandt. Paintings Drawings and Etchings"

-Goldschmidt, Gabriela (1991)
"The Dialectics of Sketching"
Creativity Research Journal Volume 4(2) pp 123-143
- Goldschmidt, Gabriela (1992)
  "On Figural Conceptualization in Architectural Design"
  In proceedings of the 11th EMCSR, Vienna, April 1992

- Goldwater, R. & Treves M. (1976)
  "Artists on Art From the 14th to the 20th Century"
  John Murray. London

- Gombrich E.H. (1952)
  "Leonardo's Method for Working out Compositions" contribution to Congress

- Gombrich E.H. (1962)
  "Art and Illusion" Phaidon Press

- Gombrich E.H. (1966a)

- Gombrich E.H. (1966b)
  "The Renaissance Conception of Artistic Progress and its Consequences" in
  "Norm and Form" cited.

- Gombrich E.H. (1982a)
  "Image and Code" In "The Image and the Eye"
  Phaidon Press

- Gombrich E.H. (1982b)
  "Mirror and Map" In "The Image and the Eye"
  Phaidon Press

- Goodman, N. (1968)
  "Languages of Art : an Approach to a theory of Symbols"
  Indiana :Bobs-Merrill

  "Theories of Visual Perception"
  John Wiley & Sons :New York

  "Filtering and Condensation tasks with integral and separable dimensions"

  "Psychophysics of Spatial-frequency Channels"
  In M.Kubovy & J.Pomerantz (Eds), "Perceptual Organization"
  Lawrence Erlbaum: Hillsdale, N.J

  "Attending to Spatial Frequency and spatial position of near threshold visual patterns"
  "Attention and Performance Vol XI" Erlbaum :N.J

- Gray, Stephen (1992)
  "Virtually Real?" Symposium. Virtual Representations for Design and
  Manufacture Coventry University.
"Computers and Architecture"

"Gregory Richard (1969)
"Eye and Brain"

"Gregory Richard (1970)

"Gregory Richard (1972)

"Gregory Richard (1973)
"The Confounded Eye" In Gregory, R., L. and Gombrich, E.H. (Eds)
"Illusion in Nature and Art". Duckworth.

"Gregory Richard (1980)
"Perceptions as Hypotheses"
Philosophical Transactions of the Royal Society of London, Series B, 290 pp 181-97

"Visuoimaginal constructional apraxia: On a case of selective deficit of imagery" Brain and Cognition, 5, pp 255-267

"Hagen, M.A. (1981)

"Hagen, M.A. (1986)
"Varieties of Realism : Geometries of Representational Art" Cambridge University Press

"An outline of an investigation into the special character of Pictures" In Mehler & Walker (Ed.) "Perspectives of Mental Representation", Erlbaum. Hillsdale. N.J.

"Haldane, J.B.S (1928)
"On Being the Right Size" In 'Possible Worlds' Chatto and Windus : London.

"Consciousness: An afterthought"
Cognition and Brain Theory, 5(1) pp 29-47.

"Hauser, Arnold (1951)
"The Social History of Art"

"Hawkins, H.L. & Shigley, R.H. (1972)
"Irrelevant information and processing mode in speeded discrimination" Journal of Experimental Psychology, 1972, 96, pp 435-437
Hayes, J.R. (1973)  
"On the function of visual imagery in elementary mathematics"  
In Chase, W.G. (Ed) "Visual Information Processing"  

Hayes-Roth, F. (1979)  
"Distinguishing theories of representation: A critique of Anderson's  
'Arguments concerning mental imagery'  
Psychological Review, 86, pp 376-392

Hebb, D.O. (1949)  
"The organisation of behaviour"  
New York: Wiley.

Helmholtz, H. von (1910)  
"Treatise on Physiological Optics"  
(Edited by James P. Southall English edition 1925, 1967)  
Dover Publications New York (1967)

"Autobiographical statements on creativity by artists, musicians and writers"  
In M.S. Lindauer (Ed.) "Research in Psychological Aesthetics and the arts: A  

Hering, Ewald. (1964)  
"Outlines of a Theory of the Light Sense"  
(Translated by Leo M. Hurvich and Dorthea Jameson) Harvard University  

"Consciousness in Contemporary Psychology"  
Annual Review of Psychology, 31, pp 1-26

"Parallel models of associative memory"  
Erlbaum Associates, Hillsdale, NJ.

Hinton, G.E. (1979)  
"Some demonstrations of the effects of structural descriptions in mental  
imagery"  

"Developing the concept of working memory"  
in G. Claxton (Ed.) "Cognitive Psychology: New Directions"  
Routledge and Kegan Paul.

Hobbes, Thomas (first published 1651)  
"Leviathan"  

Hochberg, J. (1968)  
"In the mind's eye"  
In R.N. Haber (Ed) 'Contemporary Research in Visual Perception" London,  
Holt, Rinehart and Winston Inc.
How Sketches Work: Bibliography

Hochberg, J. (1980)
"Pictorial Functions and Perceptual Structures"
In Hagen, M. (Ed.) "The Perception of Pictures"

Hochberg, J. (1982)
"How big is a stimulus?" In Beck, J. (Ed.) "Organization and representation in perception"
Erlbaum: Hillsdale, N.J.

"Form perception: Experience and Explanations"
"Figural Synthesis" Lawrence Erlbaum Associates

Hochberg, J. and Brooks, V. (1962)
"Pictorial recognition as an unlearned ability: A study of one child's performance"

Hockett, C.D. (1960)
"The Origin of Speech"
Scientific American, 203, pp 88-96

"Parts of Recognition", Cognition, 18, pp 65-96

Hoffman, J.E. (1980)
"Interaction between global and local levels of form"
Journal of Experimental Psychology: Human Perception and Performance, 6, pp 222-234

"Spatial selectivity in visual search"
Perception and Psychophysics, 30, pp 283-290

Hofstadter, D.R. (1985a)
"On the Seeming Paradox of Mechanizing Creativity"
In "Metamagical Themas" Basic Books U.S.A.

Hofstadter, D. R. (1985b)
"Variations on a theme as the crux of creativity"
In "Metamagical Themas" Basic Books U.S.A.

Holding, D.H. (1985)
"The Psychology of Chess Skill"
Erlbaum: New Jersey.

"Analogical mapping by constraint satisfaction"
Cognitive Science, 13 pp 295-356

Honey, W.B. (1945)
"The Ceramic Art of China and other Countries of the Far East"
London: Faber and Faber
  "Introduction to the Graphical Kernel System (GKS)"  

  "Ferrier Lecture: Functional architecture of Macaque Monkey Visual Cortex"  

  "Brain Mechanisms of Vision"  

  "Anatomical Demonstration of Orientation Columns in Macaque Monkey"  
  The Journal of Comparative Neurology vol 177, No 3 pp 361-379.

  "Segregation of form, colour and stereopsis in primate area 18"  
  Journal of Neuroscience, 7, pp 3378-3415.

  "Dynamic Binding in a Neural Network for Shape Recognition"  

  "Consciousness Regained"  
  Oxford: Oxford University Press

  "The Inner Eye"  
  London: Faber & Faber.

  "Normal and Pathological Processes in Visual Object Constancy"  
  In Humphreys, G. & Riddoch, M., J.(Eds)  
  "Visual Object Processing: A Cognitive Neuropsychological Approach"  
  Erlbaum Associates, New Jersey

  "Cascade processes in Picture Naming"  
  Cognitive Neuropsychology, 5, pp 67-104

- Humphreys, G.W. and Bruce, V. (1989)  
  "Visual Cognition. Computational, Experimental and Neuropsychological Perspectives"  
  Erlbaum Associates, New Jersey

  "The Specification of Colour Appearance. I Concepts and Terms"  
  Color Research and Application 2, pp 55.

  "Object recognition using alignment"  
  M.I.T. AI Memo 937.
- Jackendoff, R. (1983)
  "Semantics and Cognition"
  Cambridge M.A.: MIT Press

- Jackendoff, R. (1990)
  "Consciousness and the Computational Mind"

- James, William (1890)
  "Principles of Psychology" (Volumes 1 & 2)

  "Propositional Representations, Procedural Semantics and Mental Models." In
  Mehler, J., Walker, E.C.T. & Garrett, M. (Eds) "Perspectives of Mental

  "Mental Models"
  Cambridge University Press.

  "A Computational analysis of consciousness"
  In "Consciousness in Contemporary Science"

- Jolicoeur, P. (1985)
  "The time to name disoriented natural objects"

- Jolicoeur, P. & Kosslyn, S.M. (1983)
  "Coordinate systems of visual long-term memory representation" Cognitive
  Psychology, 15, pp 301-345.

  "Effects of Orientation on the Identification of Simple Visual Patterns"

  "Curve tracing: A possible basic operation in the perception of spatial
  relations"
  Memory and Cognition 14: pp 129-140.

- Jonides, J. (1980)
  "Toward a model of the mind's eye movement"
  Canadian Journal of Psychology, 34, pp 103-112.

  "Voluntary versus automatic control over the mind's eye movement" In

  "Integrating visual information from successive fixations" Science, 215, pp
  192-194
"Jonides J., Kahn, R. & Rozin, P. (1975)  
"Imagery instructions improve memory in blind subjects"  

"Judd, D. B. and Wyszecki G. (1963)  
"Color in Business Science and Industry"  

"Julesz, B. (1985)  
"Preconscious and Conscious Processes in Vision"  

"Julesz, B. and Miller, J.E. (1975)  
"Independent spatial-frequency tuned channels in binocular fusion and rivalry."  
Perception, 4, pp 125-143.

"Jusczyk, P.W and Klein, R.M. (Eds) 1980  
"The Nature of Thought. Essays in Honor of D.O. Hebb"  
Erlbaum Associates, Hillsdale, New Jersey

"Kahn, Gustave (Ed.) (1971)  
"The Drawings of Georges Seurat"  

"Changing views of attention and automaticity"  

"Kanizsa, G. (1976)  
"Subjective Contours"  

"Kanizsa, G. (1979)  
"Organisation of Vision : Essays on Gestalt Psychology"  
Praeger, New York.

"Amodal completion : Seeing or Thinking ?"  
In Beck, J. (Ed.) "Organization and representation in perception" Erlbaum: Hillsdale, N.J.

"Karmiloff-Smith, A. (1990)  
"Constraints on representational change: Evidence from childrens' drawing"  
Cognition, 34, pp 57-83

"Katz, David (1935)  
"The World of Colour"  

"Sight and Mind" Oxford University Press
  "Explorations of representational momentum"
  Cognitive Psychology 19: pp 369-401

  "The Role of vision in "visual Imagery" experiments: Evidence form the
  Congenitally Blind"
  Journal of Experimental Psychology: General 112 pp 265-277

  "The order of visual processing 'top down', 'Bottom up' or 'middle out'"
  Perception and Psychophysics, 33, pp 225-231

- Kline, M. (1954)
  "Mathematics in Western Culture"
  George Allen and Unwin Ltd.


- Koffka, K. (1938)
  "Principles of Gestalt Psychology" Routledge & Kegan Paul

- Kohler, K. (1947)
  "Gestalt Psychology" New American Library

- Kolers, P.A. (1972)
  "Aspects of Motion perception"
  Pergamon, New York.

- Kosslyn, S.M. (1975)
  "Information representation in visual images"
  Cognitive Psychology, 7, pp 341-370

- Kosslyn, S.M. (1978)
  "Imagery and internal representation"
  In Rosch and Lloyd (Eds) "Cognition and Categorization" Erlbaum.

- Kosslyn, S.M. (1980)

- Kosslyn, S.M. (1981)
  "The medium and the message in mental imagery",

- Kosslyn, S.M. (1983)

- Kosslyn, S.M. (1987)
  "Seeing and Imagining in the Cerebral Hemispheres: A Computational Approach" Psychological Review, Vol 94, No 2, pp 148-175

- Kosslyn, S.M. (1994)
  "Image and Brain. The resolution of the Imagery Debate"
Kosslyn, S.M. and Koenig, O (1992)
"Wet mind: The new cognitive neuroscience"

Kosslyn, S.M. and Shin, L.M. (1994)
"Visual Mental Images in the Brain: Current Issues"
In Farah, M.J. & Ratcliffe, G. (Eds.) (1994)

Kosslyn, S.M. and Shwartz, S.P. (1977)

Kosslyn, S.M., Alpert, N.M., Thompson, W.L., Maljkovic, V., Weise, S.B.,
"Visual mental imagery activates topographically organised visual cortex: PET

Kosslyn, S.M., Ball T.M., and Reiser, B.J. (1979)
"Visual images preserve metric spatial information: Evidence from studies of
image scanning" Journal of Experimental Psychology: Human Perception and
Performance, 4, pp 47-60

"Individual differences in mental imagery ability: A computational analysis" In
"Visual Cognition" M.I.T. Press

Kosslyn, S.M., Cave, C.B., & Gabrieli,J.D.E. (1989)
"Imagery in the blind side: Three case studies"
Unpublished manuscript, Harvard University. Cambridge, M.A.

Kosslyn, S.M., Cave, C.B., Provost,D. & Von Gierke,S. (1988)
"Sequential processes in image generation" Cognitive Psychology, 20, pp 319-343

Kosslyn, S.M., Chabris C.F., Marsolek C.J., & Koenif O. (1992)
"Categorical versus Coordinate spatial representations: Computational analyses

Kosslyn, S.M., Koenig, O., Barrett,A., Cave,C.B., Tang,J., & Gabrieli,J.D.E.
(1989)
"Evidence for two types of spatial representations: Hemispheric specialization
for categorical and coordinate relations" Journal of Experimental Psychology: Human Perception and Performance", 15, pp 723-735.

"Generating visual images"
J. exp. Psychol. Gen. 112, pp 278-303

Kotov, A. (1971)
"Think like a Grandmaster" (Translated by Bernard Cafferty) Batsford Press;
London.
"Does pattern matching require the normalization of size and orientation?"
Perception and Psychophysics, 30, pp 24-28

"Kulikowski, J.J. (1978)
"Limit of single vision depends on contour sharpness."

"Spatial frequency content of Visual Imagery"
Perception and Psychophysics, 28(6), pp 555-559.

"The effect of visual angle on the global and local reaction times depends on
the set of visual angles presented"

"Langer, E.J and Imber, L.G. (1979)
"When Practice makes Imperfect: Debilitating effects of Overlearning"

"Size scaling in visual pattern recognition"
Journal of Experimental Psychology: Human Perception and Performance, 4, pp 1-20

"Lawson, Brian (1990)
"How Designers Think" Butterworth :London.


("LeDoux, J. E. & Hirst, William (1986)
"Mind and Brain: Dialogues in cognitive neuroscience"
Cambridge University Press.

"Leeuwenberg, E.L.J. (1982)
"Metrical Aspects of Patterns and Structural Information Theory" In Beck, J. (Ed.) "Organization and representation in perception"
Erlbaum 1982

"Le Grand, Y. (1968)
"Light, Colour and Vision"
Cahpeman and Hall:London.

"Leonardo da Vinci (1270)

"Leslie, A.D. (1987)
"Pretense and representation: The origins of 'theory of mind'"
Psychological Review, 94, pp 412-426

"Levelt, W.J.M. (1965)
"On Binocular Rivalry"
Soesterburg, The Netherlands: Institute for Perception RVO-TNO.
  "Two visual systems in mental imagery: Dissociation of "what" and "where"  
in imagery disorders due to bilateral posterior cerebral lesions"  
Neurology, 35, pp 1010-1018

- Lindauer, M. S. (1983)  
  "Imagery and the Arts" In Sheikh, A. A. (Ed) "Imagery: Theory, Research,  
and Application" New York, Wiley.

- Livingstone, M. S. (1988)  
  "Art, Illusion and the Visual System"  
Scientific American, Volume 258, No 1, pp 68-75. (jan, 1988)

  "Psychophysical evidence separate channels for the perception of form, color,  
movement and depth"  
Journal of Neuroscience, 7, pp 3416-3468.

- Logie, R. H. (1986)  
  "Visuo-Spatial processing in working memory"  

  "Visuo-Spatial Short-Term memory: Visual Working Memory or Visual  
Buffer ?" In Cornoldi, C. & McDaniel, M. (Eds)  
"Imagery and Cognition" Berlin: Springer-Verlag.

  "Visuo-Spatial Working Memory"  
Lawrence Erlbaum Associates.

  "Visuo-spatial working memory: Visual, Spatial or Central Executive?" In  
Logie, R.H. & Denis, M. (Eds) "Mental Images in Human Cognition"  
Amsterdam: North Holland Press

- Logothetis, N.K. and Schall, J.D. (1989)  
  "Neuronal correlates of subjective visual perception"  
Science 245:761-763

  "Frontal Lobe Syndromes" In Vinke & Bruyn (Eds) "Handbook of Clinical  
Neurology Vol 2"  
North Holland Publishing Company:Amsterdam

- MacCurdy, Edward (Translator) (1938)  
  "The Notebooks of Leonardo Da Vinci"  
Jonathan Cape

- Maffei, L. and Fiorentini, A. (1977)  
  "Spatial frequency rows in the striate visual cortex"  
Vision Research 17 pp 257-264.

  "Improving Computer Systems for Fashion Designers"  
How Sketches Work: Bibliography

-Mandler, G. (1975)

-Mandler, G. (1985)


-Marcel A.,J. (1983a)
"Conscious and unconscious perception: Experiments on visual masking and word recognition" Cognitive Psychology, 15, pp 197-237

-Marcel A.,J. (1983b)
"Conscious and unconscious perception: An approach to the relations between phenomenal experience and perceptual processes" Cognitive Psychology, 15, pp 238-300


-Marks, D.F. (1972)

-Marks, D. F. (1983)

"Mental rotation by the blind: Does mental rotation depend upon visual imagery?" Journal of Experimental Psychology: Human Perception and Performance, 2, pp 515-521.

-Marqusee, Michael (1977)
Introduction to Cozens, Alexander, (1785) Paddington Press Ltd

-Marr, David (1982)
"Vision A Computational Approach" W.H.Freeman & Co.

"Theory of Edge Detection"


- Martin, Kenneth (1975)
  Kenneth Martin. Tate Gallery Catalogue and collected essays. (Exhibition 14th May - 29 June. 1975)
  Tate Gallery. London.

- Martin, M. (1979)
  Local and Global Processing: The role of sparsity
  Memory and Cognition, 7, pp 476-484.

  "Cognition"

  "Spatial frequency tuned channels: Implications for structure and function from psychophysical and computational studies of stereopsis" Philosophical Transactions of the Royal Society of London, Series B, 1980, 290, pp 95 -116 (b)

- McLuhan, Marshal (1969)
  "Understanding Media"
  Sphere Books Ltd : London.

- Mesesnel, Janez (1981)
  "France Slana. Water colors, Paintings, and Drawings 1944-1980
  Control Data Arts: Minneapolis. U.S.A.

- Michie, Donald and Johnston, Rory (1984)
  "The Creative Computer" (Chapter 4)

  "1993 Apparel User Conference" Brochure

  "Imagery in scientific thought: Creating 20th century physics." Birkhauser: Boston.

- Miller, G.A. (1962)
  "Psychology: The Science of Mental Life"
"The Visual Brain in Action"  
Oxford University Press

Milner, A.D. and Rugg, M.D. (Eds)(1992)  
"The Neuropsychology of Consciousness"  

Minsky, M. (1975)  

Minsky, M. (1985)  
"The Society of Mind"  
New York: Simon and Schuster.

"Object vision and Spatial Vision: two Cortical pathways."  
Trends in Neurosciences. 6, pp 414-417

Morris, D. (1967)  
"The Naked Ape"  
Jonathan Cape: London.

Morton, J. (1969)  
"Interaction of information in word recognition"  
Psychological Review, 76, pp 165-178

Morton, J. (1979)  

Moyer, R.S. and Bayer,R.H. (1976)  
"Mental comparison and the symbolic distance effect"  
Cognitive Psychology, 8, pp 228-246

Muir, Jean (1984)  
Exhibition Catalogue; Leeds Art Galleries, Leeds.

"What is it like to be a Bat ?"  
Philosophical Review, 83, pp 435-450

Navon, D. (1977)  
"Forest before trees:the Precedence of Global Features in Visual Perception",  
Cognitive Psychology, 9, pp 353-83

Neely, J.H. (1977)  
"Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited capacity attention"  
Journal of Experimental Psychology: General, 106, pp 226-254

Negroponte, N. (1977)  
- Neisser, Ulric (1967)  

- Neisser, U. (1976)  
  "Cognition and Reality"  
  Freeman : San Francisco.

- Neisser, U. and Becklen, R. (1975)  
  "Selective looking: Attending to specified events"  
  Cognitive Psychology, 7, pp 480-494.

  "Physical Symbol Systems"  
  Cognitive Science 4:2: pp 135-183

  "Human Problem Solving"  
  Prentice-Hall; Englewood Cliffs

  "Memory and Knowledge" In Norman, D. A. & Rumelhart, D. E. (Eds)  
  "Explorations in Cognition"  

  "Attention to action: Willed and automatic control of behaviour" (Technical report no 99) San Diego Center for Human Information Processing,University of California.

- Noton, D. and Stark, L. (1971)  

- Oatley, Keith (1988)  
  "On changing one's mind: a possible function of consciousness"  
  In "Consciousness in Contemporary Science"  

- Ogden, J.A. (1985)  
  "Contralesional neglect of constructed visual images in right and left brain-damaged patients"  
  Neuropsychologia 23 pp 273-277.

- Pach, Walter (translator)(1937)  
  "Journal of Eugene Delacroix"  
  New York.

- Paivio, A. (1971a)  

- Paivio, A. (1971b)  
  "Imagery and Verbal Processes"  
- Paivio, A. (1975)  
  "Perceptual comparisons through the mind's eye"  
  Memory and Cognition, 3, pp 635-647

- Paivio, A. (1983)  
  "The mind's eye in Arts and Science"  
  Poetics, 12, pp 1-18.

- Paivio, A. (1986)  
  "Mental Representations: A Dual Coding Approach"  
  (Oxford Psychology Series No 9)  

- Palmer, S. (1975a)  
  "The effects of contextual scenes on the identification of objects"  
  Memory and Cognition, 3, pp 519-526.

- Palmer, S. (1975b)  
  "The nature of perceptual representation: An examination of the analog/propositional debate"  
  In Schank, R.C. & Nash-Webber, B.L. (Eds.) "Theoretical issues in natural language processing" Tinlap Press: Arlington

- Palmer, S. (1975c)  

- Palmer, S. (1977)  
  "Hierarchical structure in perceptual representation"  
  Cognitive Psychology, 9, pp 441-474

- Palmer, S. (1978)  
  "Fundamental Aspects of Cognitive Representation"  

  "Canonical perspective and the perception of objects"  

  "A Functionalist Approach to Mental Imagery"  
  Paper presented at the twenty-third meeting of the Psychonomic Society. Baltimore M.D. (Also described in Baars 1988)

- Panum, P.L. (1940)  
  "Physiological Investigations concerning Vision with Two Eyes" (C. Hubscher, Trans.)  
  Hanover, Dartmouth Eye Institute (Original 1858)

  "Use of Implicit motor imagery for Visual shape discrimination as revealed by PET"  
``Peirce, C.S. (1960)
"Collected Papers of Charles Sanders Peirce"

"The brain's record of auditory and visual experience"
Brain, 86, pp 595-697

"Penrose, Roland (1962)
"Picasso: His Life and Work"

"Perky, C.W. (1910)
"An experimental study of imagination"
American Journal of Psychology 21, pp 422-425

& Jeeves, M.A. (1985)
"Visual cells in the temporal cortex sensitive to face view and gaze direction",

Thomas, S. (1991)
"Viewer-Centred and Object Centred encoding of Heads by cells in the
superior temporal sulcus of the rhesus monkey"
Experimental Brain Research, 86, pp 159-173.

"Issues of representation in Object Vision"
In Farah, M.J. & Ratcliffe, G. (Eds.) "The Neuropsychology of High-Level
Vision" Lawrence Erlbaum Associates.

"Perruchot, Henri (1961)
"Cezanne" (Translated by Humphrey Hare)
Perpetua Books.

"Neuronal responses to illusory contour stimuli reveal stages of Visual
processing" In J.D. Pettigrew, K.J. Sanderson & W.R. Levick (Eds) "Visual
Neuroscience" Cambridge University Press pp 343-351.

"Positron emission tomographic studies of the processing of single words"
Journal of Cognitive Neuroscience, 1, pp 153-170

"Visual detection and visual imagery"
Journal of Experimental Psychology 103, pp 509-514

"Components of visual memory"
Quarterly Journal of Experimental Psychology, 29, pp 117-133

"Interference with visualization"
Quarterly Journal of Experimental Psychology, 29, pp 637-650
"Mental Imagery and the Third Dimension"

Pinker, S. (1983)
"Perspective information in visualised scenes"
Presented at the Annual Meeting of the Psychonomic Society, San Diego, Nov pp 17-19


"Emergent Two-Dimensional Patterns in Images Rotated in Depth" Journal of Experimental Psychology: Human Perception and Performance Vol 6, No 2, pp 244-264

"Theories of Mental Imagery".
In Sheik, A. A. (ed) "Imagery: Theory, Research, and Application" New York, Wiley.

"Visualising objects at pre-specified orientations"
Presented at the Annual Meeting of the Psychonomic Society, San Antonio

Pitts, W. & McCulloch W.S. (1947)
"How we know universals: The Perception of Auditory and Visual forms"
Bulletin of Mathematical Biophysics, 9, pp 127-147

Podgorny, P. and Shepard, R.N. (1978)

"Perceptual Organization: An Overview"
In In M.Kubovy & J.Pomerantz (Eds), "Perceptual Organization": Lawrence Erlbaum.

Pope, K.S. and Singer, J.L. (Eds) (1978)
"The Stream of Consciousness"
New York: Plenum Press.

Popper, K.R. (1959)
"The Logic of Scientific Discovery"
Hutchinson: London.

Porter, P.B. (1954)
"The Hidden Man"
How Sketches Work: Bibliography

- Posner, M.I. (1980a)  
  "Mental Chronometry and the Problem of Consciousness"  
  In Jusczyk, P.W and Klein, R.M. (Eds) 1980  
  "The Nature of Thought. Essays in Honor of D.O. Hebb"  
  Erlbaum Associates, New Jersey

- Posner, M.I. (1980b)  
  "Orienting of attention"  
  Quarterly Journal of Experimental Psychology. 32, pp 3-25

  "Components of Attention"  
  Psychological Review, 78, pp 391-408

  "Attention and Cognitive control"  
  In Solso (Ed) "Information Processing & Cognition" Erlbaum.

  "Attentional mechanisms and conscious experience"  
  In A.D.Milner & M.D. Rugg (Eds) The Neuropsychology of Consciousness"  

  "Neural systems control of spatial orienting"  

  "Visual Attention" In Farah, M.J. & Ratcliff, G. (Eds)  
  "The Neuropsychology of High-Level Vision"  
  Erlbaum: Hillsdale, New Jersey.

- Pratt, Francis (1984)  
  "A theoretical framework for thinking about depiction"  
  In Crozier, W.R. & Chapman A. J. (Eds) (19??)  
  "Cognitive Processes in the Perception of Art"  
  (Advances in Psychology 19). North Holland.

- Pylyshyn, Z.W. (1973)  

- Pylyshyn, Z.W. (1975)  
  "Do we need images and analogues ?"  

  "The Imagery debate: Analogue Media versus Tacit Knowledge"  
  Psychological Review, vol 88, No 1, pp 16-45

  "Computation and Cognition"  
"Movement and attention in visual working memory"
The Quarterly Journal of Experimental Psychology, 38A, pp 689-703

Rawson, Philip (1969)
The appreciation of the Arts/3 "Drawing"

Read, H. (1959)
"A Concise History of Modern Painting"
Thames and Hudson: London.

Reason, J.T. (1979)
"Actions not as Planned"

Reed, S.K. (1972)
"Pattern recognition and categorization"
Cognitive Psychology, 3, pp 383-336

Reed, S.K., & Johnsen, J.A. (1975)
"Detection of parts in patterns and images"
Memory and Cognition, 3, pp 569-575

"Tacit knowledge and the effect of pattern configuration on mental scanning"
Memory and Cognition, 11, pp 137-143

Reffin-Smith, Brian (1984)
"Soft Computing, Art and Design"
Addison-Wesley Publishing Company

"Neither Pictures nor propositions: What can we learn from a mental Image?"

Rewald, J. (1991)
"Camille Pissarro"
Thames and Hudson


Richards, R. J. (1987)
"Darwin and the emergence of Evolutionary Theories of Mind and Behaviour"

"Working Memory and Human Cognition"

Robson, J.G. (1975)
How Sketches Work: Bibliography

"The Logic of Perception"  

-Rock, I., and DiVita, J. (1987)  
"A Case of Viewer-Centered Object Perception"  
Cognitive Psychology, 19, pp 280-293

"The effect of inattention on form perception"  

"Localization of cortical areas activated by thinking"  

"Mathematical Elements for Computer Graphics"  
McGraw-Hill Book Co

-Rorschach, H. (1942)  
"Psychodiagnostics, a Diagnostic test Based on Perception"  
(Translated by Lemkau, P. and Kroneneberg B.)  
Grune and Statton Inc. New York.

-Rosch, Eleanor and Lloyd, Barbara (eds) (1978)  
"Cognition and Categorization"  
Erlbaum, Hillsdale, New York.

"Two Concepts of Consciousness"  
Philosophical Studies, 49, pp 329-359

-Roskill, Mark (selected and edited.) (1963)  
"The letters of Vincent Van Gogh"  
Collins. The Fontana Library.

-Ross, J. (1976)  
"The Resources of Binocular Perception"  
Scientific American, 234, pp 80-6

Why are "what" and "where" processed by separate cortical visual systems? A computational investigation  
Journal of Cognitive Neuroscience 1 pp 171-186

"Parallel Distributed Processing. Explorations in the Microstructure of Cognition"  

"Representation of knowledge"  
"Issues in Cognitive Modelling"  
Erlbaum, New York.
"The representation of knowledge in memory" In Anderson, R.C. (Ed)
"Schooling and the Acquisition of Knowledge"
Erlbaum: Hillsdale, New Jersey.

Ruskin, John (1907)
"The Elements of Drawing. The Elements of Perspective"
J.M. Dent and Sons Ltd. London.

Russell, Bertrand (1903)
"The Principles of Mathematics" (Second edition 1937)

Ryle, G. (1949)
"The Concept of Mind"
Hutchinson (1949) Penguin University Books (1973)

Saint Laurent, Yves (1987)
Yves Saint Laurent Retrospective (Catalogue)
Art Gallery of New South Wales. Sydney, Australia.

Sagi, D. & Julesz, B. (1985)
"Where" and "what" in vision"
Science, 228, pp 1217-1219

"Disruption of short-term memory by unattended speech: Implications for the
structure of working memory"
Journal of Verbal Learning and Verbal Behaviour, 21, 150-164.

"Effects of background music on phonological short-term memory"

Samuels, M. and Samuels, N. (1975)
"Seeing with the mind's eye"
New York: Random House. 1975

Sanders, A. (1963)
"The selective process in the functional visual field"
Institute for Perception, RVO-TNO, Soesterburg, The Netherlands.

Selective attention in the speeded classification and comparison of
multidimensional stimuli

"Consciousness and Awareness in Memory and Amnesia" In Milner, A.D. &
Rugg, M.D. (Eds) "The Neuropsychology of Consciousness" Academic
Press.

Schank, R.C. & Nash-Webber, B.,L. (Eds.) (1975)
"Theoretical issues in natural language processing"
Tinlap Press: Arlington
- Scharf, Aaron (1968)  
  "Art and Photography"  

- Schiller, P.H. (1985)  
  "A model for the generation of visually guided saccadic eye movements"  
  In "Models of the Visual Cortex" Edited by D. Rose and V.G. Dobson  
  John Wiley and Sons Ltd. 1985.

- Schneider, W. & Shiffrin, R.M. (1977)  
  "Controlled and automatic information processing I: Detection, search and attention"  
  Psychological Review, 84, pp 1-66.

- Schroeder, R. (1976)  
  "Information processing of Color and Form"  
  Unpublished Honors Thesis University of Oregon  
  (Described in Posner, M. 1986)

  "How Interaction with Sketches Aids Creative Design"  
  LUTCHI Research Centre. Loughborough University.

  "Pictorial Properties in Raster Graphics: Classification and Use"  

- Segal, S.J. (1971)  
  "Processing of the stimulus in imagery and perception"  
  In Segal (ed) "Imagery: Current Cognitive Approaches"  

- Segal, S.J. (1972)  
  Assimilation of a stimulus in the construction of an image: The Perky effect revisited"  
  In Sheehan, P. W. (Ed) 1972 "The function and nature of imagery"  

- Segal, S.J. and Fusella, V. (1970)  
  "Influence of imaged pictures and sounds on detection of visual and auditory signals",  
  Journal of Experimental Psychology, 83, pp 458-464

- Segal, S.J. and Nathan, S. (1964)  
  "The Perky effect: Incorporation of an external stimulus into an imagery experience under placebo and control conditions"  
  Perceptual Motor Skills, 1964, 18, pp 385-395

- Selfridge, O.G. and Neisser, U. (1960)  
  "Pattern Recognition by Machine"  
  Scient. Am., 203, pp 60-68

- Shallice, T. (1972)  
  "Dual functions of consciousness"  
  Psychological Review, 79, pp 383-393
How Sketches Work: Bibliography

"Shallice, T. (1978)

"Specific Impairments of Planning" Philosophical Transactions of the Royal Society B, 298, pp 199-209


"Shepard, R. N. (1964)
"Attention and the Metric Structure of the stimulus space" Journal of Mathematical Psychology, 1964, 1, pp 54-87

"Shepard, R.N. (1975)
"Form, formation and transformation of internal representations" In R.Solso, Ed. "Information processing and cognition: The Loyola Symposium" Hillsdale, N.J. : Lawrence Erlbaum Associates

"Shepard, R. N. (1978)

"Psychophysical complementarity" in Kubovy and Pomerantz (Eds) cited


"Ecological constraints on internal representation: Resonant kinematics of perceiving, imagining, thinking and dreaming" Psychological review, Vol 91, No 4, pp 417-447


"Shepard, R.N. and Feng, C. (1972)
"A chronometric study of mental paper folding" Cognitive Psychology, 3, pp 228-243
Shepard, R.N. and Metzler, J (1971)
"Mental rotation of three-dimensional objects"
Science, 171, pp 701-703

"Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory."
Psychological Review, 84, pp 127-190.

Shulman, G.L. & Wilson, J. (1987a)
"Spatial frequency and selective attention to local and global information"

Shulman, G.L. & Wilson, J. (1987b)
"Spatial frequency and spatial attention"
Perception, 16, pp 103-111.

Simon, H.A. (1972)
"What is visual imagery? An information-processing interpretation." In L.W.Gregg (Ed) "Cognition in Learning and Memory" New York Wiley.

Skemp, R.R. (1971)
"The Psychology of Learning Mathematics"
Harmondsworth: Penguin.

Sloman, Aaron (1975)
"Afterthoughts on analogical representation" In Schank,R.C. & Nash-Webber, B.,L. (Eds.) "Theoretical issues in natural language processing"
Tinlap Press: Arlington U.S.

Smets, G. (1992)
"The Theory of Direct Perception and CAD in Virtual Reality" SERC sponsored symposium "Virtual Representations for Design and Manufacture"
Coventry University.(16th December)

Spies, Werner (1968)
"Max Ernst Frottages" (revised edition 1986)
Thames and Hudson.

Standing, L.G., Conezio,J. & Haber, N. (1970)
"Perception and memory for pictures: Single-trial learning of 2500 visual stimuli"
Psychonomic Science, 19, pp 73-74

Sternberg, S. (1966)
"High Speed Scanning in Human Memory"
Science, 153, pp 652-54

Sternberg, R.J.(Ed.) (1988)
"The Nature of Creativity"
Cambridge University Press.

"Spatial Cognition. Brain Bases and Development"
"Stoichita, V.I. (1979)  
"Mondrian"
Abbey Library (Murrays) :London.

"Sussman, G.J. (1975)  
"A Computational model of skill acquisition"

"Sutherland, N.S. (1968)  
"Outline of a theory of visual pattern recognition in animals and man"
Proceedings of the Royal Society, B. 171, pp 297-317

"Swedish Standards Institution (1978)  
"Colour Atlas" NCS (Natural Color System)
SIS, Stockholm, Sweden

"Mental rotation and Orientation-Dependence in Shape Recognition"

"Thouless, R. (1931)  
"Phenomenal Regression towards the Real object"
British Journal of Psychology, 21, pp 339-359

"Tipper, S.P. (1985)  
"The negative priming effect: Inhibitory priming by ignored objects"

"Tipper, S.P. and Cranston, M. (1985)  
"Selective attention and Priming: Inhibitory and Facilitatory effects of ignored primes"
Quarterly Journal of Experimental Psychology, 37A, pp 591-611

Negative priming between pictures and words in a selective attention task:  
Evidence for semantic processing of ignored stimuli"
Memory and Cognition, 16, pp 64-70.

"An attempt to clarify some (vague or ambiguous) notions of fuzzy set theory"
In Trappl, R. (Ed.) "Cybernetics and Systems '94"  
World Scientific

"Treisman, A.M. (1986)  
"Properties, parts and objects"

"A feature-integration theory of attention"
Cognitive Psychology, 12, pp 97-136

"Emergent features, attention and object perception."
Journal of Experimental Psychology: Human Perception and Performance, 10, pp 12-31
"Illusory conjunctions in the perception of objects"  
Cognitive Psychology, 14, pp 107-141

Tsal, Y. (1983)  
"Movements of Attention across the Visual Field"  
Journal of Experimental Psychology: Human Perception and Performance 9, pp 523-530

Tulving, Endel (1983)  
"Elements of Episodic Memory"  

Tversky, A. (1977)  
"Features of Similarity"  
Psychological Review, 84, pp 327-352

"Visual Routines"  
Cognition, 18, pp 97-159

"Aligning pictorial descriptions: An approach to object recognition"  

"The control operations of consciousness"  
In "Consciousness in Contemporary Science"  

"Aspects of Consciousness"  
Vol 3 "Awareness and Self Awareness"  

Underwood, G. and Stevens, R. (Eds) (1979)  
"Aspects of Consciousness"  
Vol 1 "Psychological Issues"  

"Aspects of Consciousness"  
Vol 2 "Structural Issues"  

Ungerleider, L.G. (1985)  
"The corticocortical pathways for object recognition and spatial perception"  
In C. Chagas (Ed.) "Pattern Recognition Mechanisms". Pontificae Academiae Scientiarum Scripta Varia

Valentine, T. & Bruce, V. (1986)  
"Recognising familiar faces: The role of distinctiveness and familiarity"  
Canadian Journal of Psychology, 40, pp 300-305
Van Essen, D.C. (1985)  
"Functional organisation of primate visual cortex"  

"Hierarchical organisation and functional streams in visual cortex"  
Trends in Neurosciences, 6, 370-375

Vasari (Ed by G. Milanesi) (1878)  
"The Lives of the Painters, Sculptors and Architects"  
Florence 1878

Vernon, M.D. (1952)  

Vernon, M.D. (1970)  
"Perception through Experience" Methuen and Co Ltd. Britain

Vernon, P.E. (1970)  
"Creativity"  

Vygotsky, L.S. (1978)  
"Tool and symbol in children's development"  

Walker, P. (1975)  
"The subliminal perception of movement and the "suppression" in binocular rivalry"  
British Journal of Psychology, 66, pp 347-356

"Apparent equivalence between perception and imagery in the production of various visual illusions"  
Memory and Cognition, 12, pp 156-162.

"Consciousness and Behaviour"  
Boston, MA: Allyn and Bacon.

"Implicit Perception in Visual Neglect: Implications for Theories of Attention"  

Waltz, D. (1979)  
"On the function of mental imagery"  
Behav. Br. Sci., 2, pp 569-570.

"Determinants of attention to local and global features of visual forms."  
"Chromatic Imaging: An effect of mental colouring on recognition memory"
Canadian Journal of Psychology 36 pp 291-299

Watson, James (1968)
"The Double Helix"
Signet: New York

Watt, Roger (1988)
"Visual Processing: Computational, Psychophysical and Cognitive Research"
Lawrence Erlbaum Associates, Hillsdale, New Jersey

Weiskrantz, L. (1986)
"Blindsight: A case Study and Implications"
Oxford: Oxford University Press

Weiskrantz, L. (1988)
"Some contributions of neuropsychology of vision and memory to the problem of consciousness" In "Consciousness in Contemporary Science"
Oxford Science Publications.

Welch, Robert (1986)
"Hand and Machine"
Robert Welch, The Mill, Chipping Campden, GL55 6DU.

Werner, Arnold (1987)
"Dufy" Thames and Hudson, London.

Wertheimer, M. (1923,1938)
"Laws of Organisation in Perceptual Forms"

White, J. (1987)
"The Birth and Rebirth of Pictorial Space"
Faber and Faber.

"Visual Thinking in the Ice Age"
Scientific American, July 1989, pp 74-81

"Perception: From sense to object"
Hutchinson: London

Wilenski, R. H. (1944)
"Modern French Painters"
Faber and Faber.

Wilenski, R. H. (1945)
"The Modern Movement in Art" (Second Edition)
Faber and Faber.

Wilkinson, Gerald (1975)
"Turner's Colour Sketches 1820 - 1834" Barrie & Jenkins.
Willats, John (1977)

"Getting the drawing to look right as well as be right: The interaction between production and perception as a mechanism of development"

Willats, John (1985)
"Drawing systems revisited: the role of denotation systems in children's figure drawings"

Windels, F. (1949)
"The Lascaux Cave Paintings"
Faber and Faber : London.

Wollheim, R. (1977)
"Representation: The philosophical contribution to psychology"

Yarbus, A.L. (1967)
"Eye movements and vision"
New York: Plenum.

"The Art of Memory"
London: Routledge and Kegan Paul

Young, A.W. (1994)

"A Vision of the Brain"

"Imagery in the congenitally blind: How visual are visual images?" Journal of Experimental Psychology: Learning

Zusne, L. (1970)
"Visual Perception of Form" Academic press.
INDEX OF NAMES

Abbate, F. 30
Adams, A.J. 290
Albers, Joseph 33
Alberti, Leon 60
Allport, A. 139, 187, 210, 211, 216, 232
Anderson, B. 168, 169
Anderson, J. 83, 86
Anderson, R. 143
Arditi, Aries 165
Arnheim, Rudolf 21, 23, 28, 32, 41, 70, 81, 101, 102, 116, 117, 123, 207, 241, 243, 248
Auerbach, Frank 57
Austin, A.L. 168
Ayrton, Michael 23, 40
Baars, B. 19, 21, 209, 219, 221, 226, 242, 244, 252, 256, 257, 290
Bacou, R. 9
Barlow, H.B. 262
Bartlett, F.C. 74, 120
Bates, W. 29, 33, 34, 143, 205
Beck, J. 59, 97, 194
Becklen, R. 168
Beech, J.R. 139
Biederman, I. 10, 17, 21, 113, 114, 117-118, 121, 123, 140
Bisiach, E. 132, 147, 148, 209, 215, 216, 222, 223
Black, M. 16, 47, 48, 65, 66
Blake, R. 169
Boden, Margaret 19, 21, 223, 236, 237, 240, 243, 281
Boies, S.J. 188
Botticelli 53
Braitenburg, V. 155, 157
Brandimonte, M. 137, 138
Brandt, S.A. 189
Breitmeyer, B.G. 188
Brennan, S.E. 120
Bronowski, J. 81
Brooks, L.R. 76, 148, 149, 158, 159, 161, 162, 292
Bruce, V. 97, 106, 120, 136, 187
Brunelleschi, Filippo 59, 60
Bruner, J.S. 21, 68, 72, 124, 172, 236, 266
Burr, D.C. 98
Carmichael, L. 108, 172
Carpenter, P.A. 157
Cave, K.R. 126, 145, 147
Cezanne, Paul 9, 32, 33, 34, 41, 193, 199, 238
Chambers, D. 10, 18, 137, 138
Chase, W.G. 116, 153
Cheeseman 211
Chevreul, M. E. 59
Chipman, S. 133
Chomsky, N. 79, 87
Christie, D.F.M. 141
Church, R.M. & K.W. 135
Churchland, P.M. 209
Clark, Kenneth 53
Clark, S. 37, 93, 215, 246, 253, 290
Claude 53
Clouzot, Henri 43, 245, 291
Cogniat, Raymond 43
Cohen, Harold 10, 21, 30, 33, 281, 282
Coleman E. A. 8, 245
Cowey, A. 84, 171, 193
Cozens, Alexander 8, 53, 54, 55, 65, 145, 243
Craiik, F. 40, 41
Craik, Kenneth 87
Crawford, Alan 43, 210
Crick, Francis 83, 137, 144, 216, 217, 218, 227
Critchlow, Keith 30
D'Honnecourt, Villard 8, 49
Daguerre, L. 58
Davidson, J.M. & R.J. 209
De Valois, R.L. 119
Delacroix, Eugène 53, 60, 92, 278
Denis, M. 253
Dennett, Daniel 136, 209, 216
Deregowski, J.B. 70
Diehl, G. 64
DiVita, J. 115
Dixon 185
Donne, John 207
Downing, C.J. 187
Driver, J. 191, 230
Duensing, S. 169
Dufy, Raoul 43, 195
Duncan, J. 183, 187, 190, 191, 193
Eco, Umberto 68
Edelman, G. M. 157, 216, 217, 222
Edmonds, E.A 12, 48
Egeth, H. E. 184, 185, 196
Einstein, Albert 144
Eisenberg, P. 157
Ellis, A. 156
Ellis, W. 28, 101
Engel, E. 168, 169
Engel, F. 202, 203
Ericsson, K.A. 212
Ernst, Max 4, 8, 13, 16, 21, 54, 105, 145, 248
Evans, J. 92
Evans, R. 194
Eysenck, Michael 207
Faraday, Michael 144
Felfoldy, G.L. 184
Feng, C. 142
Fiorentini, A. 84
Fish, J.C. 14, 17, 82, 92, 162, 182, 194
Fisher, L. E. 209
Fodor, J. 221
Forster, E.M. 244
Fox, P.T. 146, 164
Franks, J.J. 120
Freyd, J.J. 145, 151
Frisby, J. P. 76, 89, 165, 179
Fusella, V. 149, 166
Galton, Francis 129, 144
Ganz, L. 188
Garner, W.R. 18, 183, 184, 185, 196
Gassier, Pierre 9
Gazzaniga, M. S 215
Georges, M. 119
Gerbino, W. 125, 137, 138
Gerhardstein, P. 114
Ghiselin, B. 240, 245
Gibson, J. J. 17, 72, 98, 99, 100, 105, 106, 173, 181, 270, 274
Ginsburg, A. P. 9, 199
Giotto 49
Glimcher, A. & M. 8, 41, 42, 102
Godwin, William 21
Goldenburg, G. 146
Goldscheider, Ludwig 9
Goldschmidt, Gabriela 21, 82, 92, 212, 246, 253
Goldwater, R. 52, 92
Goodale, M.A. 106
Goodman, N. 69, 74, 80
Gordon, I.E. 97, 265
Gottwald, R.L. 184
Goya, Francisco 9, 53, 172, 174, 205
Graham, N. 200, 204
Graham, S. 149, 166
Gray, Stephen 60, 172, 270, 273
Green, P.R. 97
Greenburg, D.P. 10, 273
Grossi, D. 148
Gutman, D. 190
Haber, R.N. 200
Hagen, M.A. 99, 100, 270
Haldane, J.B.S 64
Halligan, P.W. 147, 216
Hamilton, Richard 62, 278, 292
Harnad, S. 209
Hauser, Arnold 30
Hawkins, H.L. 184, 196
Hayes, J.R. 151 - 153
Hayes-Roth, F. 83
Hebb, D.O. 85
Helmholtz, H. von 4, 59, 102, 103, 104, 144, 214
Helstrup, T 143
Henderson, A. 33
Hering, Ewald 59, 72, 182, 194
Hildreth, E.R. 76, 200
Hilgard, E.R. 211, 220, 227
Hinton, G. E. 86, 132
Hirst, William 83
Hitch, G. 2, 20, 130, 160, 161, 218, 223, 224, 233, 253, 258, 259
Hobbes, Thomas 74
Hochberg, J. 76, 98, 103, 104, 106, 108, 133
Hockett, C.D. 79
Hockney, David 292
Index of Names

Hodgkin, Howard 292
Hoffman, D. D. 17, 112, 113, 123, 140
Hoffman, J.E. 187, 203
Hofstadter, D.R. 6, 19, 236, 237 - 239, 242, 243, 244, 265
Holding, D.H. 153
Holyoake, K.J. 284
Honey, W.B. 279
Hopgood, F.R.A. 82, 199
Hubel, D.H. 84, 106, 155, 171, 188, 197
Hummel, J.E. 113
Humphrey, N. 215
Humphreys, G. 21, 89, 106, 110, 111, 136, 187, 190
Hunt, R.W.G. 194, 284
Huttenlocher, D.P. 118
Imber, L.G. 244, 256
Ingres, Jean-Auguste 9, 92, 269, 271
Jackendoff, R. 81, 209, 260
James, Charles 8, 91, 245
James, William 178, 209, 214, 226, 232
Johnsen, J.A. 33
Johnson-Laird, P.N. 19, 68, 81, 83, 87, 90 - 92, 209, 210, 211, 212, 221 - 224, 226, 241, 242
Jolicoeur, P. 83, 89, 90, 112, 115, 123, 135, 259
Jonides, J. 133, 142, 158, 188
Judd, D. B. 34
Julesz, B. 165, 197, 209, 217, 218
Jusczyk, P.W. 85
Just, M.A. 185
Kahn, Gustave 9
Kahneman, D. 183, 187, 190, 191, 192, 230
Kandinsky 236, 242, 243
Kanizsa, G. 124, 125, 126
Karmiloff-Smith, A. 266
Katz, David 194
Kaufman, L. 97, 108
Keats, John 96
Keenan, J.M. 158
Kekulé, F. 144, 240
Kelly, M.H. 145
Kerr, N.H. 157
Kinchla, R. A. 201
Kline, M. 81
Knudsen, E. I. 171
Koch, C. 216
Koenig, O 83, 84, 106, 130, 139, 163, 164, 187, 213, 299
Koffka, K. 28, 100, 101
<table>
<thead>
<tr>
<th>Name</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kohler, K.</td>
<td>28, 100, 101</td>
</tr>
<tr>
<td>Kolers, P.A.</td>
<td>168</td>
</tr>
<tr>
<td>Kotov, A.</td>
<td>154</td>
</tr>
<tr>
<td>Kramer, P.</td>
<td>200</td>
</tr>
<tr>
<td>Kubovy, M.</td>
<td>46, 123</td>
</tr>
<tr>
<td>Kulikowski, J.J.</td>
<td>170</td>
</tr>
<tr>
<td>Kunen, S.</td>
<td>33</td>
</tr>
<tr>
<td>Kurtzman, H.S.</td>
<td>33, 132</td>
</tr>
<tr>
<td>Lagerfeld, Carl</td>
<td></td>
</tr>
<tr>
<td>Lamb, M. R.</td>
<td>202</td>
</tr>
<tr>
<td>Langer, E.J</td>
<td>244, 256</td>
</tr>
<tr>
<td>Larsen, A.</td>
<td>114, 115, 136, 138, 185, 202</td>
</tr>
<tr>
<td>Lawson, Brian</td>
<td>32, 207</td>
</tr>
<tr>
<td>Le Grand</td>
<td>284</td>
</tr>
<tr>
<td>Le Grand, Y.</td>
<td>284</td>
</tr>
<tr>
<td>Leakey, R.</td>
<td>175</td>
</tr>
<tr>
<td>LeDoux, J., E.</td>
<td>83</td>
</tr>
<tr>
<td>Leeuwenberg, E.L.J.</td>
<td>82</td>
</tr>
<tr>
<td>Leonardo da Vinci</td>
<td>8, 13, 16, 17, 27, 29, 46, 47, 49, 50-56, 58, 65, 66, 96, 109, 121, 145, 151, 248, 257</td>
</tr>
<tr>
<td>Leslie, A.D.</td>
<td>213</td>
</tr>
<tr>
<td>Levelt, W.J.M.</td>
<td>170</td>
</tr>
<tr>
<td>Levine, D. N.</td>
<td>157</td>
</tr>
<tr>
<td>Lewis, V.J.</td>
<td>259</td>
</tr>
<tr>
<td>Lieberman, K.</td>
<td>158, 292</td>
</tr>
<tr>
<td>Lindauer, M.S.</td>
<td>33, 144</td>
</tr>
<tr>
<td>Livingstone, M. S.</td>
<td>84, 106, 155, 188, 193, 195, 197</td>
</tr>
<tr>
<td>Lockhart, R. S.</td>
<td>41</td>
</tr>
<tr>
<td>Logie, R. H.</td>
<td>20, 21, 126, 130, 137, 141, 154, 158 - 159, 161, 162, 218, 227, 230, 256, 258, 259, 270</td>
</tr>
<tr>
<td>Logothetis, N.K.</td>
<td>170</td>
</tr>
<tr>
<td>Luria, A.R.</td>
<td>234</td>
</tr>
<tr>
<td>Luzzatti, G.R.</td>
<td>132, 147</td>
</tr>
<tr>
<td>MacCurdy, E.</td>
<td>52</td>
</tr>
<tr>
<td>Macko, K.A.</td>
<td>155</td>
</tr>
<tr>
<td>Maffei, L.</td>
<td>84</td>
</tr>
<tr>
<td>Makirinne-Crofts, P.</td>
<td>21, 37, 48, 205, 207, 280, 293</td>
</tr>
<tr>
<td>Mandler, G.</td>
<td>212</td>
</tr>
<tr>
<td>Mandler, G.</td>
<td>208, 209, 212, 214, 244</td>
</tr>
<tr>
<td>Mandler, J.M.</td>
<td>145</td>
</tr>
<tr>
<td>Manet, Edouard</td>
<td>57, 61</td>
</tr>
<tr>
<td>Marcel A., J.</td>
<td>18, 19, 107, 108, 209, 212 - 215, 218, 222, 223, 228, 229</td>
</tr>
<tr>
<td>Marchetti, C.</td>
<td>130, 258</td>
</tr>
</tbody>
</table>
Marks, D.F. 33, 129, 209
Marmor, G.S. 157
Marqusee, Michael 54, 55
Marr, David 24, 76, 88, 89, 98, 99, 106, 110 - 112, 117, 125, 134, 169, 200, 294
Marshall, J.C. 147, 216
Martin, Kenneth 30
Martin, M. 202
Matisse, Henri 64, 124
Matlin, M. W. 207
Maunsell, J.H.R. 89
Maxwell, James Clark 144
May, J.G. 33
Mayhew, J.E.W. 165
McLuhan, Marshal 37
Michie, Donald 293, 299
Miller, A.I. 144
Miller, B. 169
Miller, G.A. 208, 211
Millet, Jean-Francois 9, 47, 63, 119, 152, 174
Milner, A.D. 106, 209, 212, 215
Minsky, M. 120, 226
Mishkin, M. 89, 155
Mondrian, Piet 64
Monet, Claude 57, 238
Moore, Henry 144
Morris, Desmond 214
Morton, J. 86, 87, 121
Moyer, R.S. 78
Mozart, Wolfgang 238
Muir, Jean 91, 205, 206, 280
Nagel, T. 209
Nakayama, K. 168, 169
Nash-Webber, B.L. 88
Nathan, S. 150
Navon, D. 136, 201, 202, 203
Neely, J.H. 229
Negroponte, N. 37, 282
Neisser, Ulric 118, 150, 164, 167, 168, 187, 190, 217
Newell, A. 92, 212
Niepce, Joseph-Nicéphore 58
Nishihara, H.K. 88, 89, 112
Noton, D. 252
Oatley, Keith 214, 215, 223
Ogden, J.A. 148
Ortony, A. 220
Index of Names

Pach, Walter 92
Paiio, A. 69, 81, 83, 85, 86 - 87, 90, 94, 121, 131, 158, 192, 251, 253
Palmer, S. 16 - 17, 21, 70, 71, 73, 76, 83, 86, 89, 111, 116, 119, 120 - 121, 123, 283
Palomino 56
Pani, J.R. 142, 167, 226
Panum, P.L 169
Parker, R.E 145
Parsons, L.M. 260, 261
Paterson, R. 209
Peirce, C.S. 4, 72, 73
Penfield, W. 225
Penrose, Roland 42, 291
Perky, C.W. 150, 166
Perot, P. 225
Perrett, D.I. 117
Perruchot, Henri 32
Peterhans, E. 147
Petersen, S.E. 189, 192
Peterson, M.J. 149, 166
Phillips, W. A. 141
Picasso, Pablo 8, 23, 32, 41 - 43, 51, 102, 153, 205, 238, 241, 245, 246, 291
Pinker, S. 83, 89, 116, 118, 123, 129, 130, 135, 136, 143, 187
Pissarro, Camille 62
Pitts, W. 111
Podgorny, P. 123, 134, 141, 146
Poggio, T. 169
Poincaré, H. 240
Pomerantz, J.R. 46
Pope, K.S. 209
Popper, K.R. 265
Porter, P.B. 10, 103, 126, 173
Posner, M.I. 21, 131, 135, 156, 157, 187, 188, 189, 203, 204, 209, 211, 216 - 218,
220, 224, 229, 255 - 256, 263
Postman, L. 172, 266
Poussin, Nicholas 53
Pratt, Francis 76, 180
Pylyshyn, Z.W. 83, 86, 87, 94, 132, 134, 135, 136
Quinlan, P. 89, 111
Quinn, J.G. 161
Raphael, 53
Ratcliff, G. 215
Rawson, Philip 19, 36, 55, 64, 194, 239, 244, 248, 249, 271
Read, H. 242
Reason, J.T. 214, 232, 256
Reed, S.K. 33, 120, 134, 135
Reffin-Smith, Brian 265
Reisberg, D. 10, 18, 137, 138
Rembrandt 53
Rewald, J. 62
Rhodes, G. 120
Richards, M. 112, 113, 140
Richards, R. J. 175
Robertson, L. C. 202
Robson, J.G. 199
Rock, I. 59, 98, 103, 115, 190
Rogers, D.F. 290
Roland, P.E. 146
Rorschach, H. 58, 105
Rosch, Eleanor 116
Rosenthal, D. 209
Roskill, Mark 64, 92
Ross, J. 168
Rothbart, M.K. 156, 188, 217, 218, 224, 255
Rousseau, Theodore 121
Rubens, Peter 53
Rueckl, J.G. 198
Rugg, M.D. 209
Rumelhart, D. E. 83, 106, 220, 238
Ruskin, John 195, 196
Russell, Bertrand 80
Ryle, G. 209
Sadaat, Sohrab 21
Sagi, D. 197
Salame, P. 160
Samuels, M. & N. 144
Sanders, A. 133, 164
Santee, J.L. 184, 185
Sargent, John Singer 61
Schacter, D.L. 216
Schall, J.D. 170
Schank, R.C. 87
Scharf, Aaron 16, 58, 59, 60, 61, 62, 278
Schiano, D.J 195, 285
Schiller, P.H. 98
Schmidt, H. 190
Schneider, W. 230, 231, 241
Schroeder, R. 196
Scrivener, S.A.R 12, 14, 16 - 17, 21, 37, 48, 82, 92 - 93, 215, 246, 253, 290
Segal, S.J. 149, 150, 166
Selfridge, O.G. 110
Seurat, George, 9, 200, 204
Shepard, R. N. 70, 86, 90, 99, 114, 131, 133 - 134, 138, 141 - 142, 144, 146, 164, 183, 185, 260
Shiffrin, R.M. 230, 231, 241
Shigley, R.H. 184, 196
Shin, L.M. 123, 129, 130, 131, 142, 161, 163, 170, 198, 252
Shulman, G.L. 136, 200, 203
Simon, H.A. 86, 92, 153, 212
Singer, J.L. 209
Skemp, R.R. 151
Slayton, K. 33, 132, 143
Sloman, Aaron 71, 73
Smets, G. 99, 105, 106, 270, 273
Snyder, R.R. 188, 229
Spies, Werner 8, 54, 55
St Laurent, Yves 8, 102
Standing, L.G. 255
Sternberg, R.J. 236, 240
Sternberg, S. 241, 242
Stevens, R. 209
Stiles-Davis, J. 165
Stoichita, V.I. 64
Streintz 60
Sung Ti 55
Surls, James 144
Sussman, G.J. 223, 227
Sutherland, N.S. 79, 109
Talbot, Fox 58
Tarr, M. J 83, 89, 116, 123
Tesla, Nicholas 144
Thouless, Robert 173
Tipper, S.P. 191
Titian 56, 57, 238
Toth, H. 47
Toulouse-Lautrec 10, 145, 268
Treisman, A.M. 21, 73, 136, 157, 163, 180 - 183, 187, 190 - 192, 209, 218, 220, 230, 259
Tsai, Y. 135, 187
Tulving, Endel 122, 245
Turing, Alan 221
Turner, William Mallord 28, 41, 43, 53, 63, 179, 195
Tversky, A. 73
Ullman, S. 83, 89, 90-91, 94, 118, 119, 154, 165, 235, 249
Umiltà, C. 222, 223, 226
Underwood, G. 209
Ungerleider, L.G. 89, 155
Valentine, T. 120
Van Essen, D.C. 11, 89, 155, 157
Van Gogh, Vincent 64
Vasari, Giorgio 49, 56, 60
Velasquez, Diego 56, 57
Vernon, M.D. 108, 172
Vernon, P.E. 240
Vygotsky, L.S. 214, 215
Walker, P. 169
Wallace, B. 132, 209
Wallace, M.A. 216
Waltz, D. 138
Ward, L.M. 202, 203
Watkins, M.J 195, 285
Watson, James 144, 242
Watt, Roger 106, 119, 165, 189, 200, 294
Weiskrantz, L. 215, 216
Welch, Robert 9, 21, 43 - 44, 47, 63, 91, 102, 210, 293
Werner, Arnold 195
Wertheimer, M. 100, 101
White, J. 60, 249
Wiesel, T.N. 84, 171
Wilding, J.M. 97
Wilenski, R. H. 57, 61, 197, 278
Wilkinson, Gerald 10, 63, 179
Willats, John 70, 76
Wilson, J. 136, 200, 203
Windels, F. 176
Wolfe, J.M. 201
Wollheim 69
Yarbus, A.L. 97, 261
Yates, F.A. 269
Young, A.W. 156, 216
Zeki, S. 106, 155 - 157, 163, 164
Zervos, Christian 245
Zimler, J. 158
Zusne, L. 59
Note: images removed for copyright purposes