

**TOWARDS A FURTHER UNDERSTANDING OF
OBJECT FEATURE BINDING:
A COGNITIVE NEUROSCIENCE PERSPECTIVE**

CAROLYN MAIR

**This thesis is submitted in partial fulfilment of the
requirements of Bournemouth University for the
Degree of Doctor of Philosophy**

January 2006

ABSTRACT

The aim of this thesis is to lead to a further understanding of the neural mechanisms underlying object feature binding in the human brain. The focus is on information processing and integration in the visual system and visual short-term memory. From a review of the literature it is clear that there are three major competing binding theories, however, none of these individually solves the binding problem satisfactorily. Thus the aim of this research is to conduct behavioural experimentation into object feature binding, paying particular attention to visual short-term memory.

The behavioural experiment was designed and conducted using a within-subjects delayed response task comprising a battery of sixty-four composite objects each with three features and four dimensions in each of three conditions (spatial, temporal and spatio-temporal). Findings from the experiment, which focus on spatial and temporal aspects of object feature binding and feature proximity on binding errors, support the spatial theories on object feature binding, in addition we propose that temporal theories and convergence, through hierarchical feature analysis, are also involved. Because spatial properties have a dedicated processing neural stream, and temporal properties rely on limited capacity memory systems, memories for sequential information would likely be more difficult to accurately recall. Our study supports other studies which suggest that both spatial and temporal coherence, to differing degrees, may be involved in object feature binding. Traditionally, these theories have purported to provide individual solutions, but this thesis proposes a novel unified theory of object feature binding in which hierarchical feature analysis, spatial attention and temporal synchrony each plays a role. It is further proposed that binding takes place in visual short-term memory through concerted and integrated information processing in distributed cortical areas. A cognitive model detailing this integrated proposal is given. Next, the cognitive model is used to inform the design and suggested implementation of a computational model which would be able to test the theory put forward in this thesis. In order to verify the model, future work is needed to implement the computational model. Thus it is argued that this doctoral thesis provides valuable experimental evidence concerning spatio-temporal aspects of the binding problem and as such is an additional building block in the quest for a solution to the object feature binding problem.

ACKNOWLEDGEMENTS

I would like to express my gratitude for sustained support and encouragement from my first supervisor, Professor Martin Shepperd, without whom this thesis would not have been completed. Sincere thanks to Dr. David Heathcote, my second supervisor, for discussions and guidance, and to Professor Alessandro Treves who kindly invited me to spend three months at the Sector for Cognitive Neuroscience, at SISSA, in Trieste, Italy, where I discovered how much there was to learn. Thanks also to my friends and my colleagues at Bournemouth University for their help and support. Finally, thanks to my children Jonathan, Louisa and Sophie for listening and caring.

TABLE OF CONTENTS

CHAPTER 1

BACKGROUND AND OVERVIEW OF THESIS

INTRODUCTION.....	1
1.1 MOTIVATION FOR STUDY.....	2
1.2 THESIS AIM AND OBJECTIVES.....	2
1.3 THESIS ORGANISATION	3
1.4 RESEARCH METHODOLOGY	5
1.5 RESEARCH SCOPE AND LIMITATIONS.....	6
<i>1.5.1 Focus of study</i>	<i>6</i>
<i>1.5.2 Application domain</i>	<i>7</i>
<i>1.5.3 Empirical investigation</i>	<i>7</i>
<i>1.5.4 General empirical limitations</i>	<i>8</i>
<i>1.5.5 Rationale for methods</i>	<i>8</i>
1.6 CONCLUSIONS	9
1.7 SUMMARY	9

CHAPTER 2

BIOLOGICAL AND ARTIFICIAL NEURAL NETWORKS

INTRODUCTION.....	11
2.1 BIOLOGICAL NEURONS.....	11
<i>2.1.1 Synapses, neurotransmitters and receptors</i>	<i>13</i>
<i>2.1.2 Spiking neurons.....</i>	<i>17</i>
<i>2.1.3 Biological neurons in networks.....</i>	<i>19</i>
<i>2.1.4 Biological neural systems (networks of networks).....</i>	<i>21</i>
2.2 ARTIFICIAL NEURAL NETWORKS	24
<i>2.2.1 Integrate and fire model neurons.....</i>	<i>28</i>
<i>2.2.2 Transfer functions</i>	<i>30</i>
<i>2.2.3 Coding information: rate or temporal spike coding?</i>	<i>31</i>

2.2.4	<i>Local, fully distributed or sparse models?</i>	32
2.2.5	<i>Learning rules</i>	33
2.2.6	<i>Learning in artificial neural networks</i>	35
2.2.7	<i>Unsupervised networks</i>	39
2.3	MODELS IN COMPUTATIONAL NEUROSCIENCE	43
2.3.1	<i>Pattern associators</i>	44
2.3.2	<i>Autoassociator (attractor) networks</i>	46
2.3.3	<i>Competitive learning networks</i>	51
2.3.4	<i>Continuous attractor neural networks</i>	55
2.3.5	<i>Modular networks</i>	57
2.4	MAJOR DIFFERENCES BETWEEN BIOLOGICAL AND ARTIFICIAL NEURAL NETWORKS	58
2.5	SUMMARY	60

CHAPTER 3

THE VISUAL SYSTEM AND VISUAL MEMORY

INTRODUCTION	61
3.1 BASIC FEATURES OF THE VISUAL SYSTEM	63
3.1.1 <i>The retina</i>	63
3.1.2 <i>The lateral geniculate nuclei (LGN)</i>	64
3.1.3 <i>Receptive fields (RF)</i>	65
3.1.4 <i>The primary (striate) visual cortex (V1)</i>	67
3.1.5 <i>Extrastriate cortices</i>	70
3.2 VISUAL INFORMATION PROCESSING	71
3.2.1 <i>Feed back, feed forward and horizontal connectivity</i>	72
3.2.2 <i>The parietal cortex</i>	75
3.2.3 <i>Neurological evidence for functional specificity in visual processing</i>	76
3.3 VISUAL SHORT-TERM MEMORY	78
3.3.1 <i>Working and visual short-term memory systems</i>	78
3.3.2 <i>Memory Masking</i>	81
3.3.3 <i>Limited capacity of VWM</i>	83

3.3.4	<i>The role of prefrontal cortex, inferotemporal cortex and other areas in VWM</i>	84
3.3.5	<i>Neuronal firing</i>	87
3.4	SUMMARY	89

CHAPTER 4

THE BINDING PROBLEM

INTRODUCTION	91
4.1 PROPOSED SOLUTIONS	92
4.1.1 Coincidence and convergence	92
4.1.2 Spatial attention and the Feature Integration Theory (FIT)	94
4.1.3 Temporal binding (binding through synchronous firing)	96
4.2 FEATURE BINDING IN VWM	101
4.2.1 Synchrony	105
4.2.2 Spatial and temporal aspects of binding in VWM	106
4.3 COMPUTATIONAL MODELS of OBJECT FEATURE BINDING IN VWM:	108
4.3.1 Rolls (1992) VisNet	109
4.3.2 Usher & Niebur (1996) object oriented attention	110
4.3.3 Camperi & Wang (1998) visuospatial working memory in PFC	111
4.3.4 Usher & Cohen (1999) maintenance and response n in STM	111
4.3.5 Renart et al. (1999; 2001) attractor models of WM	112
4.3.6 Compte et al. (2000) model for spatial working memory	113
4.3.7 Raffone & Wolters (2001) memory and perception	114
4.3.8 De Kamps & van der Velde (2001) object-based attention	114
4.3.9 Deco and Lee (2002) visual information processing	116
4.3.10 Deco and Lee (2001) spatial and object-based attention	117
4.3.11 Domijan (2003) feature binding and capacity limits in VWM	118
4.3.12 Van der Voort van der Kleij et al. (2004) model of VWM	119
4.4 SUMMARY	120

CHAPTER 5

BEHAVIOURAL EXPERIMENT

INTRODUCTION.....	123
5.1 EXPERIMENTAL GOAL	123
5.2 METHOD	124
5.2.1 <i>Stimuli and coding</i>	124
5.2.2 <i>Stimulus onset asynchrony (SOA)</i>	125
5.2.3 <i>The experimental conditions</i>	126
5.2.4 <i>Response coding.....</i>	130
5.2.3 <i>Participants</i>	132
5.2.4 <i>Materials</i>	133
5.2.5 <i>Protocol.....</i>	133
5.3 HYPOTHESES.....	134
5.4 RESULTS	139
5.4.1 <i>Presentation rates</i>	139
5.4.2 <i>Overall errors</i>	142
5.4.3 <i>The Spatial condition</i>	147
5.4.4 <i>Spatio-temporal (spatial errors) condition</i>	148
5.4.5 <i>Spatio-temporal (temporal errors) condition</i>	150
5.4.6 <i>The Temporal condition</i>	152
5.4.7 <i>Comparison of spatial and temporal errors in the Spatio-temporal condition.....</i>	153
5.4.8 <i>Comparison of temporal errors in the Spatio-temporal condition and the Temporal condition</i>	155
5.5 ILLUSORY CONJUNCTIONS.....	156
5.5.1 <i>Overall ICs</i>	157
5.5.2 <i>Illusory conjunctions in the Spatial condition</i>	159
5.5.3 <i>Spatial illusory conjunctions in the Spatio-temporal condition....</i>	160
5.5.4 <i>Temporal illusory conjunctions in the Spatio-temporal condition</i>	161
5.5.5 <i>Illusory conjunctions in the Temporal condition</i>	163
5.5.6 <i>Comparison of spatial and temporal illusory conjunctions in the Spatio-temporal condition.....</i>	164

5.5.7 Comparison of temporal illusory conjunctions in the Spatio-temporal condition and Temporal condition	165
5.5 DISCUSSION OF RESULTS	166
5.6.1 Presentation rates	168
5.6.2 General findings.....	169
5.5 THREATS TO VALIDITY AND LIMITATIONS.....	175
5.7.1 External validity.....	176
5.7.2 Construct validity.....	177
5.7.3 Internal validity.....	178
5.7.4 General limitations	179
5.7 SUMMARY	181

CHAPTER 6

OBJECT FEATURE BINDING IN VISUAL WORKING MEMORY: COGNITIVE MODELLING

INTRODUCTION.....	183
6.1 RATIONALE FOR THE COGNITIVE MODEL	183
6.2 DESIGN	184
6.2.1 The dorsal stream module (DS)	186
6.2.2 The ventral stream module (VS).....	187
6.2.3 The prefrontal cortex module (PF)	187
6.2.4 Operation of the multimodular cognitive model	187
6.2.5 Predictions	189
6.2.6 Discussion of cognitive model in relation to behavioural data	192
6.3 OUTLINE OF THE PROPOSED COMPUTATIONAL MODEL....	197
6.3.1 The autoassociative recurrent attractor network: VS module	199
6.3.3 The competitive self-organising network: PF module	202
6.4 EVALUATION OF MODEL.....	205
6.5 DISCUSSION AND CONCLUSIONS	206
6.6 SUMMARY	207

CHAPTER 7
DISCUSSION AND CONCLUSIONS

INTRODUCTION.....209

7.1 SUMMARY OF LITERATURE210

7.2 EXPERIMENTAL AND MODELLING METHODOLOGIES211

7.2.1 Evaluation of experimental methodology.....211

7.2.2 Evaluation of modelling methodology213

7.3 CONTRIBUTIONS TO KNOWLEDGE.....214

7.4 LIMITATIONS AND FURTHER WORK.....216

7.4.1 Further experimental work217

7.4.2 Further modelling work218

7.4.3 Summary of further work219

7.5 AUTHOR’S PUBLICATIONS221

REFERENCES.....223

BACKGROUND AND OVERVIEW OF THESIS

INTRODUCTION

Neuroscience is young and multidisciplinary, encompassing diverse yet interrelated sciences such as physiology, physics, mathematics, computer science, psychology and biology. There are many branches of neuroscience including cognitive and computational neuroscience. Cognitive neuroscience is concerned with the neural mechanisms of mental processes in the brain, whereas computational neuroscience is an approach to understanding the brain by simulating the nervous system using computer modelling techniques.

There are few, if any, places in the nervous system which process all the information necessary to carry out a particular task. Therefore sensory, cognitive and motor processes must emerge from parallel interactions among large and widely distributed neural populations. Understanding how information that is processed in these distinct areas is integrated, is commonly known as the ‘binding problem’ (Rosenblatt, 1961; Treisman and Gelade, 1980; Von der Malsburg, 1981; Treisman and Schmidt, 1982). There are several ‘binding problems’ ranging from the unity of consciousness to object feature binding. The latter is the focus of this thesis.

Objects comprise individual features such as colour, form and orientation, each of which is processed in distinct brain regions. However, we can perceive coherent object representations. When multiple objects are present, individual features must be correctly bound to a particular object and segregated from other objects. Despite continued research effort and a large literature, there is no agreed solution to the object feature binding problem (e.g. Roskies, 1999). In neuroscientific terms, the concern is how information, distributed in patterns of neural firing, results in coherent representations.

1.1 MOTIVATION FOR STUDY

Despite the inherent complexity of visual scenes, humans rapidly recognise and act on visual information. Neurophysiological evidence has shown that the visual system decomposes objects into visual features which are processed in distinct and widely distributed brain regions. Then, in order to respond rapidly to visual stimuli, information must be flexibly co-ordinated across and between sensory modalities in a unified neural process. The binding problem exists when more than one visual feature of an object must be correctly associated to provide a coherent representation of that object. Correct association becomes crucial when multiple objects are present and incorrect combinations known as illusory conjunctions (ICs), of features that belong to different objects must be avoided. In most circumstances binding is not a problem for the nervous system, yet because of its apparent simplicity but inherent complexity, it remains a conceptual problem for neuroscientists, psychologists and philosophers. Several solutions have been proposed, but to date, none individually solves the binding problem.

In order to understand complex systems scientists and engineers build models. However, building realistic models to perform specific functions, and understanding how representations are stored in memory, remains a challenge. Moreover, understanding which mechanisms facilitate the integrative processing involved in dynamic feature binding is demanding (Wersing and Ritter, 1999). Ultimately, research such as is presented in this thesis could advance our understanding of neural mechanisms through identifying principles appropriate for building more robust and efficient machine vision systems.

1.2 THESIS AIM AND OBJECTIVES

The aim of this thesis is to further understand how and where object feature binding takes place by means of a behavioural experiment and cognitive modelling.

In order to achieve this aim a number of objectives need to be met. These are listed below:

- To thoroughly review and critique literature relating to object feature binding and visual short-term memory
- To outline, review and critique artificial neural systems in general and models of object feature binding in particular
- To design and conduct a behavioural experiment to test hypotheses relating to spatio-temporal aspects of object feature binding
- To develop a cognitive model based on existing evidence to explain the resultant behavioural data
- To use the cognitive model to inform the design of a biologically plausible multimodular attractor network model to be implemented at a later date.

Implicitly, the effects of distributed and local cortical interactions in visual information processing and visual short-term memory (VSTM) will be explored. Ultimately, research such as presented in this thesis could advance our understanding of neural mechanisms through identifying principles appropriate for building more robust and efficient machine vision systems.

1.3 THESIS ORGANISATION

A minimum prerequisite of neuroscience is a basic understanding of the structure and workings of individual biological neurons, neuronal networks and systems of networks, as well as how they are represented in artificial neural models. Chapter 2 introduces basic concepts of biological and artificial neural networks and highlights their differences. The main emphasis is on attractor networks.

Chapter 2 provides an overview of human visual and short-term memory systems and their known information processing mechanisms. The problem of how we integrate fragmented perceptions into coherent wholes, the binding problem, arises from the inherent complexity and distributed nature of visual information processing and memory. Despite being approached by investigators from psychology (Treisman and Gelade, 1980), physiology (Nobre et al., 2003), and computational modelling (e.g. Wersing and Ritter, 1999) for over twenty years, finding a solution to the binding problem is elusive.

Chapter 4 describes existing theories and computational models of object feature binding. Hypotheses arising from these theories are tested in the behavioural experiment (Chapter 5). Limitations of the experimental design and explanatory power are provided at the end of Chapter 5. The rationale, design and limitations of a novel cognitive model, based on existing evidence, intended to explain the resultant behavioural data, is given in Chapter 6. This cognitive model is used as the foundations of the design of a computational model. Details of this model and its limitations are in the later parts of Chapter 6.

The computational model is not implemented at this stage, but is intended for further work and is detailed in Chapter 7 along with a synthesis of the experimental and modelling methodologies, the rationale for their design and descriptions of the findings. Limitations of the work carried out and proposed further work are presented prior to the contributions to knowledge and conclusions proposed in this thesis.

The present study thus addresses the issues outlined above and tests, through behavioural experimentation and cognitive modelling, hypotheses (expanded in subsection 5.2.4.6) concerned with target proximity, and spatial and temporal properties on object feature binding. We concentrate on single feature binding errors and illusory conjunctions¹ (ICs).

Thus in summary we hypothesise that:

- In the Spatial and Spatio-temporal condition, in support of the Spatial Theory, the proportion of spatial observed errors and ICs will differ from random and there will be more adjacent spatial errors and ICs than intermediate or distant errors and ICs.
- In the Spatio-temporal and Temporal condition, in support of the Temporal Theory, the proportion of temporal observed errors and ICs will differ from random and there will be more adjacent temporal errors and ICs than intermediate or distant errors and ICs
- In support of the Spatial Theory, in the Spatio-temporal condition, the proportion of observed spatial and temporal errors and ICs will differ from random, and there will be more spatially than temporally adjacent

¹ An illusory conjunction is the phenomenon in which features from one object present in a scene, are transferred to another object which is not present, to create the illusion of a composite object

errors and ICs in the Spatio-temporal condition than temporal presentation.

- In support of the Spatial Theory, there will be a difference in the proportions of temporal errors and ICs in the Temporal and Spatio-temporal conditions, and there will be more adjacent errors and ICs in the Spatio-temporal than the Temporal condition.

1.4 RESEARCH METHODOLOGY

A behavioural experiment was conducted using a within-subjects design because it affords less error variance and greater power which reduce Type 2 error rate. Fifteen volunteer participants with normal or corrected-to-normal vision completed a delayed recall task in three experimental conditions. Stimuli comprised a battery of sixty-four composite objects each with three features and one of four feature dimensions. In the experiment the three features were colour, orientation and size; a feature dimension defines the variation of each feature. Hence dimensions of the feature colour were purple, turquoise, green and yellow; of the feature orientation, vertical, horizontal, right diagonal, and left diagonal; of the feature size, four sizes ranging small to large. The objects were designed so that each feature dimension was presented 4 times within the 64 object stimuli. Results were analysed using Chi-squares tests for goodness of fit, and multiple *t*-tests to compare means.

The cognitive model, designed by incorporating and integrating existing evidence, was intended to explain the resultant experimental data and to inform the proposed computational model whose design is based adaptations of existing evidence integrated in a novel and biologically plausible approach. The cognitive model presents a unified theory of object feature binding. The proposed computational model is a high level abstraction of a complex system modelled as a multimodular recurrent attractor network (Chapter 6) that could be implemented at a later date. To the author's knowledge, no other model, cognitive or computation explains the data and neural mechanisms using this approach.

1.5 RESEARCH SCOPE AND LIMITATIONS

Scope

The cognitive model aims to i) explain data derived from the behavioural experiment and ii) provide a starting point for a computational model. Because the proposed computational model is necessarily concerned with the emergent properties of networks of networks, the detailed and biologically plausible cognitive model will be simplified to lead to a high level computational model. Hence, evidence from neuronal physiology (Chapter 2), the anatomy of visual information processing and working memory systems (Chapter 3) are integrated into the model.

Limitations

The work is concerned with vision and VSTM systems and thus does not consider other sensory systems, or episodic, or long-term memory systems. Furthermore, although details of a proposed computational model are provided (Chapter 6), the model is not implemented in this thesis. It is considered to be potentially valuable further work.

1.5.1 Focus of study

This thesis focuses on information processing in both the visual system and in visual working memory, the integration of information processing within and between these systems and their emergent behaviour. In order to gain insight into these systems, the thesis focuses on biological, neurophysiological and psychological aspects of information processing as well as the physical aspects involved in computational modelling. Neurophysiological evidence shows that individual neurons convey information via electrochemical pathways in a process which is not binary, not stable, and not synchronous (Anderson and McNeill, 1992), throughout widely distributed cortical and subcortical structures. Cognitive neuroscience attempts to explain how a specific electrochemical brain state can give rise to a given experience. Computational neuroscience uses models to evaluate and organize experimental data in an attempt to understand how the organisation of the brain is related to its functions.

The current study uses a cognitive neuroscience model to understand the behavioural results and test the hypotheses, and proposes testing at a later stage,

by means of a modular attractor neural network. The hypotheses are that object feature binding takes place in visual memory and that individual object representations are formed as memories which are distinguished by their spatial or temporal properties. Thus the multimodular network comprises intra and interconnected networks that represent the visual dorsal and ventral information processing streams and prefrontal cortex.

1.5.2 Application domain

Both cognitive and computational neuroscience typically involve the concerted efforts of neuroscientists, psychologists, physicists, computer scientists, mathematicians, and investigators in artificial intelligence concerned with understanding the brain through the design and construction of models of neurally inspired information processing systems. The application of such biologically plausible architectures could possibly enable engineers to build machines which incorporate some characteristics of human brains, particularly robustness, distributed memory, and the ability to generalise from and interpret imprecise and noisy information. Furthermore, a fuller understanding of physiological brain processes obtained through computational modelling could help physicians diagnose and ultimately treat some neuropsychological dysfunctions.

1.5.3 Empirical investigation

In order for findings to be reliable, the same results would be produced if the study were replicated. Unless a measure is reliable, it cannot be valid (Robson, 1995). Validity is concerned with whether the findings actually relate to what they are supposed to relate to. That is, are the findings 'true' or are they due to the effect of other artefacts? Generalizability refers to the extent to which findings can be applied to other contexts, situations or populations. Reliability, validity and generalizability are key issues in establishing trustworthiness in research findings. To the extent that unreliability can arise through subject or observer error or bias, the physical environment in this study was controlled, the participants were sampled from the same population, and the same experimenter conducted the experiments. Threats to validity and general experimental limitations are fully discussed in Section 5.8.

Having established reliability, the experiment aimed to measure effects of spatial and temporal position on object feature binding, and how these relate to the proximity of the objects. The empirical investigation comprised a set of experiments designed to enable the collection of behavioural data. From these data cognitive model has been developed and is used to inform a computational model which is a multimodular attractor network using a modified Hebbian rule. Thus the proposed neural model is both biologically plausible and appropriate for modelling memory.

1.5.4 General empirical limitations

The experiments were conducted in a controlled, hence artificial environment, with each participant only once. These factors affect generalizability and to some degree, reliability. However, conducting multiple observations with the same participant was not practical because of the time involved in conducting each condition. Furthermore, learning effects and fatigue were minimized because each participant took part in each condition only once. The order of the three conditions was randomised to reduce any effects that might have occurred by running the conditions in the same order for each participant. The controlled environment in which the study was conducted, and the stimuli used, meant that the results cannot be readily generalized to other situations. Further studies could be conducted using 'real' images in a natural environment. The software programme Macromedia Director, used to run the experiment was selected because of the experimenter's experience and its availability.

1.5.5 Rationale for methods

Having conducted several pilot studies, the experimental methods (Chapter 5) described in Chapter 5 are considered to be the most appropriate. The cognitive model based on existing evidence is biologically faithful and detailed. It takes a novel, unifying approach to object feature binding. Furthermore, this detailed and realistic model is designed to inform a computational model. However, this proposed neural model has not been implemented in this study due to time constraints, but is suggested as future work. Indeed, further work to the model, suggested in Section 7.4 could add refinements so that it would become more detailed, sophisticated and ultimately more insightful.

1.6 CONCLUSIONS

This thesis is concerned with object feature binding and visual short-term memory. To this end, it tests hypotheses by means of behavioural experimentation. It derives an original and detailed realistic cognitive model in order to understand and explain the neural mechanisms underlying the behavioural data. This model is designed to inform an equally detailed and realistic computational model. Section 7.4 describes further work that could be carried out on both models.

1.7 SUMMARY

This chapter has given a background and introduction to cognitive and computational neuroscience, the binding problem and visual short-term memory. The motivations, objectives for the present study were outlined, and the research methodology described and critiqued. The scope, limitations, foci and application domain of the research were presented. Finally, an overview of the thesis, its limitations and rationale for the methods used are provided. The next chapter considers biological and artificial neural networks, their history, application and differences.

BIOLOGICAL AND ARTIFICIAL NEURAL NETWORKS

'The neuron can be thought of as a computational element which sums its inputs within its time constant and whenever this sum, minus any inhibitory effects, exceeds a threshold, produces an action potential which propagates to all of its outputs' (Rolls and Treves, 1998).

INTRODUCTION

A minimum prerequisite of computational neuroscience is a basic understanding of the structure and workings of individual biological neurons, neuronal networks and systems of networks, as well as how they are represented in artificial neural models. This chapter provides an introduction and overview of the basic concepts of biological and artificial neural networks and highlights the differences between them. The main emphasis is on attractor networks.

2.1 BIOLOGICAL NEURONS

The human brain contains approximately 10^{11} neurons, each with approximately 10^4 synapses which communicate through a connection network of axons and synapses. Biological neurons operate in a chemical environment and communicate through electrical impulses (Arbib, 1987) in a process which is 'not binary, not stable, and not synchronous' (Anderson and McNeill, 1992). Understanding the mechanisms and speed by which humans are able to assimilate novel associations remains a central issue of cognitive science.

Despite variations in morphology, function and specialization, the majority of biological neurons have in common four basic components (Fig. 2.1): dendrites, soma, axon, and synapses. Dendrites are extensions of the soma (cell body) which receive input through the synapses of other neurons. The soma processes incoming signals over time and outputs the processed value to other neurons through the axon and the synapses. The axon is a thin fibre of constant diameter which sends an impulse towards other neurons, a gland or a muscle. Swellings at the ends of an axon's many branches form the pre-synaptic terminal

(bouton) containing vesicles from which a neurotransmitter is transmitted to post-synaptic neurons across the synaptic cleft.

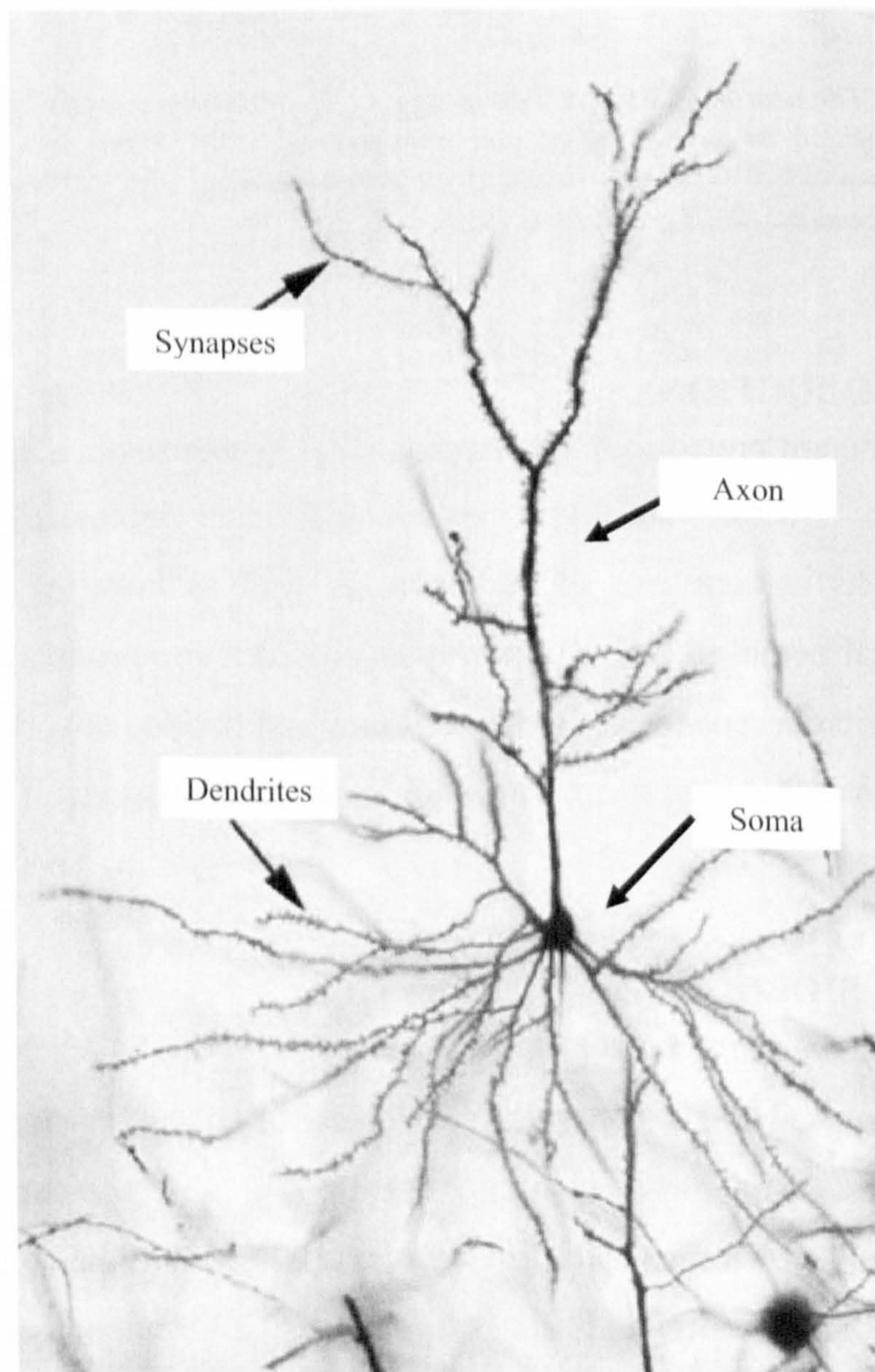


Figure 2.1 A Pyramidal Neuron. The most prolific type of neuron in the cortex is the pyramidal cell. The four main parts are the synapses, axon, dendrites and soma.
Source: <http://www.genesis-sim.org/GENESIS/cnsweb/cns2b.html>

The majority of neurons typically receive signals from approximately 10,000 other neurons. However, motor neurons receive far fewer, and hippocampal cells, far more. Every neuron has a single axon which can branch to send information to different brain regions. Once an action potential has been triggered at the beginning of an axon, some sodium (Na^+) gates open in this region. The inward flow of positive ions decreases the negativity on the inside causing Na^+ influx which further decreases negativity inside the membrane, and depolarizes it

causing more Na^+ to enter. When enough Na^+ ions have entered, a threshold is reached, the voltage difference across the membrane reverses, and the inside of the membrane becomes more positive relative to the outside. This is the refractory period in which Na^+ gates close and potassium (K^+) gates open allowing K^+ to flow out until the voltage is restored and both Na^+ and K^+ gates are shut. During the refractory period, the membrane is virtually insensitive to stimulation, which helps prevent backflow, but the electrical disturbance triggers an action potential at an adjacent membrane site away from the point of stimulation towards the axon terminal. Although the inside of the membrane is negative again, the concentrations of Na^+ and K^+ are reversed so active transport pumps work to restore the original internal membrane concentrations and voltage. These properties are important in developing biologically plausible artificial neural networks.

2.1.1 Synapses, neurotransmitters and receptors

The most common type of connection between neurons is the chemical synapse which consists of the axon terminal and specific receiving sites on the dendrites. Axon terminals can synthesize neurotransmitters which are stored in synaptic vesicles². When a neuron fires, the action potential (also known as a spike) triggers the release of neurotransmitters which drift across the synaptic cleft between the axon terminal and the dendrite of the post-synaptic neuron (Fig. 2.1.1).

A neurotransmitter is a type of molecule found in the nervous system. Neurotransmitter molecules pass between pre- and post-synaptic neurons at the synaptic cleft³. Neurotransmitters are packaged in vesicles within the synapses and on the arrival of an impulse they are released and diffuse across the synaptic gap to bind neurotransmitter receptors or other ion channels and stimulate or inhibit the firing of the postsynaptic neuron. Crudely, neurotransmitters have excitatory or inhibitory effects on post synaptic neurons which are implemented as positive or negative weight changes in some artificial neural networks.

² In cell biology, a vesicle is a relatively small and enclosed compartment, separated from the internal fluid of the cell by at least one lipid (a water-insoluble organic molecule) membrane. Vesicles store, transplant or digest cellular products and wastes.

³ The synaptic cleft is a gap at the end of a synapse where neurotransmitters are released via vesicles (fluid filled pouches).

However, in order for such a model to lead to a further understanding of brain processing, complex properties of neurotransmission should be taken into account.

Some excitatory neurotransmitters such as glutamate⁴ open channels that allow positively charged ions to enter the cell. This triggers an increase in the membrane potential that drives the post-synaptic neurons towards their excited state. In contrast, inhibitory neurotransmitters, such as gamma-amino-butyric acid⁵ (GABA) drive the post-synaptic potential towards its resting potential. Other neurotransmitters and receptors modulate the effectiveness of synapses through a sequence of biological processes which can have a variety of consequences. Most often, neurotransmitters are removed from the synaptic gap in a process of re-uptake⁶ after they have activated their specific receptors by transport proteins residing in neuronal and glial⁷ plasma membranes. Re-uptake ensures that neurotransmitters do not continue to excite or inhibit postsynaptic firing.

Most often, a neuron has receptors for one or two neurotransmitters, but different neurons may use the same neurotransmitter for different purposes. Furthermore, a neurotransmitter may have either an excitatory or inhibitory effect on the post-synaptic neuron, and which neurotransmitter is used depends on the type and location of the neuron. However, the main function of any neurotransmitter is to open the receptor channel in response to binding with a neurotransmitter. In many cases, it is the receptor which determines whether the transmitter is excitatory or inhibitory. Receptors can also determine whether a transmitter acts rapidly by direct action on an ion channel or slowly by a second-messenger system that allows for synaptic plasticity and leads to learning and memory. The speed and mechanics of transmitter inactivation after the signal has been sent are also important. Clearly, these low level factors have implications for realistic neural modelling.

Neurotransmitter receptors are often described as N-methyl-D-aspartate (NMDA), non-NMDA, or kainite-quisqualate (K-Q). NMDA channels are

⁴ Glutamate is the major excitatory neurotransmitter of the nervous system.

⁵ GABA is the major inhibitory neurotransmitter of the nervous system.

⁶ Reuptake is the reabsorption of a neurotransmitter by the molecular transporter of a pre-synaptic neuron after it has performed its function of transmitting a neural impulse.

⁷ Glia are supportive tissue of the brain. In contrast to neurons, glial cells do not conduct electrical impulses.

normally blocked by magnesium ions (Mg^{2+}), but when the cell is strongly depolarized by strong stimulation enough to induce long term potentiation (LTP), the Mg^{2+} block is removed and calcium ions (Ca^{2+}) entering via the NMDA channels trigger events that lead to the potentiated synaptic transmission. The voltage dependence of the NMDA receptor channels introduces a threshold and thus a non-linearity that contributes to a number of LTP phenomena such as plasticity leading to memory and learning.

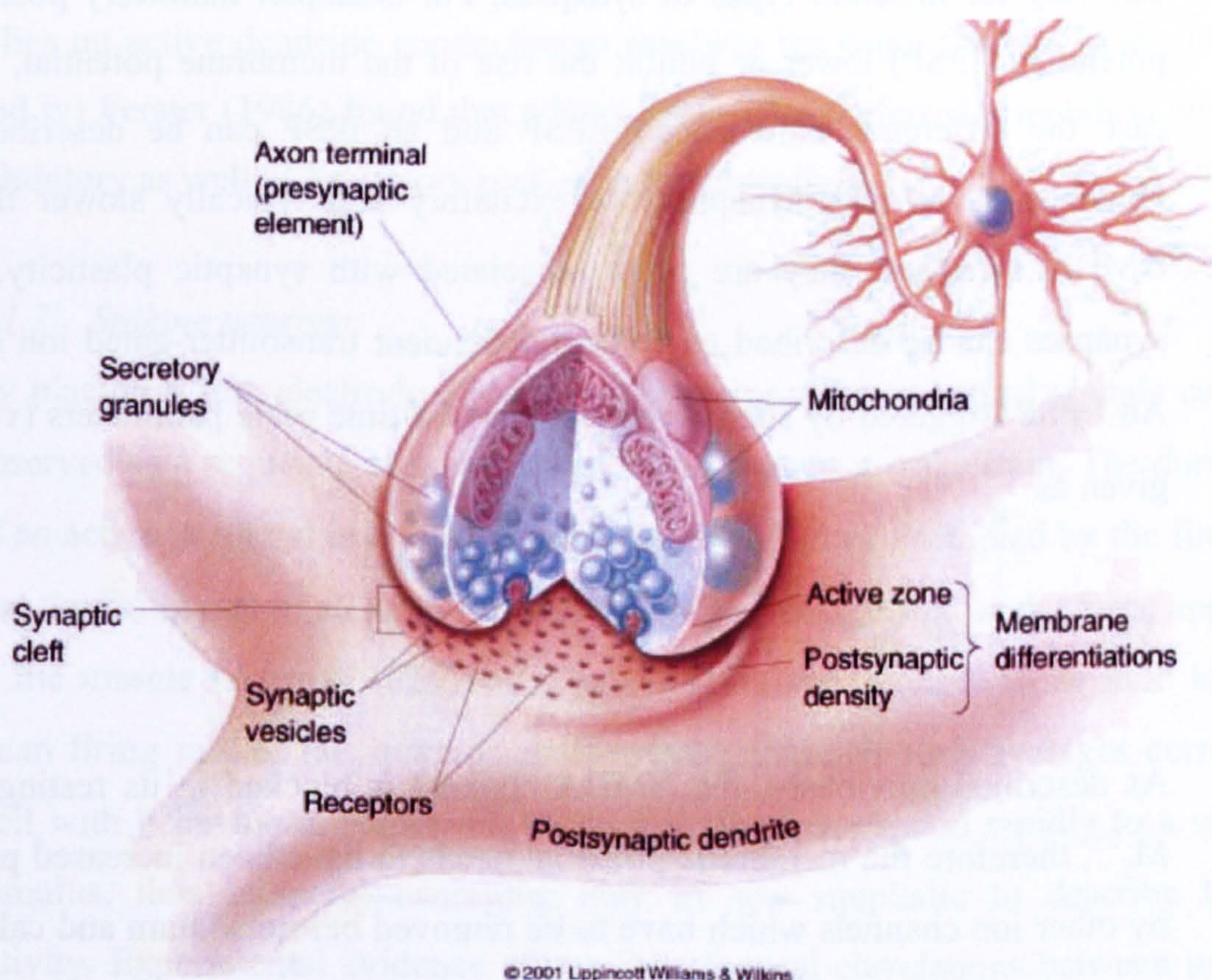


Figure 2.1.1 Schematic of a chemical synapse. When one nerve passes an impulse to the next cell a chemical event intercedes between the arrival of the impulse at the terminal of one axon and the continuation of the signal in the next cell. The juncture is referred to as a synapse. The cells are separated by a gap, the synaptic cleft. Upon appropriate stimulation the pre-synaptic cell releases a neurotransmitter. The neurotransmitter then diffuses across the cleft to receptors on the postsynaptic membrane. The receptors respond by initiating depolarization of the membrane and the impulse is propagated in the postsynaptic cell. Source: <http://www.psych.purdue.edu.html>

Excitatory synapses increase membrane potential in excitatory post-synaptic potentials (EPSPs) following the firing of the pre-synaptic neuron. This is typically after a delay of <1 ms. in which the neurotransmitter is released and

diffused, and the ion channels open. An excitatory post-synaptic potential (EPSP) from non-NMDA receptors is often described by

$$\Delta V_m^{\text{non-NMDA}} = wte^{-t/t_{\text{peak}}} \quad (2.1)$$

where ΔV_m denotes the change in the membrane potential, which is a function of time t after the delay, and w is the strength of the EPSP. The functional form of the EPSP is often denoted as α . The scale of the rise and fall, denoted as t_{peak} can vary for different types of synapses. For example, inhibitory post-synaptic potentials (IPSP) lower or inhibit the rise of the membrane potential, in which case the difference between an EPSP and an IPSP can be described by w . However as NMDA synapses are excitatory and typically slower than non-NMDA synapses, they are often associated with synaptic plasticity. NMDA synapses can be described as voltage-dependent transmitter-gated ion channels. An EPSP triggered by such a channel has two time scale parameters (τ_1 and τ_2), given as

$$\Delta V_m^{\text{NMDA}} = c(V_m)e^{-t/\tau_1} - e^{-t/\tau_2} \quad (2.2)$$

As described previously, the NMDA channel is blocked in its resting state by Mg^{2+} , therefore the membrane potential needs to have been increased previously by other ion channels which have to be removed before sodium and calcium can enter the neuron. The membrane potential also depends on several other factors including the state of the neuron, and the sum of individual potentials. For example if the membrane potential is at rest, GABA has no direct effect, but can produce divisive or shunting (subtractive) inhibition by reducing the effect of excitatory currents through non-linear interactions which add to the information-processing capabilities of neural networks. The resting potential of a neuron can be altered by the release of neurotransmitters that open specific ion channels and lead to the generation of a spike.

Conditions that cause one neuron to fire, also cause a considerable fraction of its inputs to fire⁸ (Grinvald et al., 1994). Thus in order to maintain a dynamic range of responses excitation needs to be balanced with inhibition. Inhibition is achieved through several mechanisms: i) at subthreshold membrane potential, outward currents last longer than input currents, have synaptic contacts closer to the soma, and follow the weaker electrochemical gradients (Beaulieu et al., 1992); ii) inhibitory inputs tend to make multiple synapses on a single neuron; iii) inhibition may have a disproportionately large impact by affecting the gain of the excitation of the distal dendrite (Bernander et al., 1994) which relies on active dendritic conductances reaching the soma (Amitai et al., 1993); and iv) Ferster (1986) found that a bar of the cell's preferred orientation elicited inhibitory as well as excitatory post-synaptic potentials.

2.1.2 *Spiking neurons*

By placing a fine electrode close to the soma or axon, neuronal signals can be observed as a sequence of action potentials known as a spike train. The duration of an action potential is typically in the range of 1-2ms. Prompted by the finding that the firing rate of stretch receptors in muscles was related to the force applied to the muscle (Adrian, 1926), most information was thought to be held in the mean firing rate of the spike train. However, although such averages correlate well with behavioural responses, biological systems respond rapidly to a single stimulus, thus temporal averaging may be too simplistic to describe brain activity. Experimental evidence of precise temporal correlations between pulses of different neurons (Abeles, 1994), and stimulus dependent synchronization activity in populations of neurons (Singer, 1994) showed that such temporal averaging neglects any information that might be contained in the exact timing of the spikes. A more plausible measure could be obtained by taking an average from a subpopulation of neurons with similar response properties engaged in neural response at a specific time, rather than a single neuron over time. In this way, the temporal averaging used in physiological experiments would represent

⁸ Neurons within a cylindrical radius of 50-100µm of a cortical column respond under similar conditions and are responsible for approximately half the 3000-10000 excitatory synapses to a neuron (Braitenberg and Shuz, 1991), and many inputs from outside such a column come from horizontal connections from columns with overlapping properties (Gilbert and Wiesel, 1989).

spatial averaging used in real systems. Generally, firing rates of a particular neuron tend to increase in a short time interval following the presentation of an effective stimulus. The set of stimuli that increases the firing rate is the ‘receptive field’ (see Chapter 3).

When the membrane potential reaches a threshold, it emits a spike typically characterized by a sharp increase (depolarization) to positive values followed by a sharp decrease to below the resting potential (after hyperpolarization), then a return to the resting potential level (Fig. 2.1.2). This form of a spike was originally measured by Hodgkin and Huxley (1952) on the giant axon of a squid and described in the form of a set coupled differential equations in which the activation of the potassium and sodium channels, and the inactivation of the sodium channel were modelled as dynamic variables. In these equations, conductances were the net result of individual ion channels in the membrane whose densities had to exceed the threshold in order to produce a spike, after which was the absolute refractory period.

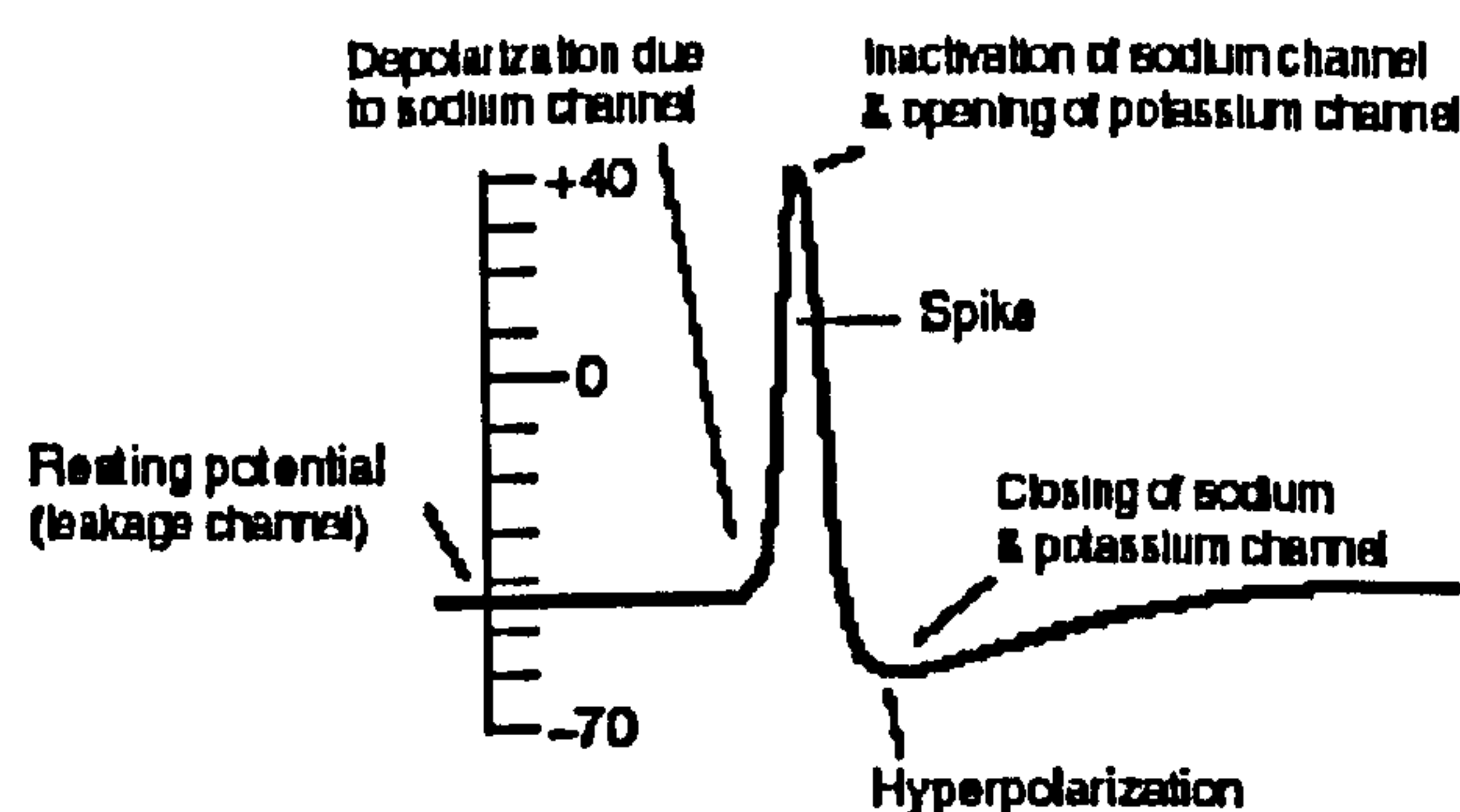


Figure 2.1.2 Typical form of an action potential, redrawn from an oscilloscope picture of Hodgkin and Huxley. Reproduced from Trappenberg (2002).

In order to produce constant firing with a waveform similar to a real neuron, Hodgkin and Huxley (1952) used a constant external current with strength $I^{\text{ext}} = 10$. In contrast to the giant axon of the squid, mammalian neurons have several types of ion channels that lead to complex responses. Yet, despite this, the influence of the stereotypical post-synaptic response to the pre-synaptic spike, described by the α function, is often considered more important in neural modelling than the precise form of the spike.

Bi and Poo (1998) demonstrated a crucial temporal relation between pre- and post-synaptic spikes by varying the time period between them. This period was equal to the time difference between the excitatory post-synaptic current (EPSC) (induced by the stimulation of the pre-synaptic neuron), and the peak of the post-synaptic action potential (induced by stimulating the post-synaptic neuron beyond its firing threshold). The critical time window was found to be approximately $|\Delta t| \approx 40$ ms. After this period, no synaptic plasticity occurred. They found that changes in EPSC amplitudes were largest for small positive (LTP) or negative (LTD) differences in pre and post-synaptic spike times, and that the absolute strength of the synaptic efficiencies in long term depression (LTD) was proportional to the initial synaptic efficiency. In contrast, the relative changes of EPSC amplitudes for LTP were largest for small initial EPSC amplitudes. The relative amplitude changes declined with increasing initial EPSC amplitude

Sections 2.1.1 and 2.1.2 outlined basic cellular and subcellular neural and neuronal structures and processes to provide a foundation for understanding more complex behaviours at higher levels of abstraction. However, individual neurons connect to each other and operate in networks of non-linear interacting systems that provide a distributed representation of information. These networks are further incorporated into larger structures, which are capable of increasingly more complex tasks. Furthermore, neural systems adjust their response according to the external environment in which they operate. Structures and processes of neurons operating in network and networks of networks are described in Section 2.1.3.

2.1.3 Biological neurons in networks

Computational functions are not accomplished in the brain by single neurons operating in isolation, but are an emergent property of architecturally specialized networks. Integration into such networks is essential for information processing, yet detailed knowledge of brain anatomy and physiology remains incomplete. Therefore, in order to enable computational tractability for modelling, some simplifying assumptions need to be made. This section provides an overview of neurons in networks.

Regions of the outer layer of the cerebral cortex, the neocortex, are divided into four physically distinguishable lobes: occipital, parietal, frontal and temporal (Figure 2.1.3a). Other divisions have been made based on their functional correlations revealed by lesions, imaging or electrophysiological measures. Despite diverse functional specificity, different neocortical areas have common neuronal organization including anatomically distinguishable layers and functionally distinguishable columns (Chapter 3 describes such specificity relevant to visual information processing).

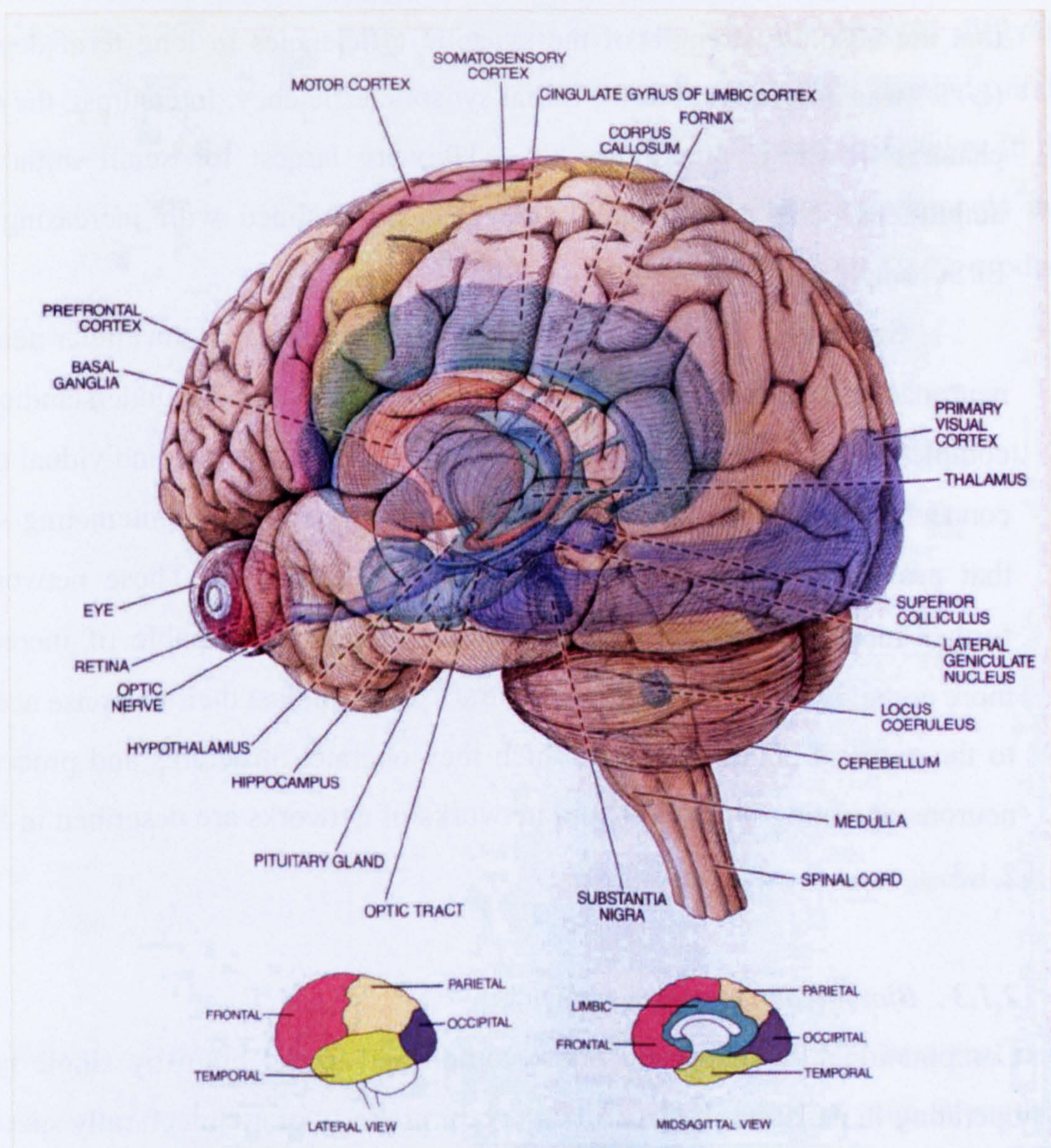


Figure 2.1.3a The neocortex can be defined in terms of its four lobes: frontal, temporal, parietal and occipital, or functional structures. Source:<http://pages.pomona.edu>

The most abundant neocortical neurons are pyramidal (75% - 90%), next prolific are stellate neurons, which may either be spiny and excitatory, or smooth and inhibitory. Other types of neurons are Martinotti cells, found only in deep layers

of the neocortex; basket cells, a variety of smooth stellate cells that synapse onto the cell body of pyramidal cells; and chandelier cells, which synapse preferentially to the initial segment of pyramidal axons. Staining reveals six neocortical layers labelled I to VI. Some layers are further subdivided into IVA, IVB, and IVC, and some further divided to $IVC\alpha$ and $IVC\beta$ (Fig. 2.1.3b).

The thickness and neural composition of the layers varies throughout neocortex. For example, layer I contains relatively few cell bodies. Pyramidal cells can be found in most other layers, but predominately in layers II and III. Stellate cells are concentrated around layer IV. Large pyramidal cells are mostly found in layer V, while most types of cells can be found in layer VI. Layer IV neurons receive many afferents from subcortical and other cortical areas. Pyramidal cells in layer V contribute to the output of cortical processing, whereas in layers II and III pyramidal cells are responsible for long-range tangential connections.

Stellate neurons tend to be more local. In fact, inhibitory smooth stellate cells are used to stabilize cortical processing.

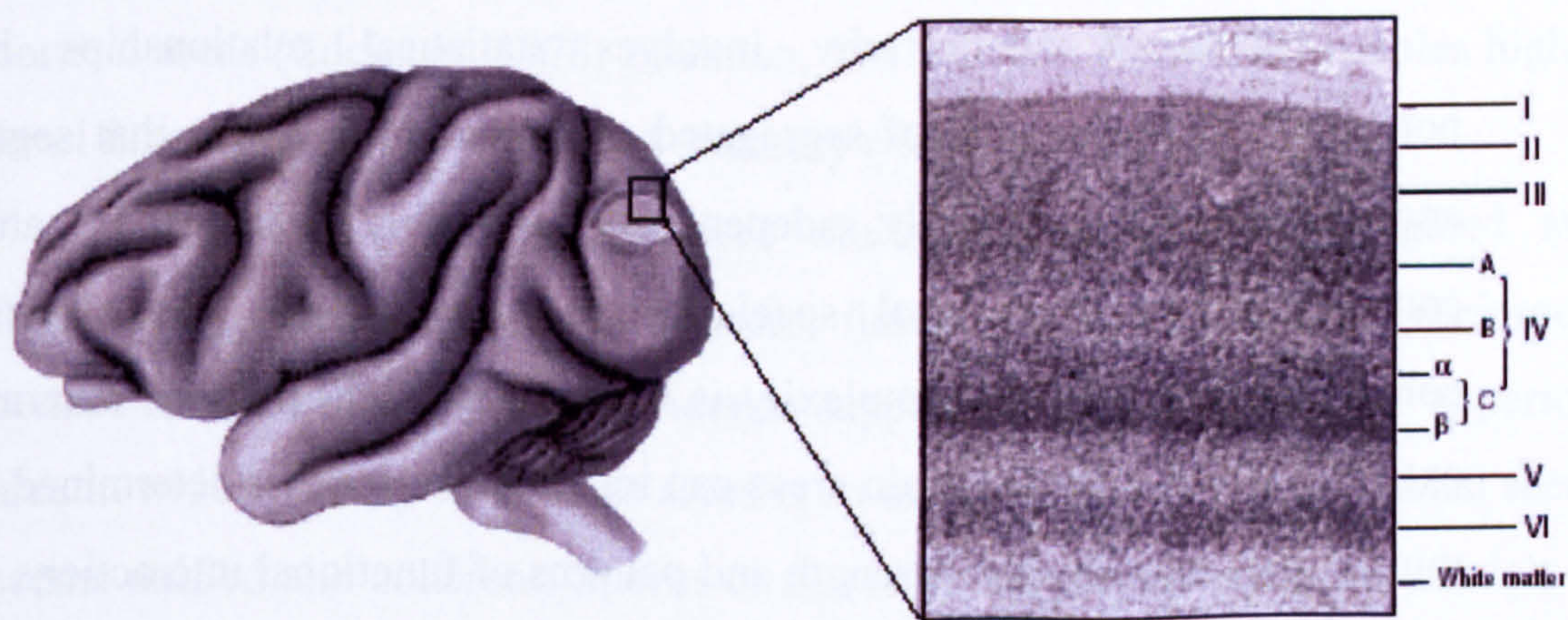


Figure 2.1.3b Stained neocortical slices showing layered structure. From Trappenberg (2002).

Despite this neuron type-specificity, many computational abilities of biological neural networks do not appear to depend critically on specific details.

2.1.4 Biological neural systems (networks of networks)

Neural integration occurs over numerous specialized neuronal groups and segregated brain areas. Functional integration is associated with patterns of

functional connectivity, expressed as deviations from statistical independence⁹ (e.g. temporal correlations) across the groups and areas. That is, the patterns of functional connectivity are statistically dependent so that knowledge about the value of one of them yields information about the value of another. Complexity captures such deviations from statistical independence across all levels of organization and quantifies the extent to which specialized units are globally integrated. Thus complexity is an important concept for integration.

Functional integration occurs within a structural substrate defined by the underlying neural network anatomy. For example, computer simulations have shown that different anatomical structures produce different patterns of functional connectivity. This concept is clearly illustrated in the visual cortex which comprises several distinct specialized areas (Felleman and Van Essen, 1991) in which functional segregation extends to the level of local columns or groups of neurons with separate specialized neural populations (see Chapter 3). How information from these specialized areas is integrated and combined into a coherent whole is known as the 'binding problem' (Treisman, 1996) and is the central issue of this thesis.

Functional connectivity involves statistical relationships between potentially large numbers of segregated elements and requires that segregation and integration be mutually independent, yet highly coherent. If patterns of connectivity maintain local specialization (segregation) and allow global coherence (integration), complexity is high. Anatomical patterns determine the extent to which different brain areas can influence each other determined by their mutual interconnectivity. Strength and patterns of functional interactions are also influenced by the physiological effectiveness of individual pathways and their particular density and pattern. In addition, connection patterns themselves might be influenced by functional interactions within a given network circuit. Evidence (Markram et al., 1997) has shown that relative timing of pre and post-synaptic

⁹ When two events are independent, knowing whether or not one of them occurred makes the occurrence of the other neither more nor less probable. Similarly, when two random variables are asserted as independent, knowing something about the value of one of them does not yield any information about the value of the other. For example, the height of a person and their IQ are independent random variables.

spikes within a range of 10 ms. can have an effect on synaptic plasticity. Thus, functional connectivity can shape anatomical connectivity.

Information in the brain is transmitted via neuronal connections. Neurons can connect vertically (upwards or downwards to adjacent or other cortical layers), horizontally (within the same layer), long-range (to other neural ensembles) or short-range (within a neural ensemble). However, within a particular cortical area, distant neuronal ensembles are linked by long-range horizontal connections which form a dense patchy network (Gilbert and Wiesel, 1989). These cortical areas are linked by pathways (Felleman and Van Essen, 1991) which are almost always reciprocal and comprise axonal bundles that link distant neurons. Edelman (1987) proposed that dynamic and stochastic processes enable the brain to be selective in terms of exploiting the most useful neuronal groups. Pathways connecting distinct brain areas provide the structural basis for re-entrant interactions in which many parallel pathways can operate simultaneously. The process of re-entrant interaction allows independent stimuli to become related through dynamic interactions between cell populations over large distances leading to spatio-temporal correlations within and between cortical areas (Edelman, 1978; Edelman, 1987). Such interaction enables higher level functions such as vision and memory (Edelman, 1987).

The visual system comprises multiple anatomically segregated and functionally specialized cortical areas linked by a network of reciprocal connections. Despite a lack of evidence supporting the existence of a 'superior' area in which activity from these areas is co-ordinated, the perceived visual scene seems unified and coherent. Sporns et al. (1989) developed computer models to address this problem. Their models showed that dynamic re-entrant interactions between functionally specified groups can lead to patterns of short-term correlations (integration) which they suggested provided evidence of links between similar object features within a similar cortical area, and multiple attributes of one or more objects across several areas. In their models, integration emerged from co-operative effects within and among specialized areas. This led Sporns et al. (1989) to propose a solution to the binding problem, but several unsolved issues remained.

Having provided an overview of biological neurons as individual processing units and as part of neuronal networks it is possible to see that complexity allows one to describe functional connectivity which in turn can shape anatomical connectivity. The second part of this chapter is concerned with artificial neural networks (ANNs). Following a history of their development, the section describes the degree to which ANNs are modelled on the basis of their biological counterpart.

2.2 ARTIFICIAL NEURAL NETWORKS

Originally, the term 'neural networks' was used to describe networks of biological neurons that constitute the nervous systems of animals. Since the 1940s however, the term 'artificial neural network' (ANN) has been used for a technology of parallel computation in which the computing elements are 'artificial neurons' based on simple properties of biological neurons (Arbib, 1998). The 10^{11} neurons in the human brain communicate through an intricate connection network of axons and synapses. Typically, biological neurons perform a nonlinear operation on their combined inputs, and output the result through sub-cellular mechanisms and biochemical reactions which are important for the transmission of information via electric potentials from a single neuron to other cells. Individual neurons cluster in order to enable dynamic, interactive, and self-organizing information processing. Clusters of neurons are incorporated into larger structures which are able to perform even more complex behaviours. The connections between these neuronal structures, the summation and transfer functions comprise a functioning neural network (Anderson and McNeill, 1992). Higher order brain function emerges from interconnected networks of networks which interact with the external environment. The brain adapts to solve novel information processing demands without *a priori* algorithms and can cope with continually changing environments. Emergent properties from neural interaction distinguish neural computation from the connectionist paradigm of parallel distributed computing as proposed by McClelland and Rumelhart (1988).

Neural models are needed in order to contribute to our understanding of the processing performed by biological neural networks, in doing so Gisiger et al. (2000) emphasize that

'The main constraint one should try to impose on these (Anderson and McNeill, 1992) networks is their 'neuorealism'; in other words, one should ensure that they reproduce biological reality to a sufficient extent, and thus can be experimentally tested.'

However, because many computational abilities of biological neural networks do not seem to depend critically on specific details, general network architectures are frequently considered in ANN modelling. The simplest network would thus be a chain, but because a single pre-synaptic spike is not sufficient to elicit a post-synaptic spike, and neural death is common, a chain is not biologically plausible. Generally, an artificial neural network (ANN) uses a densely interconnected, parallel structure inspired by information processing in the mammalian brain. A typical ANN neuron (node) produces an output when the cumulative effect of the input stimuli exceeds a threshold value (e.g. Patterson, 1996). Key elements of ANNs are the large number of highly interconnected processing elements (neurons) which communicate through weighted connections (synapses). The weights act to increase or decrease (excite or inhibit) the input signals to the nodes and simulate the metabolic growth that takes place in neurons as a result of increased cell activity (Hebb, 1949).

The first models of biological neural networks were developed by McCulloch and Pitts (1943). Because their models could perform basic Boolean functions, they were considered representative models of brain function. However, the McCulloch and Pitts neuron was a simple two-state binary threshold type with both excitatory and inhibitory inputs (Fig. 2.2), and as such, is biologically implausible. Furthermore, the units approximated threshold firing characteristics of single neurons, but they did not include the generation or reset of spikes.

Chronologically, the next important development of ANNs was provided by Rosenblatt (1958) who showed that random connections such as those in the McCulloch and Pitts' model were implausible since so many links and connections probably would not be used in biological systems. Rather, Rosenblatt proposed the Perceptron which was based on retinal physiology. The Perceptron was a threshold logic unit comprising three layers: input, association and output, such that input was a 2-dimensional array of light sensors analogous to retinal

rods and cones. This array was partially connected to a layer of association units which was connected to response units. If the cumulative inputs to the association layer exceeded a threshold, the unit fired and passed on an impulse to the response layer.

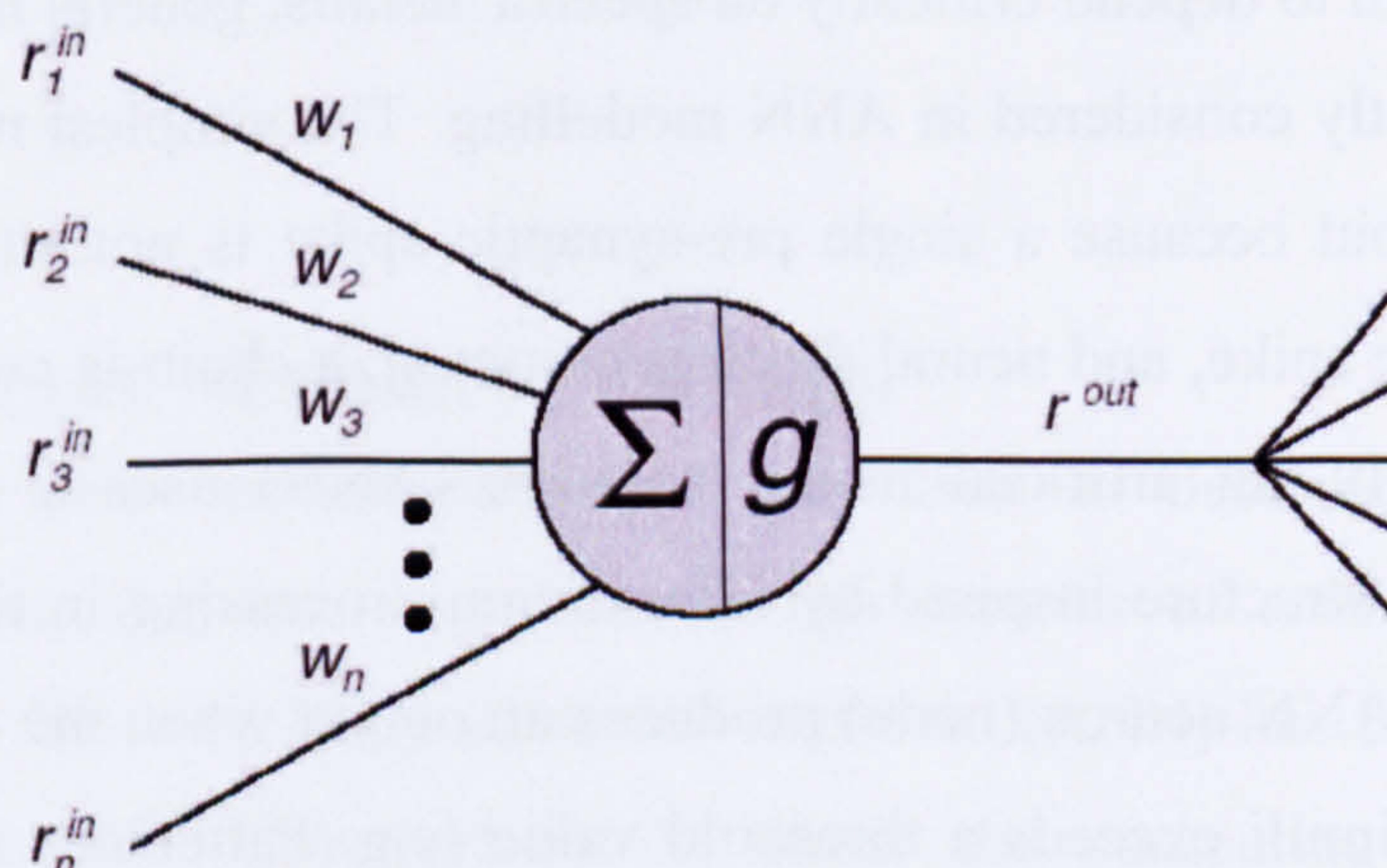


Figure 2.2 A typical **McCulloch and Pitts node** receives one or more inputs r^{in} , and produces one or more identical outputs r^{out} , each of which is a simple non-linear function of the sum of the inputs of each channel with the corresponding weight value (w_i) of that channel. The non-linear function is typically a threshold or step function which is usually smoothed (i.e. a sigmoid) to facilitate learning. Adapted from Trappenberg (2002).

Earlier, Hebb's (1949) publication *Organization of the Brain* had demonstrated that connections between neurons increase in efficacy in proportion to the degree of correlation between pre- and post-synaptic activity such that

"When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased." Hebb (1949), p. 62.

The Hebb rule states that groups of neurons which tend to fire together form a cell assembly whose activity can persist after the triggering event and which serves to represent it. Hebbian learning adjusts the network's weights such that its output reflects its familiarity with an input. The original Hebbian learning rule resulted in continued strengthening of the weights without bound. Accordingly, when Hebb's ideas were simulated at the Dartmouth Conference in 1956, the rule was generalized to prevent unbounded growth by including inhibition and weight

normalization, as well as ‘fatigue’ to make immediate subsequent firing less likely. The Hebbian learning rule has been influential in neural modelling, but modified versions are now more commonly applied.

The most basic components of neural networks are modelled on the structures of the brain. The basic units of ANNs simulate the four basic functions of natural neurons, but artificial neurons are much simpler than their biological counterpart. Real neurons have many structural, functional, physiological and anatomical variations. Despite this, all natural neurons have in common the same four basic components: dendrites, soma, axon, and synapses. Dendrites are extensions of the soma which receive input through the synapses of other neurons. The incoming signals are processed over time in the soma and transformed into an output which is sent to other neurons through the axon and the synapses (see Section 2.1). In neural models, these four components are represented by input and output nodes, connection weights and a transform function. Despite the commonality of some characteristics such as input and output nodes, a transform function and a pattern of connectivity, major differences exist among ANNs. For example, they can be single or multi-layer and can operate synchronously by updating nodes simultaneously, or asynchronously in which the sequence and timing of updating is random or governed by a probability distribution. Furthermore, they can operate by feed forward, back propagation or recurrent connectivity.

Although feed forward multi-layer networks are the leading paradigm for ANNs from an engineering viewpoint (e.g. Patterson, 1996), these models lack transparency of interpretation and biological plausibility. Furthermore, because these models exhibit learning which is neither Hebbian nor local, they cannot represent working memory. Sustained activity needed to deal with working memory is often represented in networks with attractor states¹⁰ of the network dynamics such as those developed by Hopfield (1982), Amit and co-workers (1989) and Rolls and Treves (1998).

¹⁰ Attractors describe the long-term behaviour of a dynamical system. In dissipative dynamical systems the presence of internal friction tends to contract elements in the phase-space so they approach a subset of the phase-space, an attractor, as the elapsed time grows large.

2.2.1 Integrate and fire model neurons

In order to contribute to our deeper understanding of brain processes, neural models should incorporate biological properties such as the pulse structure of neuronal output, the integrated signal flow through the synapses, axon, dendrites and soma, and noise. A single neuron in isolation has a fixed membrane potential threshold for spike generation (Mainen and Sejnowski, 1995) and is thus deterministic. Conversely, neurons in networks elicit variable spike trains with random temporal patterns, and are thus stochastic. Intrinsic noise sources which could account for this difference are synaptic transmission using neurotransmitters, and the excitation-inhibition balance.

Neural code may be deciphered by means of identifying reliable correlations between firing patterns and behavioural consequences, or between sensory stimuli and neural activity. Adrian (1929) showed that the number of spikes increases with an effective stimulus for that neuron. This increase in firing rate usually occurs a short time after stimuli presentation. Information processing in the brain includes a variety of firing rates. However, neurons convey information through spikes which have a stereotypical form. Therefore, although spikes are not likely to be the only form of information processing and transmission (Trappenberg, 2002), the influence of a presynaptic spike on a postsynaptic membrane potential (often described as an α function) is essential for processing. In the Hodgkin-Huxley (1952) model of this process, dynamics were driven by sodium and leakage channels. Such model neurons are known as ‘leaky integrators’ described by equation 2.3. Models with leaky integrate and fire neurons are known as ‘spiking models’.

$$\tau_m \frac{du(t)}{dt} = -u(t) + RI(t) \quad (2.3)$$

where τ_m is a membrane time constant determined by average conductances of the sodium and leakage channels, and $I(t)$, the input current is the sum of synaptic currents generated by firings of presynaptic cells. Simple spiking neuron models such as integrate-and-fire nodes (Tuckwell, 1988) are typically used in representations of neural coding, memory, or network dynamics (Maass

and Bishop, 1998; Gerstner and Kistler, 2002) because they can predict the mean rate of pyramidal cells recorded in in-vitro experiments over a range of time dependent inputs (Rauch et al., 2003); and can reproduce the experimental distributions of sub-threshold membrane potentials (Destexhe et al., 1994). Furthermore, such neurons are used where the precise form of the spike is not of interest, but the occurrence of the spike is.

Leaky integrate-and-fire (LIF) neurons behave by weighting and summing input spikes from presynaptic neurons to yield a time dependent membrane potential which decays if no spikes are received. As in biological neurons, if the membrane potential is excited sufficiently to exceed a certain threshold, a spike is emitted. The firing time of the postsynaptic neuron is defined by the time the membrane potential reaches the threshold. After emitting a spike, the neuron is unable to spike again for a certain period which corresponds to the biological refractory period, after which the membrane potential needs to be reset. An absolute refractory time can be incorporated by holding the reset value for a finite period. However, neural processing is more complex than weighting and summing input spikes to yield a time dependent membrane potential.

Neurons operate in noisy environments. Noise can be incorporated into neural models by adding noise to the threshold, by resetting the membrane potential to a stochastic reset potential, or by integrating noisy inputs (Trappenberg, 2002). Adding noise can also result in spike variability which occurs naturally in neural activity. Spike variability reflects input fluctuation which might occur if the dynamic operating point of a neuron is kept near threshold by excitatory and inhibitory inputs (Shalden and Newsome, 1994) rather than unreliability in spike initiation (Mainen and Sejnowski, 1995; Nowak and Bullier, 1997). Spike variability can be increased by partial or noisy reset after a spike emission which increases the gain in the relationship between input current and output spike frequency. For low spike train variability, the gain has a sharp transition and the neuron's firing rate soon reaches near maximum. This is determined by the inverse of the absolute refractory time for means of the input current exceeding the firing threshold. Hence the 'effective' threshold imposed on the IF model is where the strong increase of the firing rate starts. Low variability in spike train firing produces non-linear responses which allow

complex information processing. With increasing variance, the strong non-linear response is linearized.

The time a spike occurs can be used to represent external time-varying stimuli and external states. In fact, precise timing could be used to encode additional information about the sensory stimulus; to carry information about the global significance of the stimulus (Gray and Singer, 1989); to organize information in packets (Jeffreys et al., 1996); or for learning (Stopfer et al., 1997). In fact, by changing the relative timing of presynaptic and postsynaptic spikes in a neuron by as little as 10 ms can determine whether a synapse is potentiated or depressed (Markram et al., 1997).

2.2.2 *Transfer functions*

In the simplest neural model, each input is multiplied by a connection weight, summed, fed through a transfer (activation or gain) function to generate a result, and then output. The transfer function can be linear, step, threshold linear, sigmoid or sigma-pi or a combination of these. The simplest function is linear. It relates the sum of the inputs directly to the output. The step function returns two values, producing binary responses. The threshold linear function is bounded from below so that node activities are limited as in biological neurons, and is therefore a good approximation of functions in integrate-and-fire neurons (see Section 2.2.1). The sigmoid node, which can be discrete or continuous, bounds minimal and maximal response of a node, and is the most frequently used in modelling neuronal networks. To ensure the input does not grow without bound, a leaky integrator, a forgetting factor, allows the node to lose some of its current state over time. The forgetting factor acts on the same time scale as the integration of new input (Eq. 2.4)

$$\tau \frac{h(t + \Delta t) - h(t)}{\Delta t} = h(t) + \sum_i w_i r_i \quad (2.4)$$

In order to make the node continuous the time steps need to get progressively smaller so that $\Delta t \rightarrow 0$ and the differences become differentials. Thus for a continuous node, the Eq. 2.4 can be written as

$$\tau \frac{dh(t)}{dt} = -h(t) + \sum_i w_i r_i \quad (2.5)$$

Sigma nodes are generally used as the basic units in rate models. Despite being a rough abstraction, they are useful for modelling the population response of a network of neurons with similar properties (Trappenberg, 2002).

Information-processing capabilities of networks do not generally depend critically on the precise form of the activation function. In fact, several types of activation function can lead to similar network abilities (Trappenberg, 2002). Without external input, leaky integrator dynamics decay exponentially on a time scale τ towards a resting activation. With a positive net input, the decay slows or activation increases; with a negative net input, activation decreases or slows down recovery from sub-resting to resting level. However, without incorporating a ‘forgetting factor’, the net input grows without bound while the input is applied. Incorporating a leaky integrator that loses some of its current state over time, remedies this because the ‘forgetting factor’ acts on the same time scale as the integration of new input, but the node is still discrete.

2.2.3 *Coding information: rate or temporal spike coding?*

The irregularity of both firing and interspike intervals (ISIs) has led to speculations about the nature of neural code (Abeles, 1991; Konig and Engel, 1995) and the role of redundancy. For example, Shadlen and Newsome (1998) proposed that the high variability of ISIs of cortical neurons implied signal redundancy, whereas Panzeri and Schultz (2001) suggested that redundancy was not necessary for rapid and reliable transmission.

Neurophysiological recordings of firing rates of single cells, taken over several trials to account for noise, give an average temporal spike rate. Although such averages correlate well with behavioural responses, the methods of obtaining them are not biologically plausible, because an organism needs to respond to stimuli rapidly. Rate-based coding addresses this issue by taking the average firing rate of neuronal subpopulations with similar properties. However, if rate-based coding is used, neural responses may not adequately disambiguate information when the features have to be integrated into a coherent object while being segregated from features belonging to other objects. A proposed solution

would be to code for a representation of every possible feature conjunction, such as 'one's grandmother' (Barlow, 1972) but this would lead to a combinatory explosion of representations. To overcome this problem, von der Malsburg (1973; 1981) proposed that information might be encoded by the relative time of firing of different neurons and suggested that temporal synchronization of one subset of neurons, and separate synchronization of another subset might allow the separate activity of these subsets to be kept apart. In such a scenario decoding might be performed through neurons with sensitivity to co-occurrence of inputs.

The notion of temporal synchronisation was supported by Abeles (1982) who suggested that cortical neurons act as coincidence detectors, and coincidence, not firing rate, is the appropriate code for higher cortical functions. Later, Abeles (1991) proposed that synchronous synaptic inputs were more effective than asynchronous ones on finding that groups of neurons occasionally emit series of action potentials, 'synfire chains', which have a precise temporal relationship. Furthermore, synchronous temporal firing of neuronal populations has been proposed as a way of increasing information transfer (e.g. Engel et al., 1992; Eckhorn, 1994; Singer, 1999), visual grouping and segregation (e.g. Eckhorn et al., 1988; Gray et al., 1989), binding sensory and motor responses (Murthy and Fetz, 1992), binding features (Engel et al., 1991b), and binding different cell assemblies (Abeles 1991). However, as a solution to the binding problem, synchronous firing leaves many unanswered questions such as, the relatively long desynchronization time in the initial phase in which even weakly coupled assemblies will eventually synchronize (Grossberg and Grunewalde, 1997), and findings that stable synchronization of periodic neural oscillators cannot be used for associative coding when multiple patterns are active, a synchronously firing neuronal ensemble signifies a single event only (Ritz and Sejnowski, 2000).

2.2.4 Local, fully distributed or sparse models?

Representations in the brain can be modelled using local, fully distributed or more realistically, sparsely distributed models. In local representation one node represents a stimulus (e.g. Barlow, 1972). In a fully distributed model, the stimulus is encoded by the combination of all components that represent the stimulus. The sparse distribution uses a fraction of nodes to represent a stimulus.

Thus the proportion of cells highly activated at any one time is the sparseness of firing, and is a primary determinant of the capacity for memory storage. For non-binary units, such as real neurons, a generalized measure of sparseness of activity can be given as

$$a = \frac{\langle r_i \rangle^2}{\langle (r_i)^2 \rangle} \quad (2.6)$$

where r_i is the firing rate. Sparseness a can be defined by the average relative firing rate taken over the number of neurons in the population and the number of stimuli. Sparsely distributed representations (e.g. $a = 0.1$) used by the brain (Miyashita, 1988) can reduce interference occurring between stored memories (Trappenberg, 2002). Sparseness does not affect the interrelationships of neuronal representations, but the sparser a set of representations, the less the representational capacity and the larger the memory capacity of that population (Rolls and Treves, 1998). Possible activity changes in representations that develop with time can be examined by recording from the same populations, over periods during which some behaviourally relevant phenomenon of the existing representations may have occurred.

2.2.5 Learning rules

In memory retrieval, neurons that code for more than one representation become activated, and to avoid representational interference due to cross-talk (Schacter, 1989), binding may be involved in memory encoding (Raffone and van Leeuwen, 2001). Non-stationary synchronized activity may enable the integration of local features into global representations (van Leeuwen et al., 1997), and is more likely used for perceptual and working memory cortical coding than static or phase-locked neural oscillators (Fujii et al., 1996). Systems exhibiting such non-linear dynamic behaviours could be used for feature detection (van Leeuwen et al., 2000), perceptual segmentation and switching (van Leeuwen et al., 1997), memory retention and forgetting (2001). Although synchronous firing has been observed in and across several brain areas, the role it plays in information coding and decoding remains inconclusive (van Rullen and Thorpe, 2001). Van Rullen and Thorpe suggested spike asynchrony rather than

synchrony could be applied to fast information transfer from retina to cortex by means of temporal coding based on the relative order of the first spike in which greater impact is assigned to spikes with the shortest latencies. They showed that recognition could occur when only 1-2% of neurons had fired a single spike. However, their model neurons were not plausible in that they incorporated neither an adaptation mechanism (Smirnakis et al., 1997) nor lateral interconnections (Nirenberg and Latham, 1998), which have been shown to influence the response of retinal ganglion cells.

Despite much evidence (e.g. Optican and Richmond, 1987) implicating temporal encoding in the extraction of information, Rolls and Tovee (1994) and Panzeri et al. (1999a) suggested that temporal encoding within the spike train of a single neuron does not add much information to that which is present in the firing rate, even when short epochs are taken. In addition, Shadlen and Newsome (1998) found that 'synchronous spikes do not represent anything extraordinary' (p.3890). Rolls and Deco (2002) proposed that encoding information using firing rates is very powerful and likely to be more important than temporal encoding because information is rapidly available from an ensemble of cortical neurons. Neurophysiological evidence from single unit recording (Werner and Mountcastle, 1965) and event related potential (Thorpe et al., 1996) studies demonstrates that the majority of information is transmitted within 20-50 ms. Nevertheless, Rolls and Deco (2002) suggest that temporal encoding may be used as an additional mechanism to rate coding, as for example, a possible solution to the binding problem. Furthermore, Riehle et al., (1997) found that synchronization without firing rate modulation in monkey primary motor cortex was related to internal events, while synchronization with rate modulation co-occurred in processing external, behaviourally relevant events.

The concept of spike and rate code co-existing (Raffone and Wolters, 2001) is further supported by reports of sensory neural information transmission which typically suggest that most information is transmitted within one mean ISI, and that single spikes carry a great deal of information (e.g. Rieke et al., 1997). One may thus conclude that the brain may use different coding strategies depending on functional requirements.

Evidence outlined in this section suggests that the extent to which firing rates and correlations between different cells contribute to information

representation and processing by neuronal assemblies remains controversial. Although rate-based models do not consider the temporal aspects of neuronal firing, and therefore could miss critical information potentially available in the timing of the spikes (Gerstner and Kistler, 2002), they depend on the activity of pre- and post-synaptic nodes such as that described by Hebbian plasticity rules. Nevertheless, Hebbian learning can be implemented into rate or spike based models.

‘Rate models’ (e.g. McCulloch and Pitts, 1943) correspond to a population response in a network of neurons and use simple binary units which generate an output only when the summed input reaches a defined threshold. The minimum unit of a rate model is a receiving unit (node) which performs local processing and subsequently distributes its output to other units. The most basic unit is a sigma node (Figure 2.2.6.1) which can be used to model the average firing rate of a population of neurons with similar response properties.

2.2.6 Learning in artificial neural networks

Models in computational neuroscience typically use a local (Hebb-like) learning rule in which the signals, needed to alter the synaptic strength and lead to synaptic modification, are present simultaneously in the pre-and post-synaptic neurons. Humans associate patterns so that two stimuli occurring around the same time will become associated. Subsequently the second is retrieved when the first is presented. Pattern association such as this is implemented by the architecture of the visual cortex by means of back projections which permit top-down influences such as attention, recall and visual short-term memory (VSTM), as well as bottom-up information processing.

The brain learns from experience in order to adapt and survive in dynamic environments. However, adaptation requires considerable computational resources. Hence in designing adaptable ANNs, efficient learning algorithms are developed to try to increase learning efficiency whilst controlling computational resources. Some classes of ANNs might be considered as machine learning algorithms, because altering connection weights during training, causes the solution to be learned. Other models learn by adjusting threshold values of node activation functions, or through a combination of training algorithms.

One strength of ANNs is their ability to generalize to facilitate classification, memory and storage. ANNs generalize when they recall full patterns from partial or noisy input, when they recognize or classify objects they have not been trained on, or when they predict new outcomes based on past behaviours. Forming generalized concepts from a number of specific instances or examples is known as inductive learning. Generalization in conventional ANN systems is accomplished through the formation of classes of objects in which the universe of objects is partitioned into groups, one of which is the target group. Accurate learning occurs when the final class boundaries contain only the desired target examples and exclude all others. However, overtraining can occur if a training set is used extensively, or when too many nodes are used. In such a scenario, the network becomes over-specialized and poor generalization results.

Artificial neural networks, like their biological counterpart perform robustly because knowledge is distributed and there exists some degree of redundancy. Information is transmitted by means of connections between distributed areas. The strength of connection between the neurons is stored as a weight-value for that connection, and the system learns by adjusting these connection weights. Hence learning ability is determined by the architecture and the specific training method which can be supervised, reinforcement or unsupervised. Supervised ANNs, such as back propagation models, use each example pattern together with the desired output pattern to train the network. Comparisons are made between the actual and desired outputs and weights are adjusted through feedback to improve performance and achieve an acceptable level of classification for each new pattern not in the training set. Although highly successful in training multilayered neural nets and popular for engineering problems, back propagation is not biologically plausible because a neural signal cannot be transmitted backwards by an axon. In reinforcement learning models, the correct output is not presented to the network, instead a reward in the form of reinforcing the weights on the nodes that give the correct answer and reducing the weights on the others, is given. This results in improved network performance. Learning in unsupervised networks can be considered Hebbian because it occurs with no 'teacher' or reward to provide feedback on the desired output. Learning can take place through strengthening weights to match

prototypical input patterns as demonstrated for example in models by Kohonen (1982), Fukushima and Miyake (1982) and Grossberg (1988).

Learning rules can be described bottom-up from neurophysiological data, or top-down from the computational goal. The former allows biological detail to be incorporated, the latter allows the network to be understood in terms of global behaviour and might be useful when network behaviour cannot be predicted from single neuron or synaptic events. Biological neurons and synapses display a rich range of time-dependent processes as a function of the temporal patterns of activity in the network. Thus, both the firing frequency of connected neurons and the modulatory environment determine the intrinsic and synaptic properties that produce behaviour (Marder, 1998b). An essential part of a learning rule is the dependence on the co-activation of pre- and post-synaptic neurons (Hebb, 1949). Hebbian rules can be described as correlation-based learning rules in that modifications in synaptic transmission efficacy are driven by the correlated firing activity of the pre- and post-synaptic neuron.

Hebbian learning has been found to change the dynamic behaviour of a synapse, and not only the average amplitude of its responses (Tsodyks and Markram, 1996). This change makes the postsynaptic neuron more sensitive to transients so that synaptic plasticity may alter the sensitivity of a neuron to different neural codes. Synaptic efficacy changes with activity, and plasticity occurs across many time scales. An increase in synaptic efficacy such as facilitation, decays after about 10-100 ms post-stimulus, while long-term potentiation (LTP) persists for hours, days or longer. Plasticity can differ not only in time scale, but also in the conditions required to induce it. Some plasticity changes are independent of the post-synaptic response and depend only on the history of presynaptic stimulation, whereas LTP depends on a conjunction of pre- and post-synaptic activity as suggested by Hebb (1949).

Synaptic plasticity allows associations in the brain. A cortical neuron receives input from 5000 to 10000 synapses from neighbouring and distant cortical and subcortical areas. A small subset of synaptic channels can trigger a spike if the efficiencies are sufficient and the synaptic events fall within a certain time window. Other circumstances such as firing history which determines refractory time and shunting inhibition are also important. Input might not always be complete due to noise or partial sensory information. To model such

cases, when the signal is not strong enough to elicit a spike, weights need to be modified by a learning rule.

When Hebbian learning is implemented into rate based models the activity of neuron i is given by its firing rate r_i which is related to the membrane potential u_i by an non-linear monotonically increasing function g so that

$$r_i = g(u_i). \quad (2.7)$$

Hence the membrane potential can be calculated from the pre-synaptic firing rates r_j and the synaptic weights w_{ij} to give

$$u_i = \sum_j w_{ij} r_j. \quad (2.8)$$

Spike-based Hebbian learning can be implemented in a network of LIF neurons which comprise a resistor R and capacitor C driven by an external current I_i . The voltage across R and C is the membrane potential u_i which is set at $u_i = 0$ when at rest. Hence the temporal evolution of u_i can be given as

$$\tau_m \frac{du_i}{dt} = -u_i + RI_i(t) \quad (2.9)$$

where $\tau_m = RC$ is the membrane time constant of the neuron. Immediately after firing, the membrane potential u_i is set to a value $u_{reset} < \vartheta$ where ϑ is the threshold value. The input I_i to neuron i is due to spikes from presynaptic neurons j and each pre-synaptic spike arrival evokes a post-synaptic current with time course t . The total membrane potential is therefore the sum of all post-synaptic potentials caused by pre-synaptic firing plus the refractory effect of a negative resting potential. Firing time is updated when the membrane potential exceeds the threshold. Thus, whereas the neuronal state of rate-neurons is characterized by firing rate r , spiking neurons are characterized by their membrane potential u .

Electrophysiological recordings from several neurons simultaneously have shown that synaptic weight changes depend on the exact timing between pre- and post-synaptic spikes (e.g. Bi and Poo, 1998). Markram et al. (1997) found that potentiation or depression of synapses connecting excitatory neurons only

occurs if both pre- and post-synaptic neurons fire within a time window of ~20 ms. If the pre-synaptic neuron fires first, potentiation will take place, and if the post-synaptic neuron fires first, depression takes place. Levy et al. (2001) found that clusters of distributed synchronously firing neurons can form in a single Hebbian assembly in which excitatory neurons were fully connected. Bienenstock (1995) found evidence for binding between cell assemblies that fired with the same overall period, but which possessed different numbers of cycles. He suggested that a rich repertoire of composite memories could be built through a mechanism for encoding various combinations of multiple memories with stronger couplings within an assembly and weaker connections among different assemblies.

2.2.7 *Unsupervised networks*

Unsupervised systems learn through discovery and adaptation to structured features in the input patterns. That is, by adapting to statistical regularities or pattern clustering from the input training patterns. Unsupervised networks are biologically plausible and are further discussed in the following sections.

Kohonen networks

The human brain is a highly complex biological system organized into several layers of distinct neurons and neuronal groups, which are further organized spatially into functionally specific areas consisting of large numbers of anatomically and physiologically similar neurons. Groups of neurons within each region respond to sensory input. The Self-Organizing Feature Map (SOFM) first developed by Kohonen (1982) is a simple model of such localized mapping represented in a competitive, self-organizing network which learns from the environment without a teacher (Fig. 2.2.7.1).

The SOFM can comprise a single or multi-dimensional group of neurons in which the input vector is fully connected to each unit through adaptable weight vectors. The unit with the vector closest to the input pattern responds maximally and drives the other units to zero, and is thus, the winner. The winning unit shares the learning with its neighbours so that neighbouring nodes align their weights in the same direction as the input pattern, whereas more distant nodes align their weights in the opposite direction. The SOFM algorithm

finds the most similar neuron to the input pattern then modifies that neuron and its neighbours to increase their similarity to the input by using competitive learning and self-organization over several training trials. The SOFM was inspired in part by the Adaptive Resonance Theory (ART) models developed by Grossberg and co-workers during the 1970s and 1980s which combine co-operation and competition in unsupervised networks (Section 2.2.5.2).

Competitive learning models in which neurons organize themselves without sample outputs being provided were first proposed by Hebb (1949). Hebbian learning rules adjust the network's weights such that its output reflects its familiarity with an input. All the units are initialized with unequal small weights.

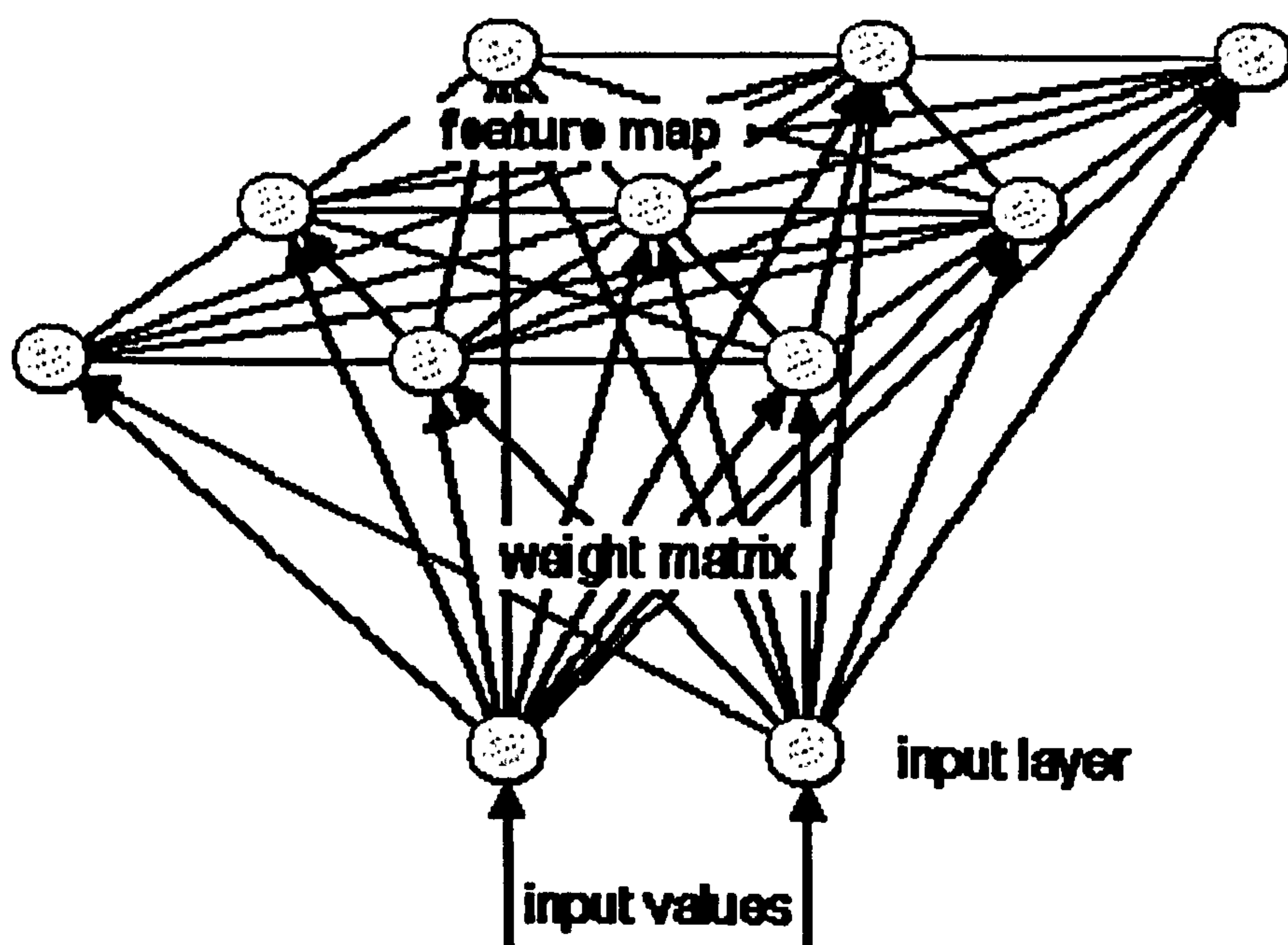


Figure 2.2.5.1 A sample structure of a Kohonen network in which each neuron of the input layer is connected to each neuron on the map. Source: <http://www.rfhs8012.fh-regensburg.de>

When an input pattern is presented, one unit in the layer responds more strongly than the others and has its weights reinforced or changed to more closely match the input pattern. Neurons in a single layer compete and there is only one winner. However, in some cases the weights of neighbouring units are also changed. Examples of competitive models are the ART networks developed in the Grossberg lab, and the single layer associative memory net developed by Hopfield (1982). These are described below.

Adaptive resonance theory (ART) networks

Form the mid 1970s to the late eighties Stephen Grossberg and co-workers at Boston University developed a series of increasingly complex and biologically plausible networks (Grossberg, 1987). Since then ART networks have been generalized and used in a broad range of applications. ART was developed as an extension to competitive and co-operative learning systems to overcome problems such as the stability-plasticity dilemma which occurs when a network needs to maintain stability while being plastic enough to adapt in a changing environment. ART networks use recurrent connections in which the output of a unit is fed back to form the input of all other units in that layer. These networks display several biological neural network characteristics such as time dependency, unsupervised learning and self-stabilizing memory. In ART models, weights are flexible, modified by past and new knowledge, and learning is a combination of competitive bottom up and cooperative top-down processes.

ART networks map n -dimensional input patterns to output categories or classes based on the input pattern's features. Similar patterns are grouped into the same class and dissimilar patterns into separate classes. The degree of similarity required for within class grouping is adjustable. Many class groupings of highly similar patterns are created when the threshold is set to a high level and fewer classes are created when the threshold is low. Learning in ART networks is continuous, unsupervised, adaptive, and occurs in real-time. A new category is formed when a novel input pattern is presented. New categories continue to be formed until the networks exhausts its pool of output category neurons at which time new patterns will be rejected. Input patterns matching existing categories initiate some degree of learning without upsetting the stability of the learned categories.

Hopfield networks

Hopfield networks (Hopfield, 1982) are single layer recurrent networks with symmetric weight matrices in which the all diagonal weights are set to zero, each neuron is connected every other neuron, and there is no differentiation between input and output neurons (Figure 2.2.5.3). The main application of a Hopfield Net is the storage and recognition of patterns. Hopfield networks store a number

of prototype patterns P as fixed-point attractors¹¹. The locations of the attractors are determined by the weight matrix W .

The stored patterns may be specified by Hebbian learning or through a gradient descent updating rule such as the Delta rule¹². A Hopfield network uses bidirectional lines between summation nodes, and weight values between -1 and 1 are randomly assigned and summed to zero. As learning proceeds, the weights change so that fewer paths are active and the network is more able to fully learn its pattern set. This 'energy function' tendency towards minimizing active pathways is measured by summing the output values of all the weights. Once the system has learned P prototype patterns they may be used for associative recall.

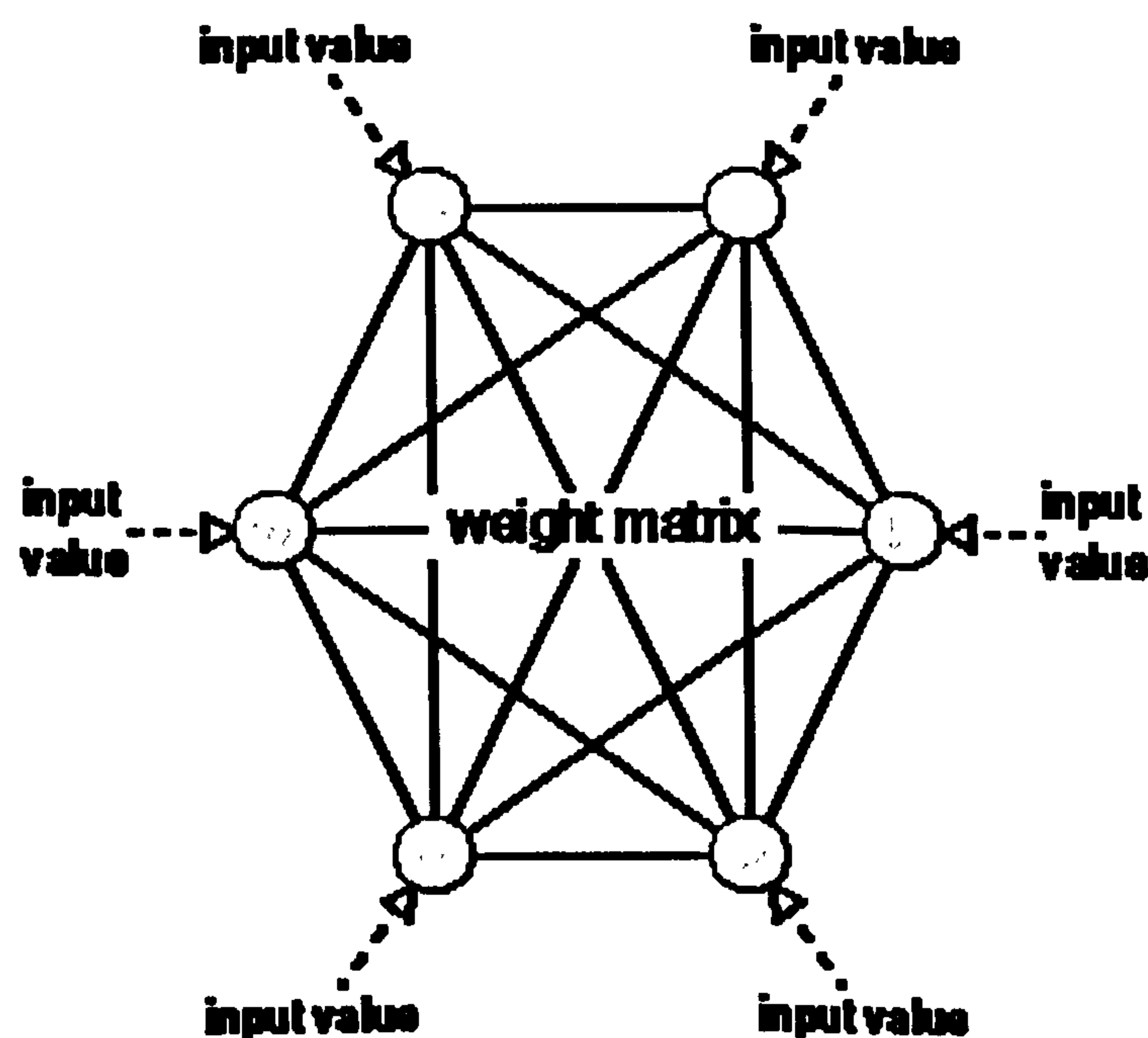


Figure 2.2.5.3 Hopfield's association network. Source:<http://rfhs8012.fh-regensburg.de>

To recall a particular pattern, the network operates recursively by feeding the output signals of the network back to the inputs repeatedly at each update time point t until the network stabilizes. The updating scheme maybe synchronous in which the outputs are computed simultaneously before the outputs are fed back to the inputs; asynchronous in which the outputs are computed sequentially in some order according to a probability distribution with

¹¹ Steady state (or equilibrium) behaviour corresponds to fixed-point attractors in which all trajectories starting from the appropriate basin-of-attraction eventually converge onto a single point.

¹² In the Delta rule weights are adjusted to minimize the squared errors over all patterns.

the outputs fed back to the inputs following each update; or a combination of the two. When combined synchronous and asynchronous updating is used, subgroups of units are updated synchronously and each group updated using an asynchronous updating scheme. During recall, a general Hopfield network can reach a cycle or a fixed point (attractor) state, and in order to function as an associative memory, the network should converge to some fixed point that is close to the input vector after a finite number of iterations. A continuous version of the Hopfield network (Hopfield, 1982) is a generalization of the discrete network described above. It uses analogue or continuous valued activation functions in place of binary functions.

Hopfield showed that as the ANN evolves, the energy must decrease or remain the same. As the number of possible states is finite, the network must eventually converge to a local minimum. This energy function is similar to the function that characterizes magnetic materials in physics using a simple Ising ‘spin’ model where the atoms of the material can assume one of two orientations, + or – (up or down). In the ‘spin’ model the weights correspond to the magnetic field influence between two neighbouring atoms. The energy minima correspond to fixed-point attractors which are the stored patterns. The state of the system at convergence determines the output pattern. Although Hopfield networks can be used as a content addressable memory in which full memory can be retrieved by providing partial information, not all associations are possible. In practice the number of different associations which can be learned and recalled is about 15% of the number of summation nodes. Increasing the size of the network increases learning capacity, but results in an exponential increase in the number of connections.

2.3 MODELS IN COMPUTATIONAL NEUROSCIENCE

In order to simplify, identify and test particular aspects of complex systems, we build models (Trappenberg, 2002). The exploratory nature of neuroscience means that finding the correct level of abstraction for model building, can be difficult. Several architectures have been proposed in the literature, but of interest to this thesis are those based on recurrent connectivity implemented in

associator networks. Relevant model designs in computational neuroscience are described in the following subsections.

2.3.1 Pattern associators

A pattern associator operates by simultaneously applying two stimuli: a first (conditioned) and a second (unconditioned) stimulus. The unconditioned stimulus is applied vertically on a 2-D grid through unmodifiable synapses to generate an external input to each neuron which is dominant in forcing the output neurons to fire. Simultaneously, the conditioned stimulus which consists of the set of firings on the horizontal axis is applied through modifiable synapses to the dendrites of the output neurons. Hebb's rule for synaptic modification is applied such that presynaptic firing on an input axon, paired during learning with postsynaptic dendritic activity, increases the strength between that axon and dendrite. The firing rate of every neuron is a value determined by the unconditioned input, and thus a function of dendritic activation. The change of synaptic weight results from the simultaneous presence of pre-and post-synaptic firing. After learning, the dendrites are activated through strengthened synapses so that they are activated when a cue is presented on the input axons. If the cue and the learned patterns are the same, the postsynaptic neurons will be activated even in the absence of the external input, because the strengthened synapses result in the firing axons producing activation of the postsynaptic dendrite. Thus those output neurons activated during learning are most strongly activated and the unconditioned stimulus can be recalled.

Biological neurons have thresholds which must be exceeded in order for the neuron to fire. Similarly, recall in pattern associator networks is most efficient when the threshold is exceeded only by strong activation. If the threshold is too low, erroneous memories could be recalled when many associations are stored. Some neural models use a binary threshold function to indicate that firing occurs when activation exceeds the threshold, but no firing occurs when activation is below the threshold. However, biological neurons have firing rates that saturate at maximum. This operation may be expressed formally as a sigmoid function (described above). Furthermore, an activation function which incorporates non-linearity is advantageous because it minimizes small activations elicited by interfering memories and enables neurons to perform

logical operations such as to respond only if two or more sets of inputs are simultaneously present.

If the post-synaptic neuron is strongly activated during learning, a pattern associator adds the vector that has the same pattern as the input vector to its weight vector so that recall depends on the similarity of the recall cue to the learned vector. In this way, pattern associators generalize because recall involves computing the dot product of the input pattern vector (which may be incomplete) with the synaptic weight vector to produce firing which reflects the similarity of the current input to the previously learned pattern. As well as the ability to generalize, pattern associators demonstrate other characteristics of biological systems such as graceful degradation. Because activation is the dot product, the system can have missing or damaged synapses and still produce perfect recall. However, both generalization and graceful degradation can only occur if the representation is distributed because recall depends on the dot product of the input pattern and weight vectors.

Recall in biological neural networks occurs within 10-20 ms. because the brain performs parallel computations. In pattern associators, a single pairing of the conditioned and the unconditioned stimuli produces the output firing which enables the association to be learned. In the brain co-occurrence is required for as little as 100 ms. and synaptic modification is present within a few seconds. Long-term potentiation (LTP) or synaptic strengthening occurs only in synapses from active afferents. This synaptic specificity is important for the correct operation of pattern associators. In biological neurons, most of the firing produced is available along the length of the dendrite so that any active synapse onto it can be modified and the cell learns an association between the total pattern of activity on its axons and its post-synaptic activation. The storage capacity of pattern associators needs to be analyzed in parallel with retrieval quality which deteriorates when too many associations are stored, or there is too little input. In contrast, autoassociators (Section 2.3.2) can store a finite number of memories beyond which retrieval of anything is impossible. Autoassociators are described in the following subsection.

2.3.2 Autoassociator (attractor) networks

Hebb's ideas in *The Organization of Behaviour* (1949) laid the foundations for the concept of memories as attractor states. Hebb described how the plasticity rule could be used to form cell assemblies. Because these assemblies could sustain activity even in the absence of the triggering stimulus, activity patterns within such assemblies could form the neural substrate of on-line or working memory. The associative memory implemented by such networks has the ability to recognize a previously stored pattern from partial input, a high degree of fault tolerance with respect to a partial loss of the network, and the efficiency of learning from single examples. Recurrent networks of associative nodes can be used to model associative memory, where the memory states correspond to the point attractors imprinted by Hebbian learning in these dynamical systems. These types of networks are called autoassociator, or attractor, neural networks. Abeles et al., (1995) proposed that cortical neural networks are mostly in attractor states with distinct firing rates and neuronal interaction. When stimuli are presented associatively, synaptic plasticity allows associative memories to form. In associative models, the first (or known) stimulus maintains its weight, and the second stimulus changes its weight as it becomes associated with the first. The main biological implementations of Hebbian learning are long-term potentiation (LTP) and long-term depression (LTD) which take place in NMDA-rich synapses (Rolls and Treves, 1998) and properties of non-NMDA neurotransmitter receptors (Heidmann and Changeux, 1982). LTP and LTD changes in synaptic plasticity can last for minutes or even weeks.

Attractor neural networks are able to store and retrieve memories as patterns of neural activity, and can form memory states (Hopfield, 1982) manifested as self-sustaining selective neural activity which corresponds with neurophysiological data (e.g. Miyashita, 1998; Fuster, 1995). Hopfield (1982) introduced the concept of attractor neural networks in which a stimulus, once presented to the network elicits a configuration of activity specific to that stimulus which is learned via Hebbian synaptic modifications. These modifications enable the neural assembly to sustain an active representation, a 'memory state', of the stimulus in its absence, and through superimposed synaptic weight changes, memories can be recalled from partial cues (e.g. Rolls and Treves, 1998). Furthermore, such networks can complete pattern on

presentation of a similar or partial one (Kesner and Rolls, 2001). Self-sustaining and highly irregular (Renart et al., 2001; van der Velde and de Kamps, 2003) selective neural activity exhibited by attractor networks (e.g. Amit, 1989) corresponds with single unit recordings of cells in monkey prefrontal (Fuster, 1995) and inferotemporal cortex (Miyashita, 1988) performing DMS tasks. Such sustained firing could be generated by a balance of total excitation and inhibition afferents to cortical cells (Shalden and Newsome, 1994), sufficiently strong recurrent synaptic excitation (Compte et al., 2000), and a predominance of NMDA receptors (Goldman-Rakic, 1995).

Autoassociators can recall the entire memory when presented with a fragment of it. Typically, the external input produces firing activity which is applied as a recall cue on the recurrent collaterals. This activity is multiplied with the synaptic weight vector on each neuron stored during learning to produce a new activation. The new activation reflects the similarity between the recall cue and a stored pattern. This enables the firing of the output vector to be associated with itself during learning. Hence during recall when part of the external stimulus is presented, some output neurons are forced to fire and additionally through recurrent collateral axons and modified synapses, other neurons can be activated.

Activations after thresholding result in a firing activity which is more similar to one of the stored patterns than the firing activity was at the first iteration, because that was firing from the recall cue alone. Autoassociators store associations between elements in a pattern, with each element in a pattern vector stored as the firing of a neuron. Therefore a memory is a set of pattern vectors. The network operates to recall one of the patterns from a fragment of it. Kohonen (1977) called these memories ‘autocorrelation memories’ because they learn correlations between the activity of neurons in the network. Each pattern is thus associated with itself by Hebbian-like learning.

Internal recall in attractor networks involves the multiplication of the activity firing vector by the vector of synaptic weights. If the learned patterns are distributed, the similarity of this firing vector to previously stored firing vectors can be provided by the output. In this way, attractor networks perform pattern completion because the recalled memory is closest in pattern similarity to the partial cue. Autoassociative recall can be perfect because recall improves with

each iteration. It is also fast because the pattern can be applied to the synapses and the axons simultaneously, and activation can be accumulated in one or two time constants of the dendrite. However, although pattern completion implemented an autoassociative Hopfield net typically requires only 5-15 iterations for pattern completion (Rolls and Deco, 2002), which corresponds to 50-200 ms. in the brain, it is still too slow to be biologically plausible.

Rolls and Treves (1998) showed that if neurons are modelled as integrate-and-fire neurons (see Section 2.5) in real-time, rather than being updated at every iteration as in a McCulloch and Pitts model, the network relaxes into its recall state more rapidly. One contributing factor for this is that in a network of integrate-and-fire neurons, some of the neurons are already close to threshold before the recall cue is applied. In an attractor network modelled with integrate-and-fire neurons, learning is 'one-shot' because a single presentation enables the network to learn the association of the activation of the dendrites with the firing of the recurrent collateral axons. The simplest and biologically plausible learning rule is a variation of Hebb's rule in which the information required to specify the change in synaptic weight is available locally at the synapse because it is dependent only on the pre-synaptic firing rate (available at the synaptic terminal) and the post-synaptic activation (available on the dendrite of the receiving neuron).

Biological plausibility in the artificial network can be increased by using a threshold linear activation function so that continuously variable firing rates can be modelled (Rolls and Tovee, 1995), and if the representation is sparse rather than fully distributed, a small proportion of the neurons fires above the spontaneous rate as in hippocampus (Treves and Rolls, 1991), and response is faster. Furthermore, by using diluted connectivity (Treves and Rolls, 1991) in which some or many connections are missing, the network matrix is non-symmetric (i.e. w_{ij} does not equal w_{ji}). Asymmetry is useful for storing sequences of patterns (Hopfield, 1982) because an asymmetric component can associate one pattern with the next in sequence to the standard connection weight which associates a pattern with itself. However the asymmetric (heteroassociative) component needs to operate on a slower time scale than the autoassociative component (Kleinfeld, 1986) because the two time scales allow

the latter to stabilize before the former moves the network to the next pattern. Both components result from a Hebbian rule which increases the weight whenever post-synaptic activity is paired with pre-synaptic activity occurring within a given time: the autoassociative components exhibit simultaneous conjunctions of pre- and post-synaptic activity; the heteroassociative component can be implemented by Hebbian learning of each conjunction of post-synaptic activity with pre-synaptic activity shifted a time step in the past.

In order for a system to maintain a steady state it needs to self-regulate through negative and positive feedback mechanisms. Amit and co-workers (Amit et al., 1994) developed attractor network systems which incorporated recurrent inhibition. In each memory state, the strong recurrent inhibition resulted in only a small subset of neurons firing at more elevated frequencies, and in the absence of external stimulation, the network stabilized in a state of low spontaneous activity. Amit et al. (1994) found that when two stimuli were frequently shown in sequence, synaptic modifications occurred so that when one stimulus was shown, neurons selective to the other also tended to be activated and the two independent memory states became correlated. These correlations have been measured in inferotemporal (Miyashita, 1988) and prefrontal cortex (Fuster, 1995) in monkeys during delayed response tasks.

In the cortex, short-term memory (STM) may be implemented by recurrent excitatory collateral connections between neighbouring pyramidal cells with a probability of approximately 0.1 (Abeles, 1991). Autoassociative memory can be used as a STM in which iterative processing keeps a representation active until the next input cue is received. Memories are stored as patterns of neural activity in the recurrent synaptic connections between neurons in the network. Recurrent connectivity keeps a representation active by continued firing even in the absence of external input. Hence this type of network can operate as a short-term memory (STM). Such continued firing mechanisms are seen in inferotemporal and prefrontal cortex (Williams and Goldman-Rakic, 1995). Such firing may be maintained by the operation of associatively modified recurrent collateral connections between nearby pyramidal cells producing attractor states in autoassociative networks (Rolls and Treves, 1998) (Figure 2.3.2).

An autoassociation network

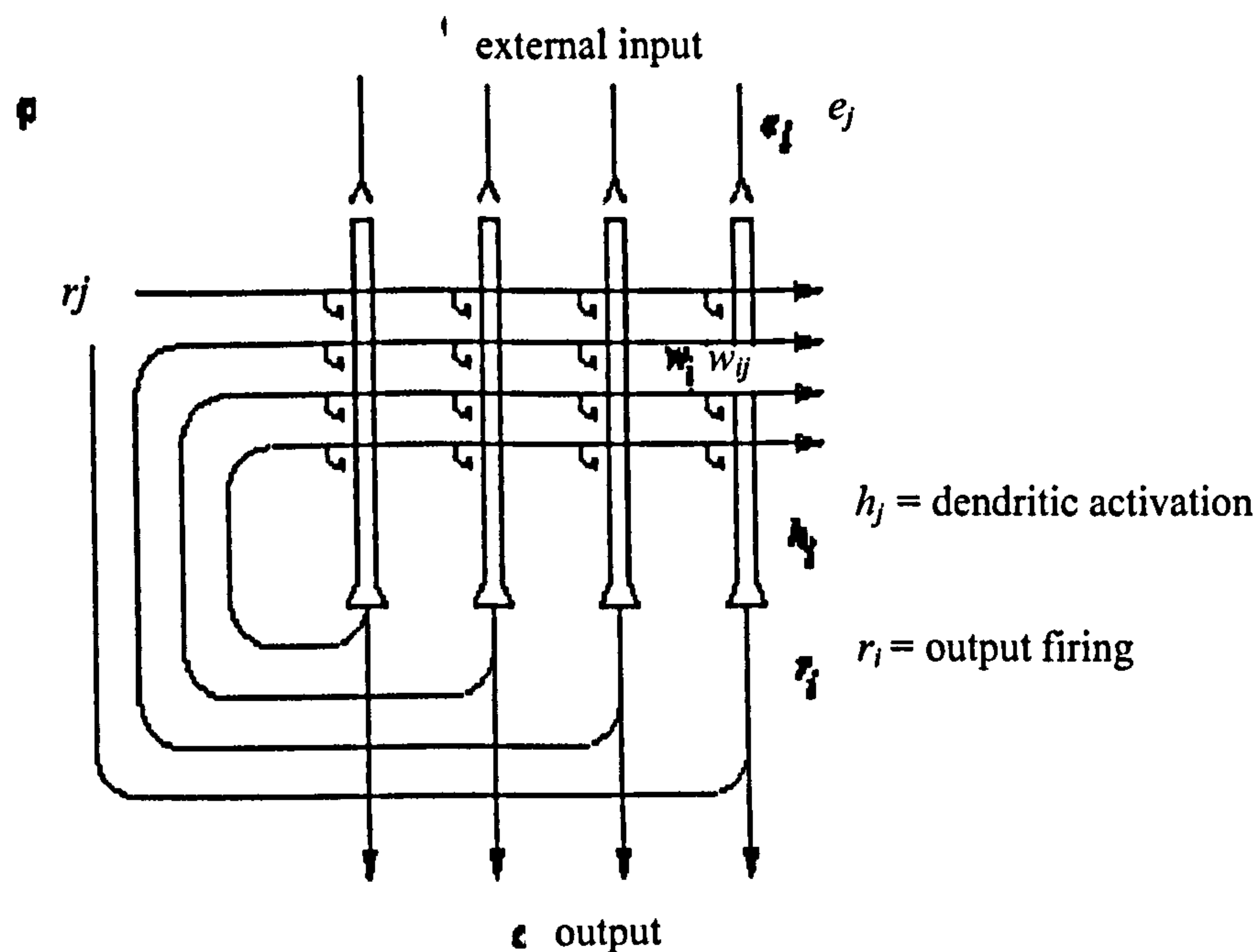


Figure 2.3.2 An autoassociation network. Source: Rolls (2003).

However, as firing rates are positive, an interfering correlation can occur between different pattern vectors, but this can be removed by subtracting the mean of the presynaptic activity from each presynaptic term. In this way, the rule increases the synaptic weight if the firing exceeds the threshold (equivalent to long-term potentiation), and decreases the synaptic weight if it does not. Biological plausibility can be increased in an attractor network by using a threshold linear activation function so that continuously variable firing rates can be modelled (Rolls and Tovee, 1995). Furthermore, if the representation is sparse rather than fully distributed, a small proportion of the neurons fires above the spontaneous rate, as seen in hippocampus (Treves and Rolls, 1994), and response is faster.

Diluted connectivity (Treves and Rolls, 1991) in which some or many connections are missing, makes the network matrix asymmetric. Asymmetry in a network (i.e., w_{ij} does not equal w_{ji}) is useful for storing sequences of patterns (Hopfield, 1982) because an asymmetric component can associate one pattern with the next in sequence to the standard connection weight which associates a pattern with itself. However, in order to perform well, the new component needs to operate on a slower time scale than the autoassociative component (Kleinfeld,

1986) because the two time scales allow the autoassociative component to stabilize before the heteroassociative component moves the network in to the next pattern. The heteroassociative retrieval cue is the previous pattern for the next pattern in the sequence.

Autoassociative components reflect Hebbian learning, whereas a heteroassociative component can be implemented by Hebbian learning of each conjunction of post-synaptic activity with pre-synaptic activity shifted a time step in the past. Thus both components result from a Hebbian rule which increases the weight whenever post-synaptic activity is paired with pre-synaptic activity occurring within a given time which can range from a few hundred milliseconds to simultaneity. Although this is similar to a trace rule which is consistent with observed conditions that induce LTP, the learning rule for sequences is more complex. The time shifted conjunctions of activity encoded in the weights must produce retrieval activations that are also time shifted. To do this the synaptic weights need to keep separate traces of simultaneous and time-shifted activity. No evidence of this in the brain has been found to date.

In the standard stochastic attractor model, the addition of noise can be helpful (Amit, 1989). Noise destabilises spurious local minima and increases the probability that the network will end in a learnt pattern. As noise increases, the size of the basins of attraction, and the number of spurious attractors decreases (Davey and Adams, 2002).

2.3.3 Competitive learning networks

A competitive model is one in which individual units, grouped into clusters are organized into functional layers. Each unit in the same cluster receives the same inputs and inhibits the others to compete in a ‘winner-take-all’ competition. The unit receiving the largest input achieves its maximum output while others are driven to zero. In this way, the ‘winning’ unit learns through a redistribution of the weights in which the total for the cluster is held constant and weight values of the winning node are shifted towards the input pattern vector. An example of a competitive network which learns by self-organization is the Self-Organizing Feature Map (SOFM) first developed by Kohonen (1982) and described previously (Section 2.2.5.1). Competitive self-organizing models represent neurons that respond to correlated combinations of their inputs. These neurons

function as feature analyzers which perform useful pre-processing in many sensory systems.

Competitive nets learn to categorize¹³ input pattern vectors based on the similarities between them. Each category activates a different neuron or a set of neurons. Hence correlated (similar) vectors activate the same neuron. This is unsupervised because learning is based on similarities in the input space and there is no external 'teacher' to force the classification.

In the visual system, categorization reflects a set of correlated active inputs that occur together. A competitive network 'self-organizes' during learning to find these sets in input space and in doing so removes redundancy¹⁴ from input space by allocating output neurons to reflect the set of inputs that occur together, and produces sparse output vectors with no loss of information. This is an important part of processing in sensory systems (Barlow, 1989) because neurons in receiving systems operate with the limited number of inputs each neuron can receive. In addition to performing categorization, competitive networks can perform the opposite function of separating (orthogonalizing¹⁵) patterns so outputs are less correlated with each other.

The architecture of a competitive network is a one-layer network with a set of inputs that make modifiable excitatory synapses with the output neurons. The synaptic weights are initialized to random values before learning. Connectivity need not be fully distributed. In fact, diluted connectivity helps maintain stability by making neurons tend to find inputs to categorize in only certain parts of the input space.

Competitive networks exhibit biological plausibility for many reasons. Firstly, the most prolific type of cortical cell is pyramidal which is excitatory. Pyramidal cells connect to and activate inhibitory interneurons whose firing results in mutual inhibition as they in turn synapse with pyramidal cells using GABA as the inhibitory neurotransmitter. In a competitive network model, the

¹³ Competitive nets perform categorization which is the process of placing vectors into categories based on their similarity. This is in contrast to classification performed by pattern associators and MLFF networks in which outputs are placed in particular classes as instructed by a 'teacher'.

¹⁴ Redundancy can be defined as the difference between the maximum information content of the input data and its actual content.

¹⁵ The process of orthogonalization is applied to signals before they are used as inputs in associative networks trained with Hebbian rules to reduce interference between patterns stored in these memories

output cells also compete through mutual inhibition to produce the firing rate vector that produces a non-linear gain function, which is at its extreme, winner-take-all. Thus learning occurs through the application of a modified Hebbian rule. Furthermore, in order to prevent the same few neurons from always winning, the length of the synaptic weight vector is normalized¹⁶.

Normalization of the inputs is useful in helping different input patterns be equally effective in learning. In the brain, normalization might be achieved by a set of input axons to a neuron coming from another network in which the firing is controlled by inhibitory feedback. At an even lower level of description, (Buhl et al., 1994) has suggested that different classes of input to a neuron activate different types of inhibitory interneurons which terminate on different parts of the dendrite. This might allow separate feed forward inhibition for the different classes of input. The feed back inhibitory interneurons also have characteristic termination sites on or close to the cell body where they control firing of the neuron by shunting (divisive) inhibition rather than subtractive inhibition.

A biologically plausible way to implement the effect of normalization is by using a modified Hebbian rule that increases synaptic strength according to conjunctive pre- and post-synaptic activity and allows the strength of each synapse to decrease in proportion to the firing rate of the post-synaptic neuron and the existing synaptic strength. This results in heterosynaptic LTD, a decrease in synaptic strength for synapses from weakly active presynaptic neurons onto strongly active post-synaptic neurons, because the amount the synaptic strength decreases depends on the existent strength of the synapses. Thus this rule allows the sum of the synaptic weights on each dendrite to be similar without the need for explicit normalization of the synaptic strengths, and moves the direction of the weight vector towards the current input pattern vector in proportion to the difference between the two vectors and the firing rate of a neuron. However, if normalization is needed, Oja's rule (Oja, 1982), normalizes the synaptic weight vector and is still a local learning rule. Oja's rule augments a standard Hebbian rule with a multiplicative term that is proportional to the square of the postsynaptic firing rate.

¹⁶ A vector can be normalized so that its length is equal to 1 by dividing it by its length. The dot product of two normalized vectors is 1 at maximum and -1 at minimum.

Non-linearity in the learning rule can assist competition (Rolls, 1989b). For example, a non-linear function of the post-synaptic firing mimics the operation of NMDA receptors in learning. In the brain LTP occurs when strong activation of a neuron has produced sufficient depolarization for the voltage-dependent NMDA receptors to become unblocked, allowing Ca^{2+} to enter the cell (Section 2.2.4). Hence synaptic modification occurs only on neurons that are strongly activated. This assists competition in selecting only a few winners.

In a competitive network a winner can be selected by searching for the neuron with the maximum activation. In a winner-take-all network patterns within a cluster are categorized as the same and the within-pattern correlation becomes 1. With the greatest competition, only one neuron remains active and the representation is at its most sparse. Although sparse representation allows many different memories to be stored, it corresponds to the biologically implausible 'grandmother cell' hypothesis (Barlow, 1972) in which a single neuron responds to a particular object (Chapter 4). More realistic graded competition can be achieved by using an activation function that increases greater than linearly, or by using soft competition implemented by means of the normalized exponential or Softmax function (Bridle, 1990).

The Softmax function specifies that the firing rate of each neuron is an exponential function of the activation scaled by the whole vector of activations. Soft competition is implemented by the exponential function because after competition, the faster firing neurons are firing relatively much faster than the slower firing neurons. The strength of competition can be adjusted by implementing a 'temperature' $T > 0$ in which very low temperatures increase competition until with $T \rightarrow 0$ the network implements a 'winner-take-all'. Activations are mapped to the range 0 – 1, and the sum of firing rates is 1. Thus under certain conditions, the firing rates can be interpreted as probabilities. For example the competitive network firing rate of each neuron reflects the probability that the input vector is within the category signified by that output neuron. The use of a competitive network with continuous output neurons and soft competition allows the network to generalize more continuously to an input vector that lies between learned input vectors. Soft competition allows neurons with small activation to learn a little and move gradually to the presented

patterns. Hence all output neurons tend to become allocated to one input pattern or one cluster of input patterns.

Competitive networks can perform separation of similar pattern vectors through orthogonalization, and can help separate of non-linearly separable vectors if placed before an autoassociator. The sparsification that can be produced by competitive networks is useful when preparing patterns for presentation to an associator or autoassociator because it can increase the number of memories that can be associated or stored. In addition, competitive networks can form topological maps in which the closeness in the map reflects the similarity between features in the inputs. In order to form self-organizing maps, short-range excitatory connections and long-range inhibitory connections need to be added between the neurons. The effect is to encourage neurons that are close together to respond to the similar features, and distant neurons to respond to different features. When these tendencies are present during learning, the feature analyzers built by modifying the synapses from the input to the activated neurons, tend to be similar when they are near and different if distant. A common characteristic of the brain is short-range excitatory connectivity between neurons mediated by inhibitory interneurons. The density of excitatory connectivity declines as a function of distance, but some connections within an area might be long-range to connect with more distant neurons. This suggests topological maps could develop in the brain to minimize the total length of connections between neurons, and the complexity of their connectivity rules. These characteristics might underlie the separation of cortical visual information processing into different processing streams (Rolls and Deco, 2002).

2.3.4 Continuous attractor neural networks

Another class of recurrent network model which has specifically organized connectivity structure is the continuous attractor neural network (CANN) model. Such models are important in computational neuroscience because there is strong evidence to suggest that such mechanisms are used frequently for a variety of information processing mechanisms in the brain. CANNs have been studied in conjunction with many diverse brain functions including local cortical processing (Hansel and Sompolinsky, 1998), saccadic motor programming (Kopocz and

Schöner, 1995), working memory (Compte et al., 2000; Wang, 2001), and spatial representation (Skaggs et al., 1995; Zhang, 1995; Redish and Touretzky, 1996).

Whereas associative memory systems have point attractors surrounded by a basin of attraction and are useful for associative memory, a CANN is not useful as an associative memory because network state perturbations can trigger different attractor states. Nevertheless, CANNs can hold memories for a short period of time and can operate as short-term, temporary memory stores.

CANNs are characterized by local cooperation and global competition implemented by weights values that depend on the distance from nodes: near nodes are positive; distant are negative. A positive value between close nodes creates localized positive feedback loops, an 'activity packet', in which any initial node activity is reinforced between them (Trappenberg, 2003). Continuous attractor neural networks (CANNs) are able to maintain a localised packet of neural activity representing the current state of an agent in continuous space without external sensory input (Amari, 1977). CANNs use excitatory recurrent collateral connections between the neurons to reflect the distance between the neurons and the state space, and global inhibition to keep the number of neurons in an activity packet relatively constant. When activity packets are stationary, two packets in a single feature space are stable even if they are different sizes, providing they are distant (Amari, 1977). In addition, they can remain stable if they share active neurons, but are in different feature spaces. Stringer, Rolls et al. (2004) showed that with a bounded non-linear transfer function more than one active activity packet could become stable. The activity packet suppresses other network activity through long range inhibition. This balance between excitation and inhibition is the source of the formation of an activity packet. An important feature of an activity packet is that it maintains activity after the external input is removed. It is therefore an attractor state of the network and any activity that is symmetric around the centre of this attractor state will lead to the same asymptotic activity packet after removal of the external input. If network structures are perfectly shift invariant, an activity packet can be stabilized around each node of the network by applying an initial external input centred on the node on which the activity packet should be stored. The number of attractor states thus scales with the number of network nodes.

CANNs have a one layer structure from which the output is fed back as input to the network in addition to potential external input. Thus even with a constant input the network state can change within each execution, and the network is dynamical. The behaviour of the system depends strongly on the values of the connection weight parameters for which there are many choices. The learning algorithms are typically Hebbian. During learning, the activity is kept constant by the external input and the recurrent connections have no effect. During updating, the network state evolves until it reaches a fixed point where the node activities no longer change. Such fixed states can correspond to stored patterns in the network and can be retrieved from partial or noisy cues.

2.3.5 Modular networks

As demonstrated in some of the models described above, modular level networks can represent the large scale organisation of the brain. Modular networks can comprise a combination of network designs. In order to combine such representational networks, the functional and anatomical connectivity needs to be considered. Anatomical connectivity is difficult to establish, but a number of techniques such as staining, have been proposed. For the purpose of this study, we use existing evidence for anatomical connectivity between brain regions. Functional connectivity patterns can be established by using simultaneous recordings from different brain areas which show correlated firing patterns if they are functionally connected.

Modularisation allows a problem to be broken down into smaller problems each of which can be solved independently. However, the problem of recombining these sub-solutions is itself not trivial. For example, FF networks can be combined to represent an architecture known as the ‘mixture of experts’ in which a column of parallel ‘expert’ modules receive the same input. A gating module sums the output from the expert modules in addition to receiving some of the input. The output is then combined by an integration module which can be represented by a sigma-pi node. Recurrent networks can also be combined but still remain a single system with subsystems.

The previous two sections have discussed biological and artificial neurons and networks. Where possible, differences between the two have been described within the sections. Further differences exist and these are synthesized

in the following section in order to highlight some of the problems facing computational neuroscience.

2.4 MAJOR DIFFERENCES BETWEEN BIOLOGICAL AND ANNs

Despite essentially continuous research and development since the introduction of computers in the late 1940s, the connection between brain processing and computer processing is still undergoing theoretical development. ANN researchers typically proceed by abstracting the essence of a theory aimed at explaining results obtained from neurophysiological or behavioural experiments and then derive a method for designing processors which can obtain the same or similar experimental results. However, to date, no completely satisfactory brain theory exists.

As described in Section 2.1, biological neurons are diverse in their type, topology and functionality. For example, there are functional distinctions in retinal neurons (Dowling, 1987), and morphological distinctions between simple and complex neurons in visual cortex (Gilbert and Wiesel, 1983). Pyramidal and spiny stellate neurons seem only to be excitatory, smooth stellate neurons are inhibitory (Shepherd, 1979) and there are electrophysiological distinctions between properties of Purkinje and granule neurons (Ito, 1984). In addition, structural and functional distinctions exist between neurotransmitter receptors for all of these neuron types. This diversity is poorly represented by the homogeneous processing elements of most ANNs.

Depending upon the ANN's size and architecture, a processing element may have a few or thousands of inputs, whereas a typical neuron may have anywhere from a 1,000 to 10,000 inputs, but the relative number of inputs is not where the major differences lie. Biological neurons have more complex and variable connections to other neurons than do processing elements in ANNs. Many ANN models have units that connect to all others in that layer or in the adjacent layers. In the brain however, several areas such as the cerebral cortex have inputs from a horizontal layer that consists of a tight bundle of axons from other areas of the brain.

Many ANN models use biologically unrealistic learning algorithms, such as combined excitatory *and* inhibitory output units (Miall, 1989). Some elementary learning mechanisms can be formally implemented by algorithms expressing the regulation of synaptic strength as a function of experience such as a learning procedure which formally relies on the output of the network to back-propagate and make corrections to the strength of all the connections (Rolls and Treves, 1998). Hebb's more biologically plausible, local learning rule enhances and stabilizes frequently occurring activity while removing counter-productive circuits.

The transfer functions of processing elements in ANNs do not fully simulate the intricate workings of the synaptic integration of neurons. In an ANN, the operational rules governing the behaviour and the changes in connection strengths are frequently the same for all processing elements within the model. The simple mathematics of the artificial processing element's input values and relative weights typically does not match the workings of the synapse via the use of different neurotransmitters and multiple states of the receptors. Furthermore, the possible use of second messengers in regulating synaptic efficiency and the possibility that the cytoskeleton may be used in learning have no analogous counterparts in ANNs.

Biological neurons and synapses display a rich range of time-dependent processes (Marder, 1998b) as part of a complex non-linear dynamical system, but in most ANNs nodes are trivial units that simply multiply input with synaptic weight to produce an output that is a binary, linear, or nonlinear function of the input. In real networks, recent history may be at least as important as current inputs because most neurons require more than a single input spike to activate them from rest. The hidden parameters that regulate the dynamics of the biological synapse vary from synapse to synapse and at least some of these hidden parameters can be changed by LTP because synapses are history dependent (Maass and Zador, 1998). Moreover, each input has an effect on the neuron's membrane potential that greatly outlasts the duration of the input spike. Once a neuron is induced to spike by receiving sufficient excitatory inputs, it is refractory for a brief interval so that further input is less effective. In contrast, behaviour of most artificial processing elements can only affect its current behaviour by influencing the synaptic weights so that the output is independent

of previous states and can change instantly from one level to another (Miall, 1989). Synaptic strength and intrinsic properties in biological networks are continuously varying on numerous time scales as a function of the temporal patterns of activity in the network. Thus, both firing frequency of the neurons in a circuit, and the modulatory environment determine the intrinsic and synaptic properties that produce behaviour (Marder, 1998b).

2.5 SUMMARY

The aim of computational neuroscience is not to build an artificial brain, but to understand brain processes. In order to realize this, biological parameters and evidence from neurophysiological studies are taken into account when modelling. Given these constraints, and the desire to further understand the complex mechanisms and emergent behaviour of biological information processing systems, the appropriate level of abstraction can be applied so that modelling is both tractable and plausible. Beyond engineering applications, several ANN models have been constructed in an attempt to further understand brain function. However, although useful for some problem solutions, major differences between artificial nodes and neurons, and between networks of biological neurons and ANNs still remain. Notwithstanding substantial progress in many areas of neuroscience over the past decade, building realistic models that are able to perform specific cognitive functions that might lead to a further understanding of the operations and mechanisms involved in brain systems remains a major challenge.

This chapter has described the composition and function of single biological neurons and artificial nodes both as individual units and in networks. A brief outline of the development of ANNs, transform functions, integrate-and-fire models and learning algorithms was given. Then how these are implemented in artificial neural network architectures was described. Lastly, major differences between real and artificial neural network systems are described. The next chapter describes the visual system, visual working memory (VWM) and computational models of VWM relevant to this thesis.

CHAPTER 3

THE VISUAL SYSTEM AND VISUAL MEMORY

INTRODUCTION

The primate visual system, a complex collection of over thirty visual areas occupying approximately 50% of posterior cortex (Zeki, 1993), is typically conceptualised as a hierarchy of processing stages, each consisting different highly interconnected areas working in parallel to analyse different visual attributes (Felleman and Van Essen, 1991). These areas are organised into two anatomically distinct cortical streams of visual processing emanating from the primary visual cortex (V1): a dorsal, occipito-parietal stream, and a ventral, occipito-temporal stream (Ungerleider and Mishkin, 1982).

Behavioural (Ungerleider and Mishkin, 1982), electrophysiological (Haxby et al. 1991) and neurological (Giersch et al., 2000) studies provide evidence for the functional specificity of these streams: the dorsal stream processes spatial information, and the ventral stream is involved in the representation of visual objects. However, findings from neurophysiology (e.g. Milner and Goodale, 1995) suggest that the dorsal pathway controls visually guided action and often processes information implicitly, while the ventral pathway mediates conscious visual perception including that experienced during action (Figure 3).

Disparate and specific cortical areas and pathways understood to be involved in visual processing form part of an even more complex intricate network of networks within the brain. This highly complex system is further affected by the external environment in which it exists. How such a widely distributed system incorporating precise functional and anatomical specificity rapidly codes and decodes sensory information remains poorly understood despite a large literature and continued research. These issues beg the question:

what and where are the mechanisms that integrate and segregate neuronal activations and environmental stimuli to allow rapid and coherent percepts?

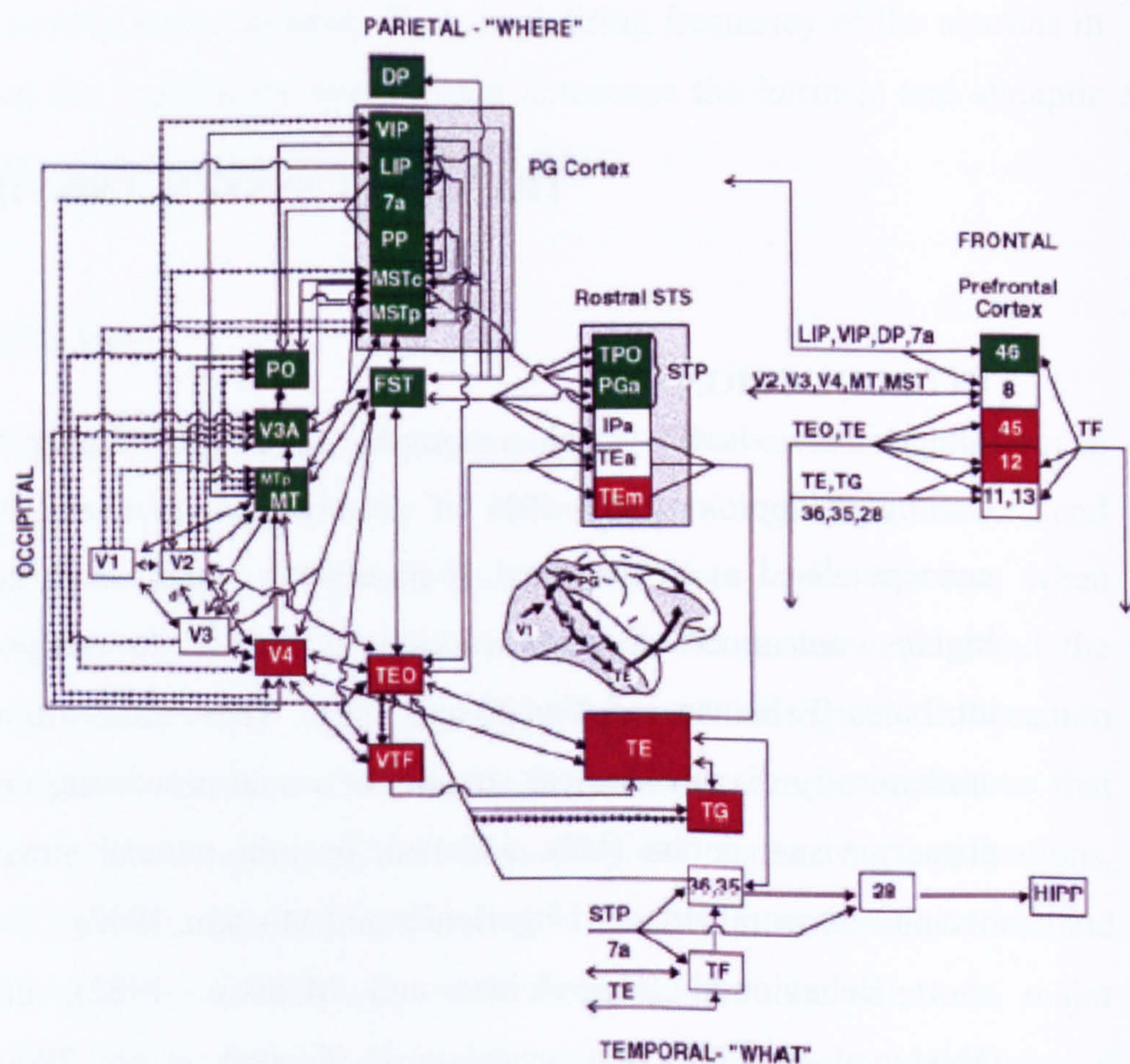


Figure 3 Visual processing pathways in monkeys. Areas in the dorsal stream have primarily visuospatial functions, and areas in the ventral stream have primary object recognition functions. Lines connecting the areas indicate known anatomical connections, with heavy arrowheads indicating feed-forward connections from lower-order areas to higher-order ones and open arrowheads indicating feedback connections from higher-order areas to lower-order ones. Solid lines indicate connections from both central and peripheral visual field representations, and dotted lines indicate connections restricted to peripheral field representations. Shaded region on lateral view of the brain indicates extent of cortex included in the diagram. Source: Ungerleider (1995).

The remainder of this chapter provides an overview of visual processing beginning with the basic features of vision: the retinae, the lateral geniculate nuclei, receptive fields, simple and complex cells, and the visual cortex. The remainder of the chapter describes the processes involved in visual working memory and the final part of the chapter provides an outline of relevant computational models of object feature binding.

3.1 BASIC FEATURES OF THE VISUAL SYSTEM

In order to begin understand the problem of object feature binding in visual working memory, basic details of the visual system are necessary.

3.1.1 The retina

Detailed vision occurs in the macula lutea, the central retinal area of the eye (Fig. 3.1.1), which corresponds to the central 5° of vision. The fovea, which corresponds to the central 1° of vision, is at its centre. The retina's photoreceptors (rods and cones) and ganglion cells (ganglia) transform light from objects into neural signals. Rods are widely distributed throughout the retina and have slow latency and dynamics. Although they are colour insensitive, rods are highly sensitive to low level brightness and are mainly activated in dark environments. In contrast, cones are concentrated in the fovea, react rapidly to high levels of brightness, and are essential for colour vision. Both rods and cones send output to horizontal and bipolar cells which respond with continuous graded activations over short distances. Retinal cones project to parvocellular cells, while both rods and cones project to magnocellular ganglia. Separation of visual information into streams begins with retinal ganglia (Shapley and Perry, 1986). The majority of ganglia are magnocellular and are concerned with motion analysis; the remainder, are parvocellular cells and are responsible for coding colour and fine detail. Axons from retinal ganglia form the optic nerve that transmits retinal information into the central nervous system.

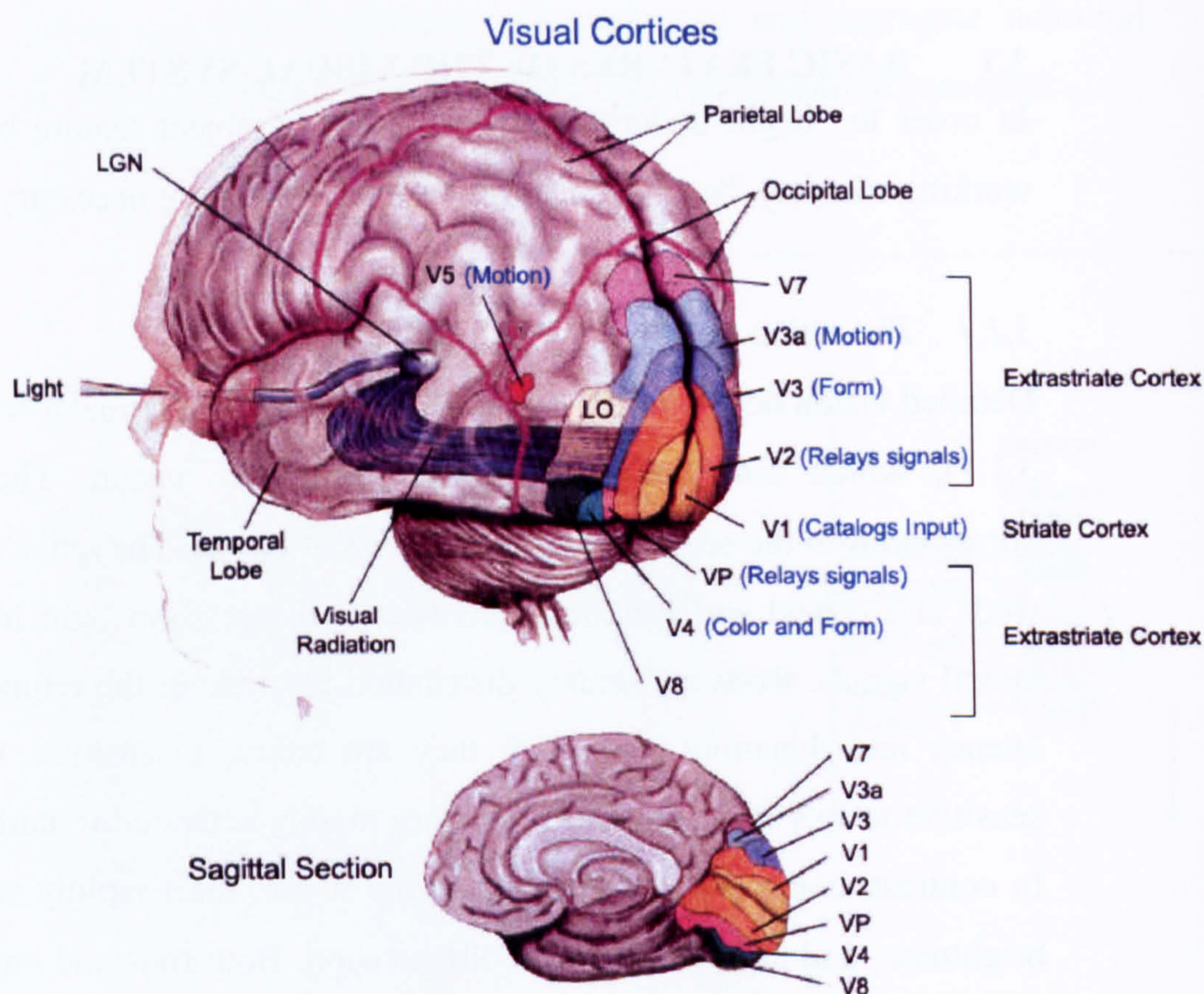


Figure 3.1.1 Visual cortices <http://www.colorado.edu/epob/epob3730rlynch/07vision.html>

3.1.2 The lateral geniculate nuclei (LGN)

The LGN are the relay nuclei between the optic fibres from each eye and V1. Half of the fibres in the optic nerve cross at the optic chiasm to send axons to the lateral geniculate nucleus (LGN) on the opposite side of the brain so that information from the right (left) visual field will be sent to the left (right) LGN and the left (right) visual cortex (Fig. 3.1.2). Each LGN receives inputs from and sends outputs to the posterior parietal cortex, an area associated with spatial perception.

Parvocellular cells project to the four upper (parvocellular) layers of the LGN, magnocellular cells project to the two lower (magnocellular) layers of LGN. The parvocellular stream deals with shape perception and colour encoding, and has sustained linear responses. The magnocellular stream deals with motion and has fast dynamics and poor response to sustained visual stimulation. Separation between these paths is maintained from the retina through the LGN, to striate (V1) and extrastriate (V2, V3, V4 and V5) areas. The magnocellular, dorsal stream projects from V1 to V2, middle temporal (MT or V5), and medial superior temporal (MST) The parvocellular, ventral stream projects from LGN to

V1, V2, V4, dorsal inferotemporal cortex (TEO), and anterior inferotemporal cortex (TE) (Figures 3 and 3.1.2). As information passes from one visual area to another, retinal topography is maintained, although there are some distortions. For example, at the end of the parvocellular path (Desimone et al., 1985) there is a disproportionate mapping of fovea in a large area of V1, and a smaller representation of retinal periphery.

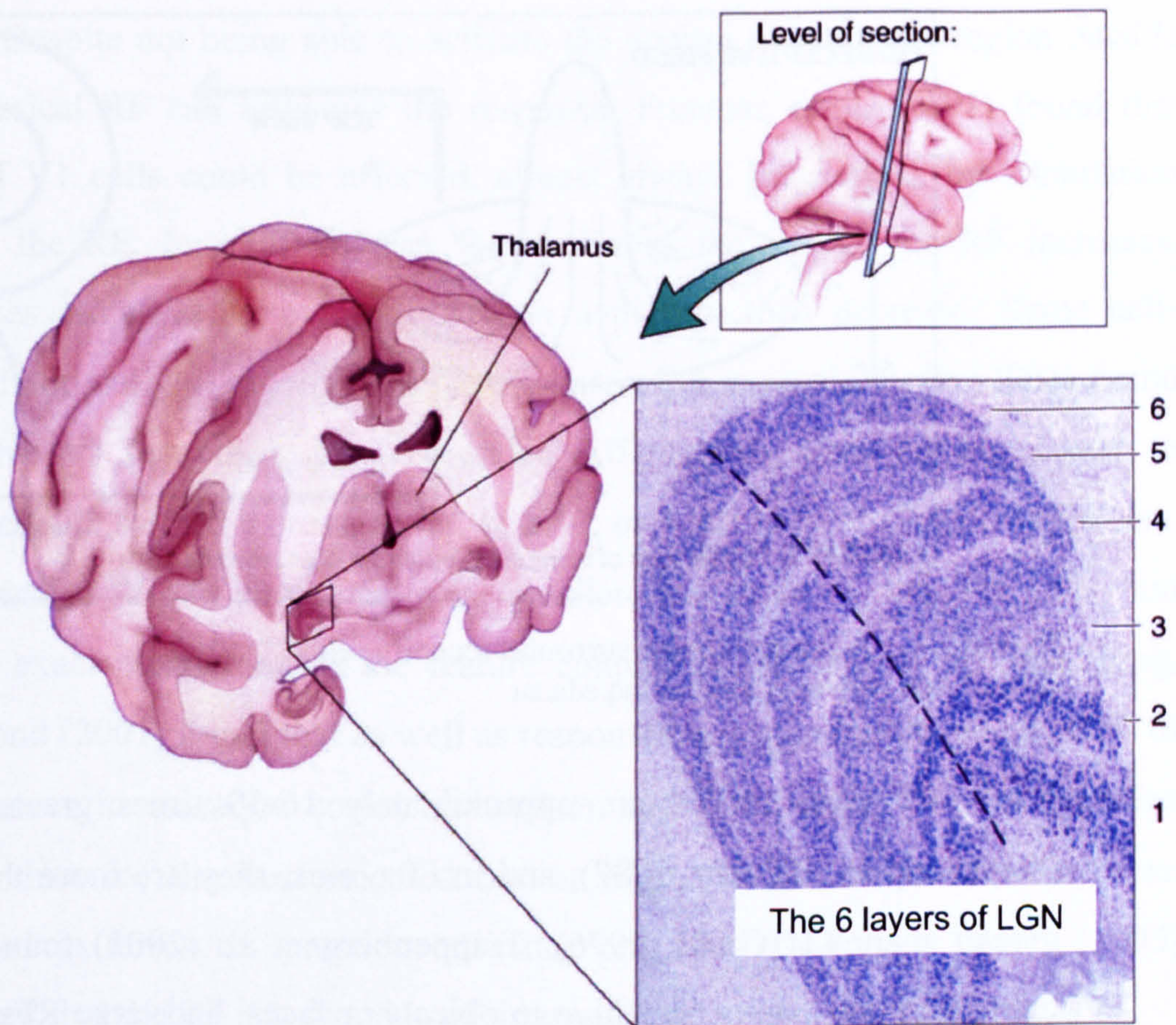


Figure 3.1.2 Layers of the lateral geniculate nucleus. Adapted from: <http://www.bioon.com>

3.1.3 Receptive fields (RF)

A receptive field (RF) is 'the area of visual space within which one can influence the discharge of a neuron' (Hartline, 1940). Early studies (Barlow, 1953; Kuffler, 1953) measured neuronal firing rate to determine an RF's preference. Findings showed retinal ganglia and LGN cells have roughly circular centre-surround RFs (Hubel and Wiesel, 1962) with centres responsive to bright stimuli and surrounds sensitive to dark (ON-centre) or the reverse (OFF-centre) (Fig. 3.1.3).

Hubel and Wiesel (1968) found that RF properties vary from area to area. Cone receptors in the retina have centre-surround organisation and respond optimally to spatial discontinuity in the image on the retina within their RF. LGN cells also have centre-surround organisation, and use lateral inhibition to produce contrast enhancement. In V1, cells respond to elementary features, whereas in 'higher' areas, cells are tuned to different aspects of complex stimuli (Maunsell and Newsome, 1987).

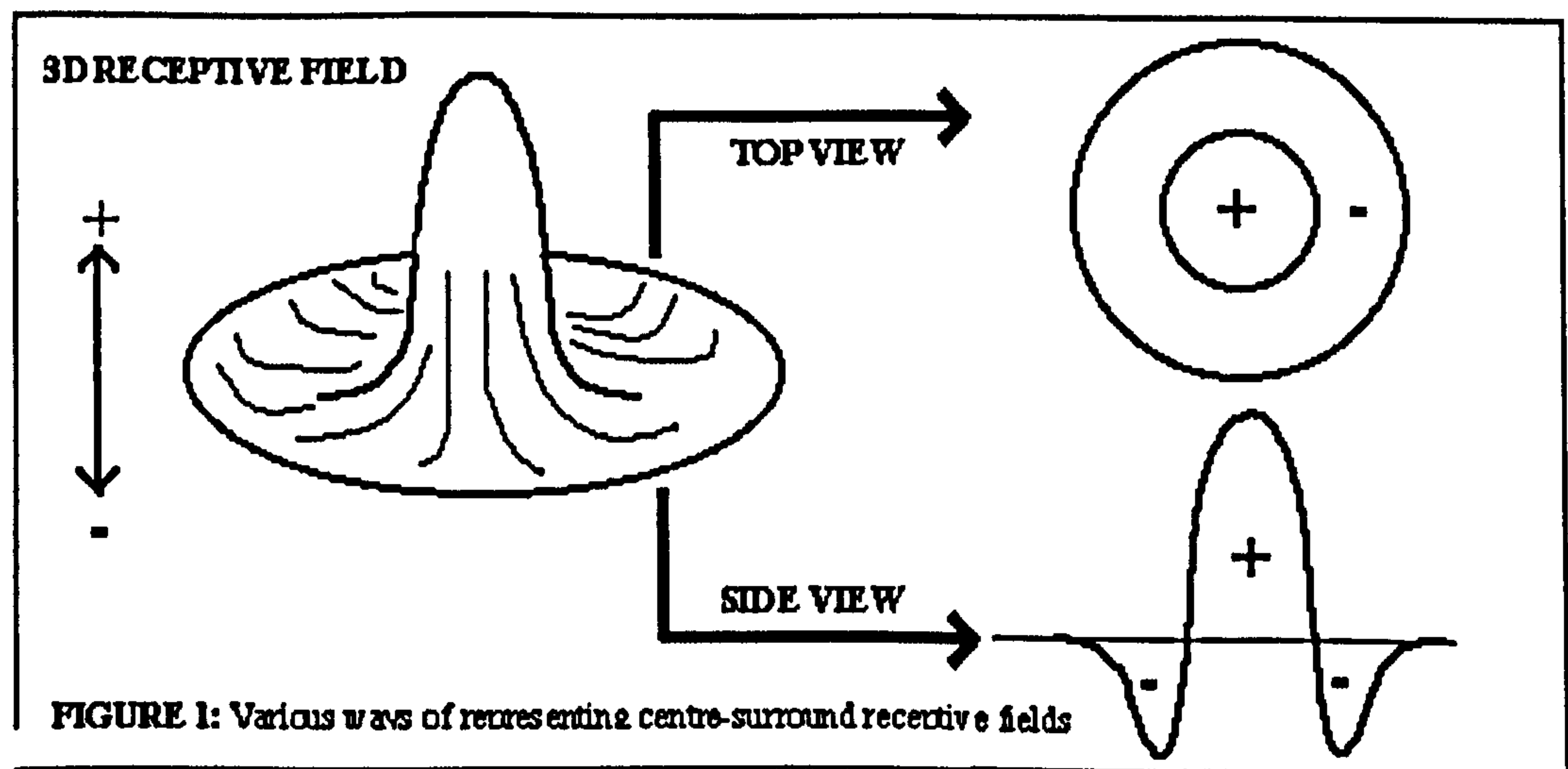


Figure 3.1.3 A typical centre-surround receptive field.
Source: <http://www.vision.bhs.mq.edu.au>

RF sizes in V4 are approximately 16-50 times greater than in V1 (Desimone and Schein, 1987), and in IT cortex, they are more than 1000 times greater than V1 (Gross, 1976). Trappenberg et al. (2001) found that cells in inferotemporal cortex, sensitive to objects or faces, had large RFs (~56 degrees) when viewing a single object on a blank screen, regardless of position on the retina. This would allow memory operations to be performed when the object was viewed from a different position. Interestingly, RFs were smaller (12 degrees) if an object was shown in a natural, rather than blank background, and if more than one object was present, RF response was greater in foveal than parafoveal neurons (Sato, 1989).

In addition to sensitivity to light, the response rate of RFs is determined by orientation and other features. For example, cells in parvocellular pathways are 'simple cells' that are orientation specific. They have centre-surround elongated RFs that respond optimally to a bar or edge (Hubel and Wiesel, 1959).

In contrast, cells in the magnocellular pathway are 'complex'. Complex cells respond to changes in light, do not have definite 'on' and 'off' regions (Hubel and Wiesel, 1962), but respond to oriented edges in their RFs, frequently tuned to the direction of motion (DeValois and DeValois, 1988). This early visual processing makes information explicit about the location of characteristic features such as elongated edges, orientation and colour, as well as removing redundancy from the visual input by not responding to areas of uniform brightness (Marr, 1982; Barlow, 1989).

Despite not being able to activate the neuron directly, the region outside the classical RF can influence the response. Freeman et al. (2001) found that 56% of V1 cells could be affected, almost always suppressed, by stimulation outside the RF. In addition, they found that as the size of the RF increases, responses increase monotonically, reach a plateau then decrease. Some cells respond optimally to specific complex features. For example, Rolls (2001) found that some V1 neurons in cats responded differentially when presented with two lines that formed a cross, an 'L', or a 'T', or when presented with single bars. Furthermore, RFs of simple cells are smaller than those of complex cells, and require exact positioning of the feature conjunction to achieve optimal firing. Richmond (2001) found that as well as responses to single bars or edges, 80% of V1 neurons in macaque cortex showed an increased firing rate for complex stimuli than for the best oriented bar or edge. Traditionally, RFs had been described in terms of spatial coordinates only, but DeAngelis et al., (1995) suggested that the RF is a function of time and space.

Humans can simultaneously and precisely perceive fine detail and localize visual signals despite the fact that much of the visual cortex is involved in identifying individual objects independent of their features, and other regions are dedicated to localizing objects in space independent of identity. The mechanisms that facilitate such perception are still not fully understood. The following paragraphs provide a brief outline to information processing in visual cortex in order to appreciate the complexity of the visual system.

3.1.4 The primary (striate) visual cortex (V1)

From the eye to the LGN, processing proceeds to the (V1) which occupies approximately 11-12% of the neocortex (Felleman and Van Essen, 1991) and has

a topographical representation (Figure 3.1.4a). V1 contains both ‘simple’ and ‘complex’ cells which, in addition to the specific preferences stated earlier, cells may be ‘end-stopped’ (require an oriented bar or edge to terminate in their RF).

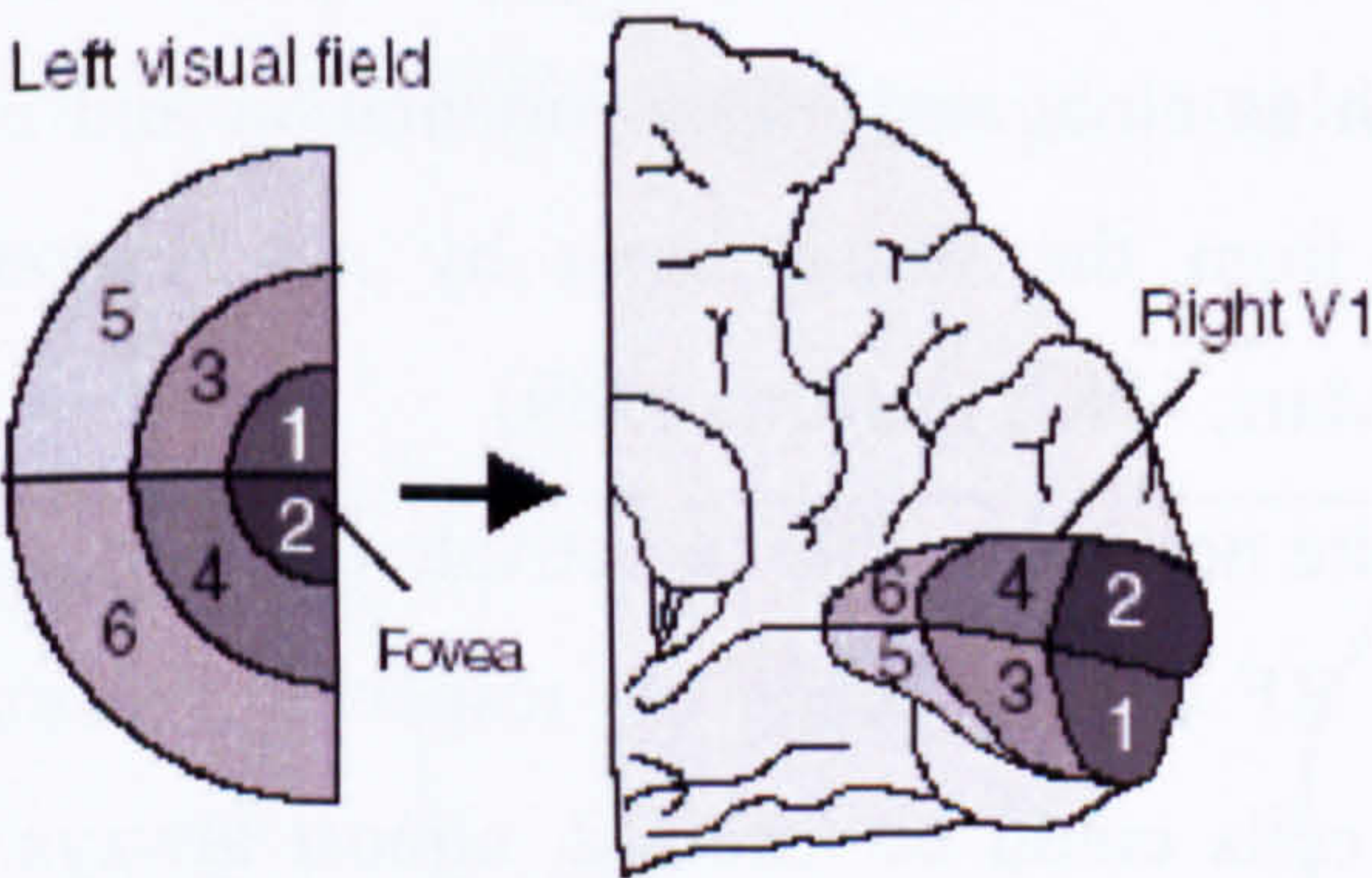


Figure 3.1.4a Topographic representation of the visual field in V1. Source: Trappenberg (2002)

Spatial information for object localization projects mainly to the parietal lobe, whereas spatial information for object recognition projects more to the temporal lobe (Mishkin et al., 1983) (Figure 3.1.4b).

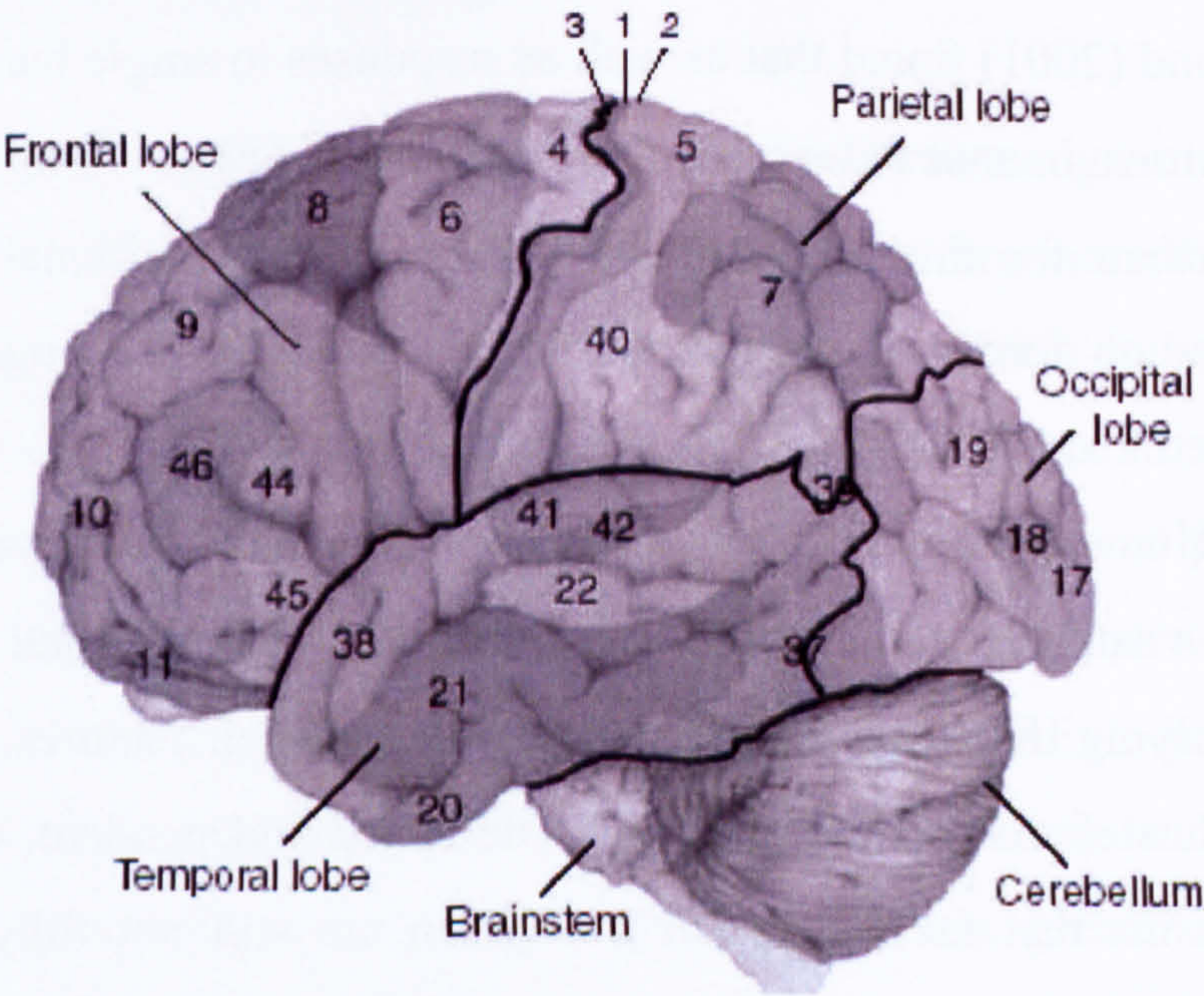


Figure 3.1.4b A lateral view of a human brain including neocortex, cerebellum, and brain stem. The numbers correspond to Brodmann’s areas. Source: Trappenberg (2002)

The RFs of inferotemporal neurons are not sensitive to spatial position, but show only a coarse localization for individual objects compared to the fine spatial

representation provided by neurons in V1 and V2. Because the spatial frequency of V1 cells is higher than V2 cells, an integrated representation of spatial detail and retinotopic localization probably occurs in ensembles of simple cells in V1. V1 neurons respond to different features such as position, size, orientation, colour or motion as well as ocular dominance (LeVey et al., 1975). The visual (striate) cortex has a laminar structure that supports interactions among local circuits that form functional columns of cortical maps (Hubel and Wiesel, 1977) (Figure 3.1.4c).

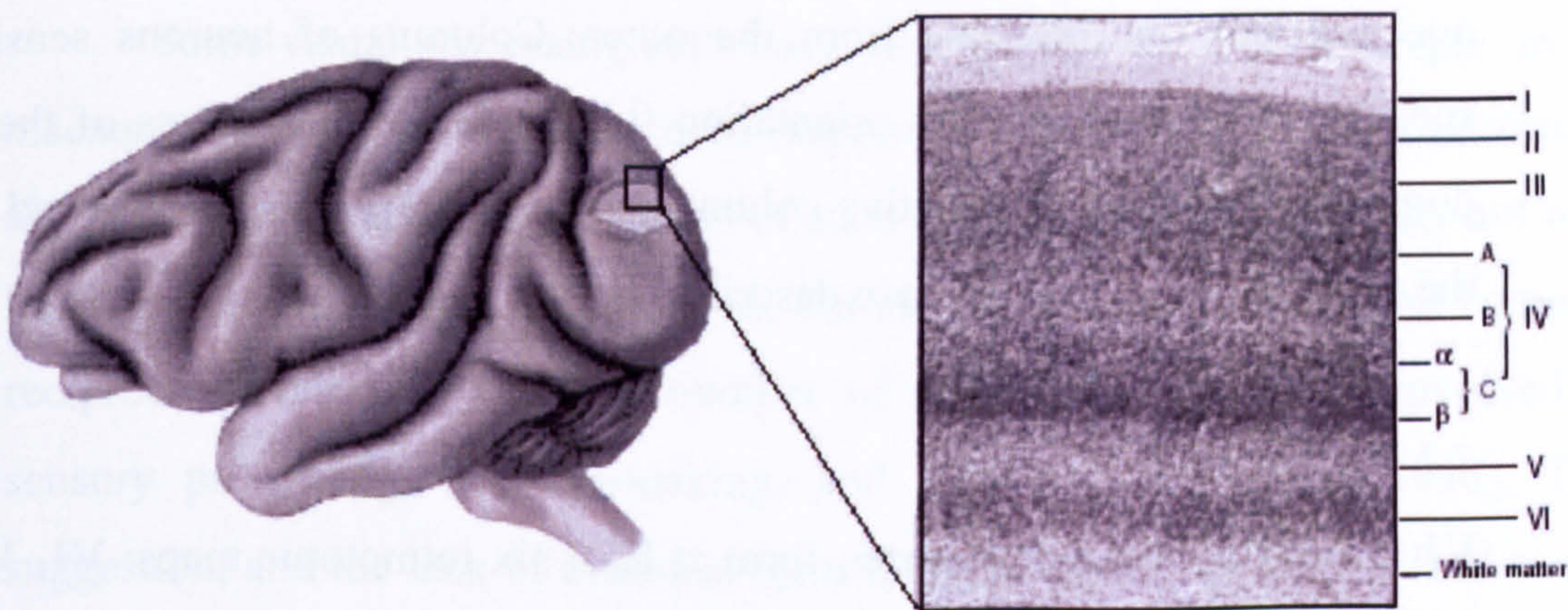


Figure 3.1.4c Nissl stained visual cortex showing layers 1-6 of V1. Source: Hubble (1988)

A set of orientation columns is ocular dominant, corresponding to a region of the retina in one eye, and forms a hypercolumn which is approximately 1 mm^2 at the surface. Hypercolumns extend through the six cortical layers of the cortex (Figure 3.1.4d).

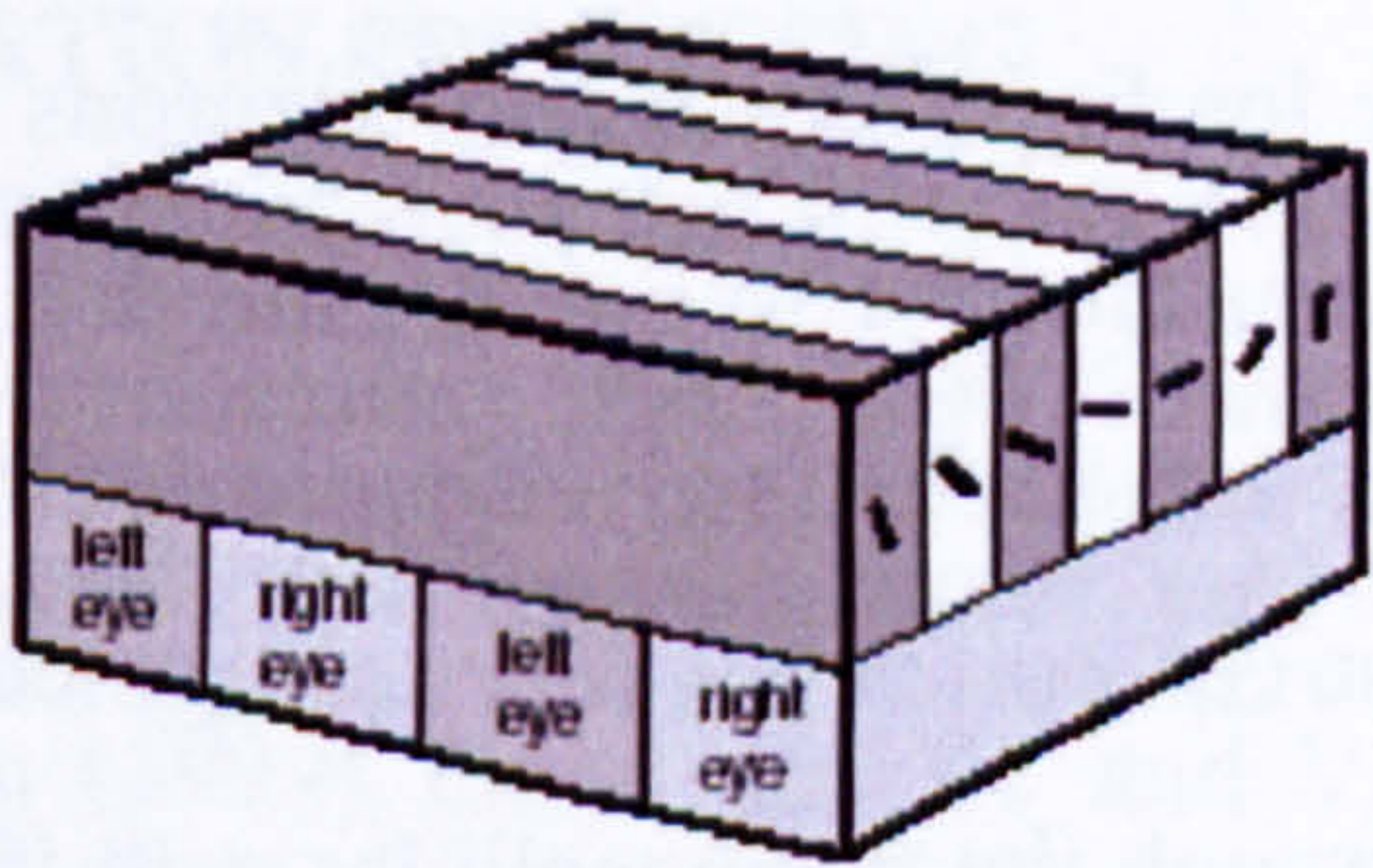


Figure 3.1.4d The relation between ocular dominance and orientation columns. Source: Trappenberg (2002)

V1 neurons may function collectively incorporating contextual information from outside their classical RFs to perform pre-attentive visual segmentation. This

may be achieved by evoking higher neural responses such as boundaries between regions, smooth contours, or pop-out targets against backgrounds, to salient image locations. Ross et al. (2000) showed how distributed features can be bound through feedback from pyramidal cells in V1 layer 2/3 back to layer 4 via layer 6, helping to bind cells into functional columns. When embedded in a cortical hypercolumn, non-classical RF properties may arise through the combination of long-range horizontal grouping interactions in V1 layer 2/3 with shorter-range on-centre/off-surround interactions from V1 layer 6 to 4. Thus layer 6-to-4 inhibition may contribute to the end-stopping effect which has been found with feedback from cortex to LGN. Cortical striations correspond to dark regions from one eye and light regions from the other. Columns of neurons sensitive to specific wavelength but not orientation (blobs), lie in the centres of the ocular dominance orientation-sensitive columns (interblob) of neurons. Beyond V1 are the extrastriate areas. These are described below.

3.1.5 *Extrastriate cortices*

V1 and early extrastriate cortex form at least six retinotopic maps: V1, V2, V3, V3a, V4 and V5. As in V1, neurons in V2 are both simple and complex, but V2 has three types of coarse tangential stripes (Tootell et al., 1983): dark thick (reciprocally connected to layer 4B of V1), dark thin (reciprocally connected to blobs of layers 2-3 of V1), and a pale stripe (reciprocally connected to interblobs) (Livingstone and Hubel, 1988). Thin stripes are orientation insensitive, but more than half code colour and are doubly opponent as are the blobs of V1. Neurons in pale stripes are orientation selective, but insensitive to motion or colour. Rolls (2001) found that over 50% of these cells were end-stopped, which would make them suitable for form processing. Neurons in thick stripes are also orientation selective, and insensitive to colour, but are rarely end-stopped. Neurons in the thick stripe are sensitive to stereoscopic depth and retinal disparity.

Extrastriate cortex has four segregated pathways: i) for motion processing connecting through the magnocellular path from retina to LGN, V1 layer 4C α to layer 4B, to V2 thick stripes to MT then MST; ii) binocular processing for depth information also connecting through the magnocellular pathway, but from layer 4B connects to the thick stripes in V2 and then MT; iii) form processing connects

the parvocellular pathway to V1 layer 4Cb to V1 interblobs to V2 pale stripes to V4 and IT; iv) colour processing connecting the parvocellular pathway to layer 4Cb to V1 blobs to V2 thin stripes to V4 to IT (Livingstone and Hubel, 1988).

V2 receives a point-to-point mapping from V1 and represents motion, orientation, colour and depth, whereas V3, V4, V5 and V6 are specialised for one submodality alone. V3, V4 and V5 project to their satellite areas V3A, V4A and V5A. The outputs from the fovea to V1 are relayed to V4, an area rich in cells which respond to colour (Zeki, 1973), but this area has also been implicated in a number of other processes including form perception, learning, attention, and memory (Walsh, 1999).

Some visual processing cells have very specific RFs. For example, in the macaque temporal cortex, some cells respond only to images of faces (Perrett et al., 1987). The functional properties of such cells most probably reflect not only local processing in each cortical area, but also each area's diverse and complex reciprocal connectivity with a number of other brain structures involved in sensory processing, decision-making, and memory (Logothetis, 1998). This suggestion, and the lack of evidence for a specific 'master area' may lead one to conclude that perceptual integration occurs not in a serial, hierarchical, feed forward system, but through the interaction of partly separate, parallel and functionally specialised systems (Felleman and Van Essen, 1991; Farah, 2000). In fact, perception appears to rely both on distributed and dynamic connections as suggested by Hebb's cell assembly theory (Hebb, 1949). The following section considers visual information processing.

3.2 VISUAL INFORMATION PROCESSING

Visual processing is the sequence of steps that information takes as it flows from visual sensors to cognitive processing. Zeki (1997) proposed that all visual cortices from V1 through IT, and other temporal areas, subserve different aspects of visual perception. Pollen (1999) found that V1 and V2 provided fine and medium grained luminance representations, but not explicit representations necessary to identify complex objects over 2-D or 3-D space. He suggested that because of the increased size of RFs in 'later' areas, 3-D representations could originate in V4 and be completed in the temporal lobe. Furthermore, he proposed

that V1 was activated by non-perceived stimuli. Hence neural networks within distinct cortical areas may be modified by recursive pathways to produce explicit perceptual experience.

Early research suggested that V1 simply pre-processed information for higher visual areas. However, Zipser et al. (1996) found responses of cells in V1 could be altered by stimuli in the RF surround. These responses might be neural correlates of perceptual phenomena such as perceptual 'pop out', or figure-ground segregation. Furthermore, evidence that the visual cortex is capable of dynamic modulation of orientation (Toth et al., 1997), position or size (Gilbert, 1993) tuning was found by adding stimuli to the RF and is believed to be partly attributable to horizontal or feedback connections. RF based processing alone is not sufficient for perception, feedback modulations are also needed (Lamme et al., 1998).

3.2.1 Feed back, feed forward and horizontal connectivity

Evidence for back-projecting pathways comes from anatomical studies (Felleman and Van Essen, 1991; Gilbert, 1993) which found that connectivity between functionally similar regions tends to predominate in the visual cortex. However, connectivity between functionally diverse regions has also been found (Yoshioka et al., 1996). On the basis of its laminar structure, cortical feed forward and feed back connections suggest a hierarchical organisation of visual processing in which V1 is at the lowest level receiving feed forward input mainly from LGN, and feeding forward through several partly segregated pathways to extrastriate areas. However, evidence suggests a more complex processing takes place in higher areas as neurons in the magnocellular pathway have larger dendritic fields and higher spine densities than those in lower areas.

Feed forward connections rapidly transmit information to extrastriate areas from V1. In fact, minimal response latencies in V2 (45 msec.) are only 10 msec. longer than in V1 (Nowak and Bullier, 1997), and are even shorter than those of V1 supergranular layer cells (55 msec.). Many parietal areas show comparatively short response latencies, and in temporal cortex, minimal latencies are 80 msec. These findings suggest that V1 is not the only source of visual input. In fact, when V1 is inactivated, areas V2, V3 and V4 remain inactive, but

neurons in MT and approximately 33% of neurons in V3a remain active (Payne et al., 1996) (Figure 3.2.1).

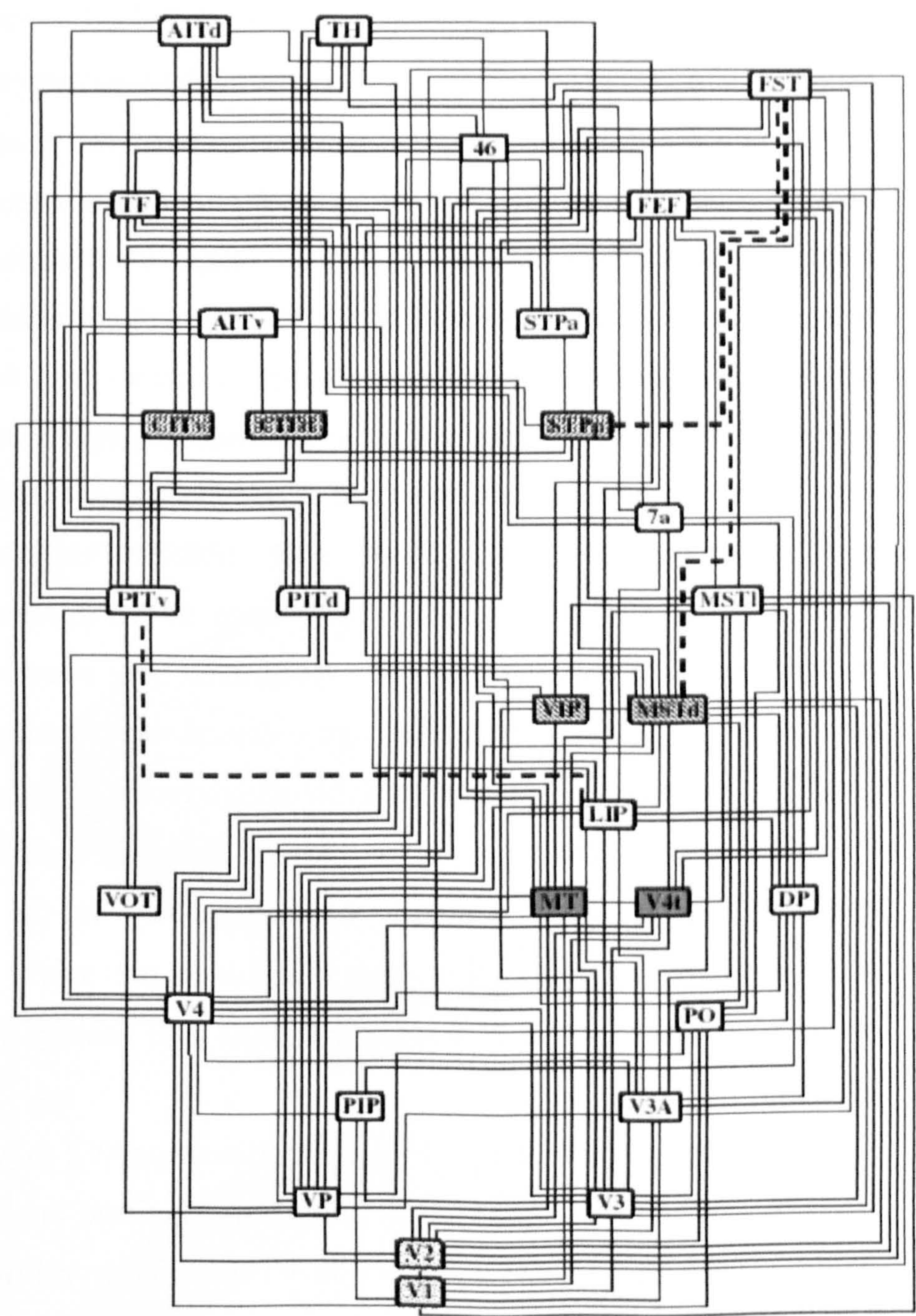


Figure 3.2.1 Connections between visual cortical areas. In the diagram, the dorsal stream is on the right-hand side, the ventral stream on the left. Adapted from Hilgetag et al. (2000)

Feed forward and feedback connections are mostly excitatory (Salin and Bullier, 1995). Thus inhibitory, usually horizontal connections in each area are important for system stability (Lund and Wu, 1997). When excitatory and inhibitory feedback is present, recursive feedback loops can develop. These are ‘the fundamental unit of neuronal activity allowing sensory input to be compared

against established criteria so that they match the input to a template or recognize incongruities and recursively respond until the incongruity is eliminated' (Miller et al., 1960).

Milner (1974) proposed ascending and descending pathways leave mutually consistent trails of facilitated synapses in the complementary pathway to perform iterative pattern recognition. Furthermore, select groups of neurons in a series of visual areas could establish a steady state adaptive resonance (reverberation) between regions if their patterns matched. Reverberation would be suppressed if the patterns did not match (Grossberg, 1980).

The evolutionary development of feedback loops probably arises from the need for an organism discriminate and determine sensory data on the basis of past experience and motivational state (Pandya and Yeterian, 1995). Paradoxically, segregation of an object from its background without prior recognition is difficult, yet frequently recognition only occurs after separation. Grossberg et al.(1994) proposed that figure-ground segmentation and recognition occurred simultaneously and interactively within recursive loops. These operate in both a top-down and a bottom up manner subserving a number of functions, and once resonance is achieved, iterative activity in the loops becomes complementary. However, this hypothesis does not solve the problem of overcoming incongruities between expectation and sensory input (Mumford, 1992).

The main effect of feed back projections from V2 to V1 is to increase selectivity of V1 neurons for small stimuli that activate the centre of the RF (Pollen, 1999). Sublaminae of layer 6 in V1 have reciprocal projections to and from LGN (Lund et al., 1975). From V2 to V1, the predominant back projections originate in layer 6 and the lowest tier of layer 5, and project to the supergranular and infragranular layers in V1. Infragranular terminals end on dendrites near the soma, suggesting strong and excitatory connections; supergranular layers terminate on distal dendrites, which are weaker and modulatory. However, the disproportionately large number of supergranular projections may provide a compensatory effect.

Feed forward input to layer 4 drives neurons, and system stability is maintained through inhibition of input to other layers in order to prevent

feedback loops (Crick and Koch, 1998) which would make the system uncontrollable. However, some direct feedback loops have been found in other areas such as the supergranular layers of V1 and V2, and in V3 and mediotemporal (MT) areas which directly and reciprocally exchange information with V1 layers 4B and 6 (Felleman et al., 1997).

3.2.2 The parietal cortex

Areas of the parietal cortex have been associated with visual processing. The posterior region, mainly concerned with visuospatial processing, has been subdivided into functionally segregated regions including lateral intraparietal area (LIP), ventral intraparietal area (VIP), medial superior temporal area (MST), and area 7a (Figure 3.2.2).

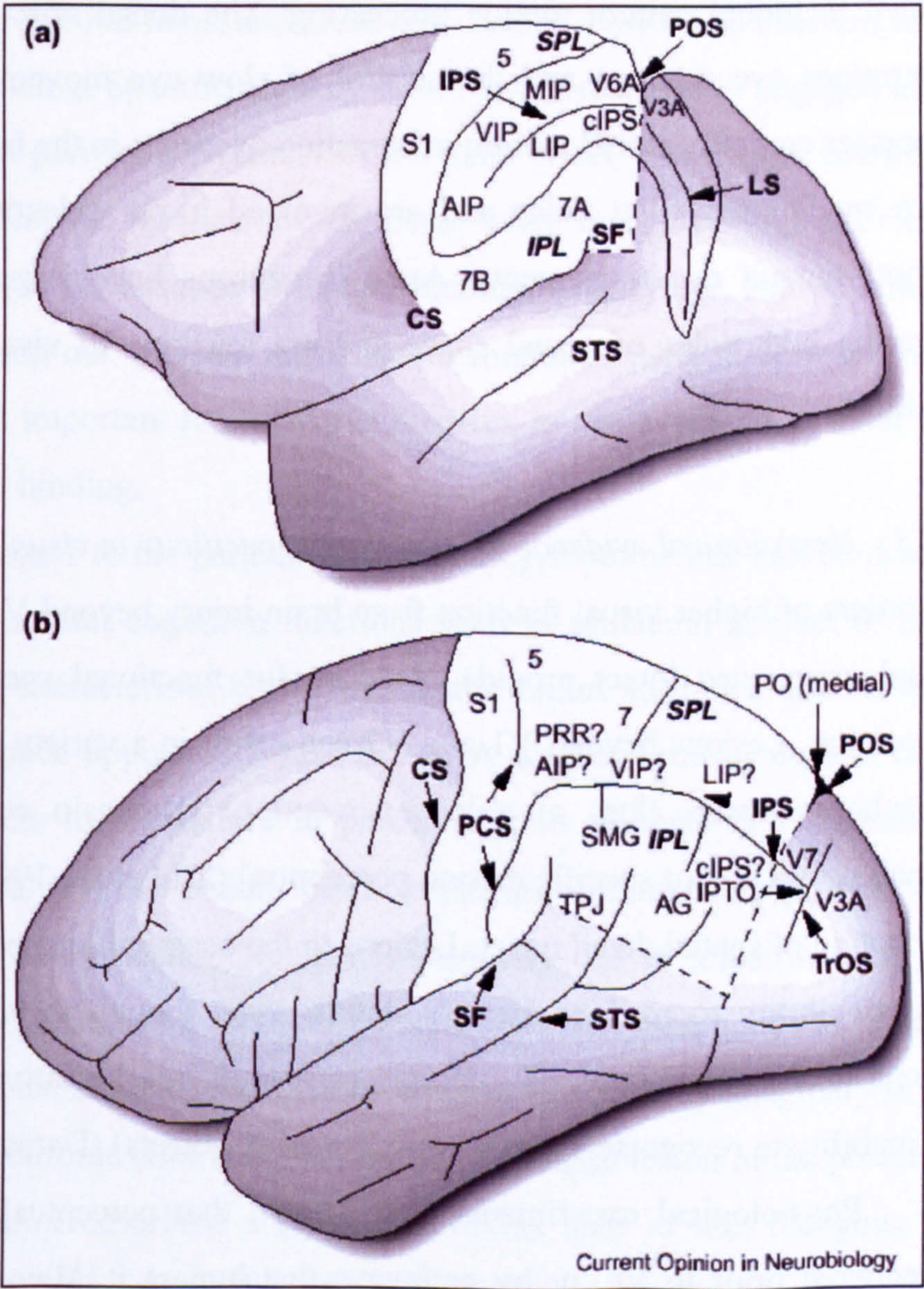


Figure 3.2.2 Lateral view of (a) macaque monkey brain and (b) human brain showing parietal lobes in white. Bold text indicates major sulci, italicized text indicates lobules, and plain

text indicates functional or anatomical areas. Parietal boundaries are based on anatomical criteria]. The central sulcus (CS), Sylvian fissure (SF), and parieto-occipital sulcus (POS) provide unambiguous boundaries. Remaining boundaries extrapolated from other landmarks. Source: Culham and Kanwisher (2001) Neuroimaging of cognitive functions in human parietal cortex. *Current Opinion in Neurobiology*, 11, 157-163.

The responses of LIP neurons implement the transformation of retinotopic spatial coordinates into a head-based spatial representation, allowing a mapping between auditory inputs and visual space. This is useful for guiding eye movements. In addition, some LIP neurons show shape selectivity (Serenio and Maunsell, 1998) which may facilitate the preparation of making actions to objects (Murata et al., 2000). VIP neurons also encode spatial information in a head-centred reference frame, and are sensitive to tactile stimulation and motion (direction and speed). MST neurons can identify self-movement, and neuronal activity in MST is related to global pattern motion processing. The dorsal area of MST (MSTd) determines eye position and the control of slow eye movements. MSTd also processes complex global motion information. Neurons in the lateral area (MSTl) have smaller receptive fields and are involved in the selection of targets for smooth-pursuit eye movements. Area 7a neurons have large receptive fields covering both sides of visual space and are sensitive to visual input and eye position.

3.2.3 Neurological evidence for functional specificity in visual processing

Disorders of higher visual function from brain injury beyond V1 that leave basic visual experience intact provide evidence for functional specificity in visual processing. Lesions beyond V1 and V2 can result in a variety of disorders such as achromatopsias (loss of colour perception) (Damasio et al., 1990); and akinetopsias (loss of specific motion perception) (Zihl et al., 1991), while leaving perception of spatial detail intact. Lesions to the occipital lobe can result in alexia (loss of ability to read or identify letters, even though they can be copied); lesions to the occipitotemporal junction can result in associative visual agnosias (an inability to recognise, but not copy, complex objects) (Damasio et al., 1982).

Physiological experiments have shown that perceptual experiences are engendered prior to V4, or by pathways that bypass it. Heywood and Cowey (1987) showed that removal of V4 impaired discrimination of form and pattern, but not achromatic intensity thresholds. Merigan and Pham (1998) found that

restricted lesions in V4 resulted in relatively mild deficits for elementary visual stimuli presented in isolation, but profound disruption of shape discrimination requiring multiple cues.

Damage to the right parietal lobe may result in a failure to attend complex visual stimuli in the left hemifield even when tests with individual stimuli showed that visual fields were intact and the patients were not hemianopic. Furthermore, patients with lesions to extrastriate areas presented with simultagnosia, that is they saw only fragmentary components of the visual field at any single moment (Critchley, 1953). Similar phenomena have been shown in behavioural studies on rhesus monkey with V1 intact and V4 lesioned (Schiller, 1993). Nobre (2003) used fMRI to investigate the involvement of parietal cortex in binding features during visual search. They identified brain areas common to visual search conditions, and areas specifically sensitive to efficient searches and feature binding. Specifically they found that visual search engaged an extensive network of parietal activations with a main effect for efficient conditions, and a simple effect for inefficient conditions, when compared with efficient pop-out conditions in the absence of feature binding. Conversely, feature binding during efficient pop-out was not sufficient to modulate parietal cortex. These results clarify the important role of parietal cortex in visual search, but not necessarily for feature binding.

Lesions to the parietal lobe display symptoms that can be described as a failure of spatial cognitive functions such as unilateral neglect or hemineglect which are characterised by the systematic failure to notice objects or events in the hemispace opposite to their lesion. A weaker form of neglect is extinction which refers to the failure to perceive or to respond to a stimulus or event contralateral to the lesion when presented simultaneously with a stimulus ipsilateral to the lesion. Neglect is primarily due to impairment in the spatial attentional processing system, specifically the incapacity to attend to the contralesional side, independent of whether the scene is imaginary or perceived. Balint's syndrome (Balint, 1909) is due to bilateral lesion of the posterior parietal cortex. Balint described a patient who was able to see, recognise and name objects 'normally', but behaved as if she were blind. Balint's patients are unable to attend simultaneously to multiple objects (simultagnosia); are unable to change visual fixation from one object to another; and have spatial disorientation.

Humphreys and Riddoch (1992) showed that simultagnosia can be associated with object-based attention. Gerstmann's syndrome (Gerstmann, 1924 (1971)) is associated with a lesion to the left inferior region of the parietal cortex. There are four typical symptoms: left-right ataxia (inability to distinguish left from right), finger agnosia, dysgraphia and dyscalculia.

More recent findings (Friedman-Hill et al., 2003) suggest that posterior parietal cortex in addition to prefrontal cortex, may be involved in inhibiting irrelevant information through interaction with and modulation of the activity of ventral visual neurons that affect object representations.

3.3 VISUAL SHORT-TERM MEMORY

Before describing the known mechanisms of visual short-term memory specifically, the way the brain implements short-term memory in general is outlined. According to Rolls (2003), the brain implements a short-term memory by maintaining neuronal firing during a short period in the absence of a stimulus. Prefrontal and inferotemporal cortex have been differentially implicated in such processes (see Section 3.3.4).

3.3.1 Working and visual short-term memory systems

In order that information can be processed or manipulated, it is maintained in working memory (WM) which is generally understood to be a multi-component system. The brain implements WM during a period following stimulus exposure by maintaining the firing of neurons by associatively modified recurrent collateral connections between nearby pyramidal cells in autoassociative networks (Rolls and Treves, 1998).

WM is associated with prefrontal cortex (Fuster, 1988) and is considered to be comprised of two subsidiary 'slave systems' (one for limited amounts of verbal information: the 'phonological loop', and the other for limited amounts of visuospatial information: the 'visuospatial sketchpad') controlled by the 'central executive' system (Baddeley, 1986) both of which involve prefrontal cortex (Gruber and von Cramon, 2003). However individual areas of cortex are specialized for specific WM processes (see Section 3.3.4).

Short-term memory refers to the limited, temporary retention of information for manipulation or later action (Bugmann and Taylor, 1994). *Visual short-term memory (VSTM)* is similar to the concept of the ‘visuospatial sketchpad’¹⁷ and is one of three broad memory systems including iconic and long-term memory. VSTM refers to the retention of visual information, and is distinguished from iconic memory (Sperling, 1960) which is fragile, decays rapidly, and cannot be actively maintained because it is vulnerable to interference from masking (see Section 3.3.2). VSTM stores visual memories which are robust to subsequent stimuli and which last over many seconds.

Object-location memory (OLM) is an important aspect of spatial memory, enabling us to remember the positions of objects in our environment (Smith and Milner, 1981; 1989). However, separate processes are involved in remembering the fixed positions of objects in the environment and remembering the identity of objects. Furthermore, these processes possibly integrate in a third process in memory (Moscovitch et al., 1995). Neurological studies have shown that the right parahippocampal gyrus (Milner et al., 1997) and the right hippocampus (Stepankova et al., 2004) are associated with processing object-location information in memory. Damage to the right medial temporal lobe (Kopelman et al., 1997) impairs OLM. However, Kessels et al. (2002b) suggested that OLM does not rely only on systems in the right hemisphere after finding that lesions in the right hemisphere impaired position memory, whereas lesions in the left hemisphere disrupted object-location binding.

Delvenne and Bruyer’s (2004) findings support independent, parallel feature-specific memory stores (Treisman et al., 1977; Stefurak and Boynton, 1986; Wheeler and Treisman, 2002), and suggested a linkage between object features might be stored in VWM provided the features are contained within the same contour at the same spatial location (Wu, 2002a). They further claimed that VWM capacity depends on how perceptual mechanisms parse visual input into integrated objects, and suggested that this could occur via the synchronization of neural assemblies (Singer and Gray, 1995) underlying an attentional mechanism.

¹⁷ Visuospatial Sketchpad is a component of the Working memory Model (Baddeley and Hitch, 1974). The visuospatial sketchpad (VS) is the section of one’s normal mental facility which provides a virtual environment for physical simulation, calculation, visualization and optical memory recall.

Kosslyn et al. (1989) described relations between objects in the environment in terms of coordinate and categorical spatial representations in which coordinate representations, purely spatial and thus encoded in the right hemisphere, contain fine-grained information used to guide actions; whereas categorical representations, dependant partly on language processes and processed by the left hemisphere, refer to relative spatial relations. More recently however Postma et al. (2003) proposed that OLM involves distributed processes: encoding object, categorical and coordinate position information, binding object and categorical position information, and object and coordinate position information.

Temporal aspects of VSTM have been less studied than spatial aspects. VSTM creates temporal continuity in a changing world, but it is severely limited. Most studies of VSTM concentrate on the representation of a single visual display, whereas in the real world, visual information occurs over both space and time. VSTM can integrate information across sequential displays successfully in iconic memory, but less well in VSTM (Brockmole et al., 2003). Stimulus onset asynchrony (SOA) has been found to be critical for correct recall. For example, an SOA of ~100 ms allowed successful recall, but this deteriorated with SOAs between 100 and 500 ms (Di Lollo, 1980). However, by increasing the SOA further to up to 2500 ms, Brockmole et al. (2002) found that performance gradually increased until it reached an asymptote at ~1300 ms. They suggested that this because a long SOA was necessary for subjects to form mental image of the first array which would be integrated with the second image when it was presented. This is known as the 'image-percept hypothesis' (Brockmole et al., 2002). Jiang et al. (2005) challenge this idea and suggest a 'limited integration hypothesis' to account for sequential memory in VSTM. They propose that the 'image-percept hypothesis' accounts for neither the limited capacity of VSTM evidenced by much research (e.g. Luck and Vogel, 1997) nor the proposed long time (1300 ms) necessary to form an image of the first array. They conclude that *limited* integration can take place in VSTM, and that multiple visual percepts in the real world are retained by representing events with spatio-temporal continuity, such as target tracking. However, in sequential displays, the presentations might be perceived as separate events and not rapidly integrated successfully, hence the need for the long SOA.

3.3.2 *Memory Masking*

In backwards masking experiments, the participant is shown a brief presentation followed by a presentation of the mask. However, the type of mask, the stimuli, the stimulus onset asynchrony (SOA) of the presentation of the mask, and the task itself determine the efficacy of the effect of the mask on memory and different memory systems. For example, sensory masking lasts ~100 ms (Sperling, 1960); conceptual masking lasts about 300 ms (Potter, 1976); and attentional dwell-time lasts for ~500 ms (Duncan et al., 1994).

One model of backwards masking supposes the mask interferes with information held in a temporary store in early visual areas which supply information to later stages. In this model, the stimulus is stored as a decaying trace which can be processed as long as it exceeds a threshold. The mask reduces the trace activity with greatest effect on duration when the activity is close to the threshold. That is, some time after SOA (Francis, 2000). Alternatively, Enns and diLollo (2001) suggest the mask interrupts processing of the earlier stimulus by replacing it. An alternative model proposes that masking 'catches up' with the earlier stimulus and interferes with it (Breitmeyer, 1984) by propelling via the fast magnocellular pathway, an inhibitory signal that 'catches up' with the information being processed from an earlier stimulus in the slower parvocellular pathway. Bugmann and Taylor (1994) proposed a model in which irregular spike trains were the product of neuronal firing in response to coincidence detection which triggered a random spike train. Amplification, as demonstrated in visual cortex, was provided through positive feedback loops in order to maintain firing above threshold. Backwards masking activated lateral inhibitory mechanisms and interrupted the firing of neurons sensitive to the stimulus. More recently, Bugmann and Taylor (2005) presented a simplified five-layer pyramidal network model (Figure 3.3.2) of backwards masking in V1 in which the response to mask interfered with sustained recurrent activity in response to an earlier stimulus.

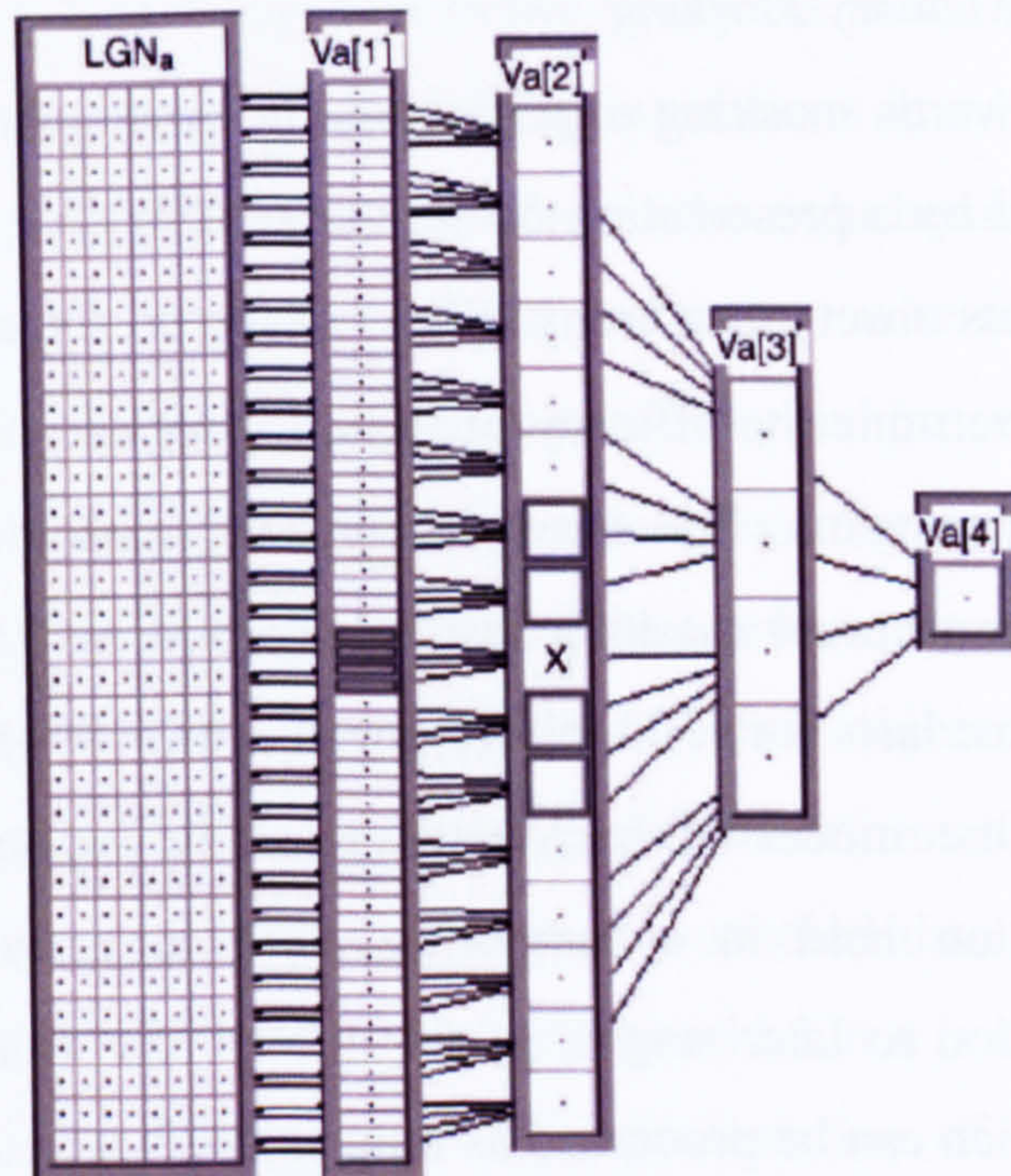


Figure 3.3.2 Pyramidal network model of the initial stages of the visual system. Each neuron has four feed forward and lateral excitatory inputs. The accentuated cells in the third layer (Abbott and Nelson, 2000)) are the four lateral inputs to the cell marked with an X, and the four small accentuated cells in the second layer ((Abbott and Nelson, 2000)) are its four feed forward inputs. The neuron in the last layer has five self-feedback connections instead of lateral inputs. Source: Bugmann and Taylor (2005).

Neurons were modelled as leaky integrate fire nodes tuned to reproduce the selectivity of V1 simple cells in the five layers: LGN/retina, input layer 4 of V1, output layer 2/3 of V1, and input and output cells of V2. A 10 msec. stimulus presentation resulted in a 70-80 msec. response amplified by lateral connections from neighbouring nodes.

Results showed that although the mask, simulated as a reduction in the membrane potentials of all cells in layer 2 of V1, 25 msec after the mask presentation, did not prevent activity reaching the last network layer, it reduced duration of firing in all layers as seen in cortex (Macknik and Livingstone, 1998)

Wodman and Vogel (2005) presented results from a study in which a mask was used to interrupt maintenance and encoding in VWM. Their findings suggested that encoding was not influenced by concurrent maintenance and that the limited work capacity of VWM (Section 3.3.3) is first allocated to new object representations that are encoded and made resistant to masking by subsequent stimuli. In keeping with Baddeley (1986) Woodman and Vogel suggested that encoding and maintenance are independent processes operating in the same

limited capacity store. They propose, as did Hebb (1949), that VWM maintenance is self-sustaining, and therefore it does not need top-down input, but encoding does.

3.3.3 *Limited capacity of VWM*

Visual attention selects behaviourally relevant information and suppresses the remaining irrelevant information, thus facilitating the limited processing capacity of the visual system. In spatial cognition, attention is guided by top-down processing to a certain spatial location in the visual field. This allows object recognition and response in a complex and dynamic environment.

Research on VWM has typically focussed on the limited storage capacity (Sperling, 1960; Luck and Vogel, 1997) of up to approximately four objects regardless of the number of features each (Vogel et al., 2001). This approach suggests that integrated objects rather than their individual features are stored. Lee and Chun (2001) suggested that capacity relates precisely to the number of objects, rather than to the spatial locations of individual stimuli. However, Jonides et al. (1993) showed that observers could remember a limited number of spatial locations independent of the objects in VWM. Heathcote et al. (1994) suggested location would be held in a separately from object identity (Funahashi et al., 1993). These findings are supported by Saiki (2003) who reported that an object's features plus spatio-temporal location are not bound together. Luck and Vogel (1997) did not determine whether capacity is object or spatial location based, but proposed that limitation of VWM might depend on how the display is organized perceptually, and on the number of integrated objects, rather than their individual features. However, (e.g. Jiang et al., 2000; Wheeler and Treisman, 2002) failed to replicate these findings.

Luck and Vogel (1997) demonstrated that in early learning stages visual patterns are coded in terms of their independent substructures, but with increased practice they are stored in terms of more integrated features (van Leeuwen et al., 1988). In this way the role of working memory might be 'unitization' (Czerwinski et al., 1992) in which representational patterns are assembled, and increasingly complex sets of features are bound.

3.3.4 *The role of prefrontal cortex, inferotemporal cortex and other areas in VWM*

Despite much research, the mechanisms required to solve VWM tasks are not yet clear (Renart et al., 2001). However, evidence that prefrontal cortex plays an eminent role in WM has come from neurophysiology (Fuster and Alexander, 1971; Goldman-Rakic, 1987; Fuster, 1988), PET scans (Haxb et al., 1995), and fMRI studies (Cohen, Perlstein et al., 1997). Funahashi et al. (1989) showed that prefrontal cortical neurons have memory fields¹⁸.

Indeed, areas of prefrontal cortex are specialised for different memory tasks: dorsolateral and inferior convexity prefrontal cortex may be related to the memory of spatial responses or objects (Goldman-Rakic, 1987; Wilson et al., 1993; Goldman-Rakic, 1996) or both (Rao et al., 1997) and the principal sulcus/frontal eye field/arcuate sulcus region to the memory of places for eye movements (Funahashi et al., 1989). Prefrontal cortex can be divided into ventral areas, involved in maintaining object information, and dorsal areas, concerned with processing spatial information (Wilson et al., 1993; Smith and Jonides, 1997) (see Figure 3). Context information is also processed in distinct neural pathways (Hayes et al., 2004): the right parahippocampal gyrus is specifically activated during the retrieval of spatial location information, while retrieval of spatial and temporal contextual information is associated with activation of prefrontal cortical areas. Ungerleider et al. (1989) showed that prefrontal cortex is strongly interconnected to extrastriate areas and with inferotemporal cortex which is specialized in higher order visual processing and is affected by memory demands.

Dorsal and ventral prefrontal cortical areas are reciprocally connected to dorsal and ventral stream visual areas, and ventral prefrontal cortex has reciprocal connections with ventral visual areas (Figure 3). Ventral and dorsal prefrontal cortex are considered to be differentially involved in maintaining information in VSTM, but do not distinguish between spatial and object information (Goldman-Rakic, 1996; Rao and Ballard, 1997). Retaining either spatial or featural information differentially activates separate prefrontal cortex

¹⁸ Memory fields are indicated by maximal firing of a neuron to the representation of a visual target in one or a few locations in the visual field, with the same neuron coding the same location in repeated trials, and different neurons coding different locations.

regions (Ungerleider et al., 1998): neurons in anterior prefrontal cortex are activated in WM for features of objects or faces (Wilson et al., 1993), whereas ventrolateral prefrontal cortex exhibits sustained delayed activity related to spatial and object information (D'Esposito, 2001), and dorsal prefrontal cortex is involved in operations to manipulate the memory in VWM in monkeys (Friedman and Goldman-Rakic, 1988) and humans (MacAvoy et al., 1991). These activations reactivate representations in posterior visual areas via feedback projections or loops (Phaf and Wolters, 1997).

However, cortical visual areas that are specialized for processing either object or spatial properties have neurons selective for, or modulated by, the other (Moran and Desimone, 1985). In fact, Goodale and Milner, (1992) suggested that both spatial and object information are unified in one cortical visual pathway, while the other is used for action. Rainer et al. (1998) proposed that prefrontal cortex neurons can represent precise conjunctions of 'what' and 'where' information combined through anatomical interconnections between the dorsal and ventral pathways (Maunsell and Van Essen, 1983), through prefrontal cortex regions connected with these pathways, or through a combination of both (Barbas and Pandya, 1989).

In addition to ventrolateral prefrontal cortex, inferotemporal cortex shows delay period related activity in visual memory tasks (Miller and Desimone, 1994). Both areas are reciprocally connected and are strongly implicated in visual short-term memory (VSTM). Inferotemporal cortex is specialized for higher order visual processing, but is affected by memory demands (Rolls and Treves, 1998), whereas prefrontal cortical neurons maintain activity for an original stimulus in the presence of distractors. This is not the case in inferotemporal cortex, where neuronal firing is disrupted by an intervening stimulus (Miller et al., 1996). However, prefrontal cortex is able to activate representations in inferotemporal cortex (Tomita et al., 1999) after interference (Miller and Cohen, 2001). Accordingly, Kessler and Kiefer (2005) propose that the activity in prefrontal cortex is not for representations *per se*, but for links to where the information can be retrieved.

Recently, medial temporal lobe (MTL) has been implicated in the maintenance and recovery of long and short-term memory tasks. Sakai and Passingham (2004) reported a double dissociation between MTL and lateral

prefrontal cortex regarding reactivation of representations and interference protection in WM and concluded that MTL mediates reactivation from long and short-term memory after disrupted rehearsal while lateral prefrontal cortex protects from interference. However, Kessler and Kiefer (2005) suggest this conclusion is contingent on the stimuli used in the task at hand. They suggest cortical long-range connections between prefrontal cortex for controlled endogenously-initiated recovery, and MTL where retrieval is triggered by an external cue. They add that the type of stimulus affects different memory systems

To perform delayed-match-to-sample tasks, a memory of the sample must be retained so that it can be evaluated against the test stimulus. Studies of cortical cells in monkeys performing delayed-match-to-sample tasks have found that both prefrontal and inferotemporal cortex are differentially important. Miller and Desimone (1994) claimed that intervening stimuli disrupted selectivity in persistent activity in inferotemporal cortex, but not in prefrontal cortex cells. However, Miller et al. (1996) found that prefrontal cortex plays a relatively larger role in WM, whereas inferotemporal cortex is involved in both WM and perception. In addition, prefrontal cortex cells were found to be less stimulus selective than inferotemporal cortex cells which conveyed more information about whether the test stimulus matched a sample or not (Miller et al., 1996). Furthermore, cells in prefrontal cortex, but not inferotemporal cortex, exhibited a progressive increase in firing rate as the trial progressed. Their findings suggested a role for behavioural functions rather than recognition or coding complex objects in prefrontal cortex, and a role for evaluation and coding of an expected future event in inferotemporal cortex.

Whereas Miller and Desimone (1994) had argued that match enhancement and match suppression subserved two distinct types of WM: in both prefrontal cortex and inferotemporal cortex, match enhancement occurred only for the stimulus that matched the sample, and match suppression occurred for any stimulus repetition. In inferotemporal cortex, suppressed responses to repeated stimuli also occur during passive fixation and under anaesthesia (Miller et al., 1991a). Miller, Erickson et al. (1996) found more visually responsive cells in prefrontal cortex than in inferotemporal cortex producing match to non-match modulation effects with suppressed responses to matched stimuli. Miller and Erickson's (1996) 'biased competition' model of attention and WM proposes that

top-down inputs to visual cortex, probably emanating from prefrontal cortex, bias responses in favour of stimuli that are actively sought or that are relevant to behaviour. Chelazzi et al. (1993) suggested that these top-down inputs extend to extrastriate areas such as V4, in addition to inferotemporal cortex.

3.3.5 *Neuronal firing*

At the cellular level, the cerebral cortex is populated by a variety of neurons and interneurons. The majority of interneurons use the inhibitory neurotransmitter, GABA; whereas pyramidal cells use excitatory amino acids. Silito and Murphy (1986) found that pyramidal-nonpyramidal interactions not only facilitate orientation specificity in V1 neurons, but may also be involved in the formation of memory fields in prefrontal cortex. Memory fields, represented by maximal firing of a neuron to the representation of a visual target in one or a few locations in the visual field, may be maintained by associatively modified recurrent collateral connections between neighbouring pyramidal cells which process attractor states in autoassociative networks (Amit, 1995; Rolls and Treves, 1998).

Much information is present in the first 20-50 msec. in single cortical cells and faster at population levels because different cells in the population carry independent information. The strength of synaptic connections can undergo substantial changes on a short time scale depending on the history of the presynaptic input (Marder, 1998b). Goldman-Rakic (1996) showed that subsets of prefrontal cortex neurons are activated phasically in the presence of visual stimuli, then tonically during the delay period, and are reactivated phasically in relation to a memory-guided response. The majority of prefrontal cortex neurons respond in more than one phase. For example, during the delay period in which prefrontal cortex neurons exhibit persistent enhanced firing activity, information is encoded and maintained through the firing patterns of associatively modified recurrent collateral connections between nearby pyramidal cells in autoassociative networks (Rolls and Treves, 1998). Goldman-Rakic et al. (1990) proposed that the neurons carrying out these component processes occupy distinct positions within the laminae of a cortical column which is functionally specialized to a particular memory (e.g. a spatial location).

Cortical neurons typically fire at under 100Hz. which gives little range to specify fine detail and spatial content. However, there is mounting evidence for

perception to result from multiple, near-simultaneous percepts in different cortical areas. From such simultaneous action, a match occurs between sensory data and memory so that a neural representation (a WM) can be achieved. Neurons that can represent a WM have been found in inferotemporal cortex and prefrontal cortex (Wilson et al., 1993). Firing may be maintained for a few hundred milliseconds in inferotemporal cortex (Rolls and Tovee, 1994), but for tens of seconds in prefrontal cortex (Fuster, 1997) by the operation of associatively modified recurrent collateral connections between nearby pyramidal cells producing attractor states in autoassociative networks (Amit, 1995; Rolls and Treves, 1998). In order to maintain the STM when new stimuli are perceived, separate networks for the perceptual and for the short-term memory functions must exist (Rolls, 2000). Renart et al., (1999) also proposed the need for two networks: one for perceptual functions located in inferotemporal cortex, and one for WM functions in prefrontal cortex.

Raffone and Wolters (2001) showed how retention of a limited number of single or bound entities (Luck and Vogel, 1997) may occur in cortical circuits of VWM. Their cortical network model exhibited oscillatory gamma dynamics and assumed that binding originated either in early processing stages, or through long-term memory guided synchronization of oscillatory reverberations. In support of earlier findings (Courtney et al., 1997; Renart et al., 2000), Raffone and Wolter's (2001) model emphasized continuous and recursive interactions between memory and perception. They found that the frequency of the oscillatory reverberations depended on conduction and synaptic delays in the simplified inferotemporal-prefrontal circuit in which feedback produced sustained oscillations after stimulus offset. They proposed that this delayed feedback could be mediated by circuitries with stable transmission along diverging and converging synaptic links, known as multisynaptic synfire chains (Abeles, 1982).

Miyashita (1988) argued that internal representations are based on local neuronal spike rate distributions which persist in the absence of their eliciting stimulus. However, Raffone and Wolters (2001) argued that composite patterns may emerge from pre-existing 'chunking fields' of different orders organized hierarchically to allow dynamic binding of lower level reverberatory assemblies and higher order representations based on perceptual grouping criteria. They

proposed that for cortical information processing, spike rate and temporal coding were not mutually exclusive, but were complementary. They explained that the correlated coherence of individual spikes could organise simultaneously active assemblies into internally coherent and mutually incoherent assemblies. In addition, coherence, based on time-averaged spike rates could reflect the recruitment of new neurons into already active assemblies, and attentive modulation in higher visual areas amplified by increased firing of rate-coherent assemblies and suppressed firing rate of other assemblies. Thus the binding and unbinding operations would be less likely to produce illusory conjunctions.

3.4 SUMMARY

This chapter has outlined the physical aspects of the visual system, neurological evidence supporting function specificity, masking and visual working memory. The visual system, is typically conceptualised as a hierarchy of information processing stages working in parallel to analyse different visual attributes. However, humans recognise and react rapidly to internally or externally generated stimuli which may be novel, embedded in complex scenes, or both. How such a widely distributed, functionally and anatomically specific system codes and decodes sensory information to allow coherent percepts remains unsolved. Experiments in VSTM suggest that most perceptual features (e.g., spatial frequency, orientation, colour) are stored accurately over relatively long periods of time (Magnussen, 2000) and encoded in multiple independent channels, each representing a different feature. Furthermore, evidence suggests that encoding takes place later than V1 in the visual hierarchy (e.g., Bennett & Cortese, 1996). We can thus assume that distributed and dynamic integration mechanisms allow the world to be perceived as a coherent whole. The next chapter is concerned with how and where these mechanisms operate in order to further understand object feature binding.

THE BINDING PROBLEM

INTRODUCTION

Perceptual grouping, identifying constituents that together form a coherent whole, is ‘the most significant function’ in visual processing (Miikkulainen et al., 2005, p.9.). Different grouping tasks take place at different levels of the visual processing hierarchy using spatial, temporal and chromatic relationships (Geisler and Super, 2000). Given this complexity, how are coherent percepts rapidly and accurately represented in the cortex?

In Chapter 3, we described how early vision involves local, parallel feature analyses of the retinal image across the visual field, and described how cognitive functions such as perception, memory and language are based on highly parallel and distributed information processing (e.g. Engel et al., 1999). We reported evidence from physiological data describing a visual system that decomposes complex objects into visual features that are widely distributed across cortex (van Essen, 1985) and processed in distinct brain regions (Hubel and Wiesel, 1968; Zeki, 1978) by functionally specialized neurons, optimally tuned to a particular stimulus attributes (Livingstone and Hubel, 1988). In addition, we described how information on features such as orientation (DeValois and DeValois, 1988), colour, form and motion (Zeki, 1993) is registered at multiple spatial scales. Visual working memory systems were described in terms of their functionality and anatomy.

Given the inherent complexity and distributed nature of visual information processing and memory, we are faced with the problem of how we integrate fragmented perceptions into coherent wholes. Despite the inherent complexity of visual scenes, humans rapidly recognise and act on visual information (Roelfsema et al., 1996). This phenomenon, known as the ‘binding problem’, remain poorly understood despite being approached by investigators from psychology (Treisman and Gelade, 1980), physiology (Nobre, et al., 2003), and computational modelling (e.g. Wersing and Ritter, 1999) for over twenty years.

The notion of binding was first introduced in the context of feature integration (Treisman and Gelade, 1980) and perceptual segmentation (Von der Malsburg, 1981), and has been described as ‘a class of problems’ (Roskies, 1999) that includes vision, perception and cognition. Evidence for the existence of a binding problem for object features has come from experiments investigating the phenomenon of illusory conjunctions (ICs). ICs can be defined as the incorrect combination of correctly perceived features that gives the illusion of an object which is not actually present. ICs cannot be accounted for by working memory (WM) limitations (Prinzmetal, 1995), or by guessing (Ashby et al., 1996). In fact, Prinzmetal et al. (1995) reported ICs without diverted attention and with long durations. They suggested that an incorrect perception of feature location resulted in an IC.

While some studies question the existence of a binding problem (e.g. Shadlen and Movshon, 1999), others have suggested a variety of mechanisms to account for it. These include coincidence (Hubel and Wiesel, 1962), convergence (Barlow, 1972), spatial attention (Treisman and Gelade, 1980; Reynolds and Desimone, 1999), and temporal synchrony (Von der Malsburg, 1981; Singer and Gray, 1995; Gray, 1999). Despite ongoing research, no single proposal is complete. However, they need not be mutually exclusive.

4.1 PROPOSED SOLUTIONS

The following section describes in chronological order the theories proposed as solutions, and discusses research in which they are tested.

4.1.1 Coincidence and convergence

Coincidence and convergence have been proposed as the neural basis for the integration of information (e.g. Barlow, 1972; Rao et al., 1997). Coincidence detectors are neurons which behave like logical AND-gates responding only when specific features are present simultaneously in an appropriate spatial arrangement (Barlow, 1972). Hubel and Wiesel’s (1962) model of cortical receptive fields (RFs) showed that a simple cell with an oriented receptive field is activated only when each of its constituent thalamic inputs is simultaneously active (see Chapter 3).

It is commonly accepted that at early stages of visual processing neurons respond to a stimulus with an appropriately oriented and coloured contour in a specific location. Because neurons at later processing stages are translation invariant, coding for features grouped into whole objects could take place through the convergence of their relevant feature detectors (Barlow, 1972). This may occur in early vision if objects are transient and segregated in a scene. Component feature detectors would then be co-activated and Hebbian learning (Hebb, 1949) could establish stable convergent connections. However, in real life, objects are not segregated, but form part of a complex scene. In this case, transient synchronization could allow segmentation to occur and Hebbian learning would then stabilize the feed forward connections of synchronized inputs only which would determine the type and size of RFs at the next hierarchical level of processing. Thus convergence would be present at all levels of visual processing. This presents a potential problem in that the larger RFs of cells in higher cortical areas would need to be modulated to avoid being driven by appropriate stimulus attributes in separate objects. Fukushima (1975) suggested that modulation could be provided through neurons which respond to low order combinations of features at each stage of visual processing. Within the same paradigm, coincidence detection within the 'what' and 'where' visual pathways has been proposed as the neural basis for the integration of information (e.g. Rao et al., 1997).

However, despite the presence of neurons sensitive to particular conjunctions of features in the visual system (Kobatake and Tanaka, 1994; Tanaka, 1997), the brain does not contain enough neurons to code every possible conjunction (Feldman, 1985; Singer and Gray, 1995). In addition, integration through convergence would not allow dynamic conjunctions in response to novel stimuli. Nevertheless, the incorporation of fixed-feature combination neurons, as in a hierarchical feature analysis system, can implement spatial relations between features crucial for object recognition and discrimination between objects with the same features in different spatial locations. The system starts with a low-level description of the object and builds features based on earlier layers in hierarchical layers so that neurons in 'higher' layers respond to more complex combinations. Advantages of such systems include speed, and biological plausibility. Individually hierarchical convergence and coincident population

coding could not be complete binding mechanisms. However, Gray (1999) proposed that they could operate together in addition to another mechanism that would distinguish between active neural groups.

4.1.2 Spatial attention and the Feature Integration Theory (FIT)

In the spatial attention 'solution', distributed representations of object features are conjoined by attending to a region of visual space (Treisman and Gelade, 1980). This is particularly important when multiple objects are present in a scene. Feature Integration Theory (FIT) (Treisman and Gelade, 1980) is an example of such a theory that suggests visual scenes are decomposed into independent maps of primitive features, and that attention, focused onto a single location, binds together the features from different maps that are present at that location.

Typically research on visual attention has focused on the processes involved in task specific sensory input. Melcher et al. (2005) however found that neural sensitivity to a feature automatically increases throughout the visual field when attention is to a specific object feature. Thus they proposed that units of implicit attentional selection are spatio-temporally co-localized feature clusters which are automatically bound throughout the visual field. Traditionally, studies using event-related potential (ERP) measures have shown that spatial information has priority in attention as location-based selection influences the feed forward information flow through the visual system, whereas feature-based modulations are secondary. However, Hopf et al. (2005) emphasised the flexibility of temporal visual attention by claiming that different sequences of location, feature and object based selection mechanisms are employed to meet the demands of the task in hand.

In FIT, priority maps are used for each feature dimension to obtain an overall activation map where focused attention selects and integrates the features present at particular locations. The activation map indicates the likelihood of the stimulus being the target at that location. FIT suggests that without focused attention no direct link exists between the features from different maps, and features from different objects may become erroneously combined into illusory conjunctions (ICs). Thus FIT proposes that selective visual attention, thought of as a 'spotlight' (Duncan, 1984; Duncan et al., 1994), enhances information within a selected region and iteratively filters out information outside that region

(Treisman, 1982). In such a case, selective visual attention would allow simultaneous analysis of different objects in a complex scene and override the limited processing capacity of the visual system by suppressing irrelevant information. However, other studies have found that participants are able to integrate features when stimuli are presented simultaneously at the same location (Rees et al., 1997) and without attention (Houck and Hoffman, 1986).

Researchers frequently use visual search tasks in order to better understand mechanisms of feature binding. In visual search tasks, observers examine a display to detect an *a priori* target and search time is recorded. The search task can be either 'feature' (target differs from distractors by one feature only) or 'conjunction' (target is defined by the conjunction of two or more properties). Usually two levels of visual processing, pre-attentive and attentive, are considered. Treisman and Gelade (1980) describe a dual-stage binding in which the target 'pops out' and basic visual features are identified in a parallel pre-attentive automatic stage, followed by a serial attentive stage in which features are combined to produce complex object representations. Thus in a feature search task, the target should be detected at the parallel stage so that there is an almost flat slope for reaction time-set size; whereas in the conjunction search task, the serial stage processes each element sequentially until it finds the target conjunction.

According to Treisman and Gelade (1980) attention is explicitly serial. Wolfe et al. (1989) also claimed that the parallel feature and computation stages generate a priority map which serially guides competitive visual attention. Duncan and Humphreys (1989) proposed that although some feature bindings can occur during the parallel stage of visual coding, object recognition comes from competition between feature assemblies of neurons. Their proposal was supported by electrophysiological measures with monkey cortex (Chelazzi et al., 1993) in which competition, after initial parallel processing, led to target identification. Desimone and Duncan (1995) proposed the 'biased hypothesis' in which attention is an emergent property of bottom up and top down neural mechanisms. In support, Corchs and Deco (2001) showed that parallel competition was performed prior to feature binding, and the focus of attention emerged from the dynamic behaviour of synchronous activity of pools of neurons corresponding to different properties of the same object.

The existence of multiple feature maps in visual perception constitutes a binding problem of its own. Kahneman et al. (1992) showed that repeating a visual item facilitates responding to its identity, but mainly the identity and the location remain the same. They proposed the concept of 'object files', an episodic trace containing information about the relationship between an object and its location. However, Hommel points out that the stimulus onset asynchrony affects object-location binding (Hommel and Colzato, 2004), and that object files can be addressed by any feature they contain, not only location. They suggested that feature conjunctions are represented by several separate binary bindings of loose clusters (Hommel, 1998). In such a scenario, the likelihood that a feature can become part of one or bindings increases if it is task-specific.

Several studies (e.g. Duncan, 1984; Vecera and Farah, 1997) have found that concurrent identification of properties of two different objects, regardless of spatial proximity, frequently leads to interference that lasts across a few hundred milliseconds of temporal separation (Raymond et al., 1992). However, this interference does not occur when the task is to identify two different properties of the same object (Duncan, 1984). It thus appears that a gain in activity for one object is accompanied by a loss in activity for others, and winning objects in one system become dominant in other systems (Duncan et al., 1997b).

Although binding by spatial attention minimizes the combinatorial problem, it does not state the actual neural correlates of attentional binding, and therefore cannot be considered a complete solution to object feature binding.

4.1.3 Temporal binding (binding through synchronous firing)

Adrian (1926) demonstrated that increased stimulus intensity resulted in an increased number of spikes. From this he suggested that as firing rate reflects stimulus intensity, firing rates are used to encode information coding not only for the presence of a stimulus, but also for the probability of its presence (Barlow, 1972). However, rate coding mechanisms alone may be considered inadequate because they suffer from the superposition problem which arises when distributed representations of two or more separate objects overlap. Golledge et al. (2003) applied information theory techniques to multiunit recordings from pairs of V1 recording sites in anaesthetised cats using single or separate bars to test the hypothesis that correlated firing serves as a neural code for binding.

Finding that firing rates carried between 89 and 89% of the information, and correlations carried a further 4-11%, Golledge et al. concluded that firing rates, rather than correlations, were the main code for feature binding in V1.

Recent evidence however suggests that both rate and spike coding may be used by the brain (Pouille and Scanziani, 2004). Reliable information can be encoded in spike timing using synchrony across a subset of neurons (Abeles, 1991). This concept is reflected in the temporal binding model. The key assumption of the temporal binding model is that time is an important constituent for neural information processing. In such a model, activity among distributed neurons synchronizes in response to object features which form a group (Milner, 1974; von der Malsburg, 1995; Singer, 1999). Engel et al. (1997) suggested that temporal correlations reflect a dynamic coupling of distributed neurons which increases efficacy, integrates and selects coherent 'chunks of perceptual information'. The Temporal Binding Hypothesis (Von der Malsburg, 1981) implies an holistic view of the visual system in which the activity of individual neurons has little causal efficacy because functional significance depends on the context of other neurons in the assembly. In addition, the hypothesis supports the idea of strong interactions across different streams at all processing levels, and the modulatory effect of top down influences. Such mechanisms would preserve the general advantages of distributed coding schemes, enhance processing speed, alleviate superposition problems, and provide an efficient mechanism for the selection of assemblies for further processing (Singer and Gray, 1995; Singer et al., 1997).

Several hypotheses propose that neurons which respond to the same sensory object might fire in precise temporal synchrony while no synchronisation should occur between cells which are activated by different objects (e.g. Engel et al., 1991b; Niebur et al., 1993). Central to such hypotheses is the concept that binding information is conveyed through neuronal firing patterns in which temporal properties may be characterised by some phase, frequency, or both (Sohal, 1995). However, synchronization alone cannot be used as a binding mechanism when multiple patterns are active, since a synchronously firing neuronal ensemble signifies a single event only (Ritz and Sejnowski, 2000).

Single cell recordings have demonstrated that features belonging to the same object are coded by functionally specific cells firing in synchrony often

observed as coherent oscillations. Neural oscillatory activity has been found between cells coding different aspects of the same visual stimulus in different parts of the visual cortex (Engel et al., 1991a), and LGN (Silito, Jones et al., 1994). Furthermore, synchronisation, in response to visual stimuli can extend beyond a single visual area, to different cerebral hemispheres (Engel et al., 1991a), to different areas of the same hemisphere (Eckhorn et al., 1988), to subcortical visual structures such as the retina, the LGN and the superior colliculus (Neuenschwander and Singer, 1996), hippocampus (Buzsáki and Chrobak, 1995) and frontal cortex (Abeles, Bergman et al., 1993). Furthermore, neurons in cortical and subcortical centres can synchronise within the millisecond range (Engel, König et al., 1992).

Munk et al. (1996) found that increased cortical activation enhances synchronous oscillation and suggested that multiple mechanisms may be involved. Evidence for the role of oscillatory synchronisation in feature binding has been obtained from low-level visual areas under anaesthesia or during passive fixation tasks (Eckhorn et al., 1988), and from experiments with awake monkeys (Frien et al., 1994).

Synchrony can arise through the synchronous firing of individual neurons, or through synchronous firing within a population of neurons. The latter is more biologically plausible because of the stochastic nature of neural firing. Corchs and Deco (2001) proposed that synchronous oscillations are induced by, and are dependent on an attentional mechanism. Their model used a parallel multimodular architecture in which each module comprised a pool of interconnected phase oscillators so that visual attention would be interpreted as the result of the synchronous oscillations of particular pools. They stressed that a variety of oscillations are observed in neural systems and precise synchronisation of neuronal responses in visual cortex has been found to be induced by internal interactions, not time-locked to the temporal structure of the stimulus. However, Lisman et al. (1998) suggested individual neurons could fire synchronously within a short time window and the their population fire in synchrony.

Oscillations in cortical neurons have been found to arise from collective behaviour and are thus more robust (Wilson and Cowan, 1972). Joliot et al. (1994) proposed that neural firing within a single oscillation cycle are bound together even when the exact timing of firing of individual neurons does not

match, whereas as events occurring in different cycles are perceived as separate. Gray (1994) proposed that synchronous oscillations are a natural result of recurrent excitatory and locally inhibitory connections of neurons which may allow a feature detecting neuron to associate itself with a group of other feature detectors to facilitate the description of a coherent object. Within a cell assembly, neurons fire in a synchronised pattern with zero phase shift, but there is a time shift between assemblies so that each remains coherent and the scene is not confused. The synchronisation hypothesis states that overlapping or adjacent segments of a scene that do not belong together should be coded by temporally separable oscillations which would be possible if neurons representing adjacent objects were activated at different relative phase shifts (Von der Malsburg, 1981). Phase, not frequency, of firing is important for information to be transmitted via 'synfire chains' in which distinct groups of neurons stimulate each other in turn by creating reverberations, or by activating a group of neurons in a separate chain (Abeles al., 1993).

Von der Malsburg (1983) found that resonant oscillation occurs when a component of an input signal is recognised, and that the spread and coupling of multiple oscillations results in fusion of all perceptual components into a whole. Later, in the temporal binding model (Von der Malsburg and Schneider, 1986) predicted that synchronisation of spatially separate cells in individual visual areas across large cortical distances would allow binding between visual areas involved in representing different object features. Experimental data demonstrate that synchrony is strongest between cortical sites in which neurons seem to be involved in co-processing of sensory input (Gray and Singer, 1989).

In order to determine the activities of two neural populations with similar properties, Gray et al. (1989) used electric recordings of moving one or two bars of light in two sites with non-overlapping RFs on cat visual cortex. They found that the RFs of visuo-cortical cells, activated by separate bars of light moving in opposite directions, showed low cross-correlation of oscillatory neural activity, and bars moving in the same direction evoked neural oscillatory activity. Furthermore, the strongest activity was recorded when the bar was connected to form a single object which suggests that oscillations are involved in binding. Engel et al. (1991a) found that that when a single long bar was used as input, both populations fired synchronously, but when two short bars were used in the

same location but moving opposite directions, the firing was not synchronous. However, when the two bars were moved in the same direction, there was weak synchronous activity. Such findings led Singer (1996) to hypothesise that neural synchrony, subserved by oscillations, is the neurophysiological correlate of feature binding (Gray et al., 1989).

Usher and Donnelly (1998) manipulated temporal properties in the visual input by flashing an object synchronously or asynchronously with its background. The time-scale of the flashing was shorter than the time taken for visual system integration and hence ruled out conscious perception. Observers were more able to report correctly where the object appeared in one of four background areas when object and background flashed asynchronously. They suggested that this was due to the timing of the inputs which caused the temporal properties of neural firing to change. When the object and background flashed asynchronously, the phase shift between the firing of neurons representing them helped distinguish between them. Similar results have been found (e.g. (Fahle, 1993; Lee and Blake, 1999), but other research has produced conflicting results (e.g. Fahle and Koch, 1995).

An explanation of this difference offered by Blake and Lee (2000) was that the salience of temporal structure on spatial grouping could be modulated by the presence and strength of spatial cues. They posited that temporal flicker cannot overrule explicit spatial structure: it promotes grouping primarily when spatial structure is weak or ambiguous. In their study, responses of neurons to visual stimuli were modulated in time-locked fashion in response to externally time-varying stimuli either by means of temporal coding of the fine temporal structure of the dynamic visual input, or through the average firing rate of a neural assembly fluctuating in a time-locked fashion to time varying stimuli. Either way, Blake and Lee (2000) suggest that the temporal structure of visual input provides a robust source of information for spatial grouping and that temporal and spatial coherence are jointly involved in binding features.

Within the large literature supporting synchrony as a binding mechanism, findings are usually based on evidence from pairs of neurons with overlapping or collinear RFs. However, using broader sampling and stimuli containing partially occluded objects, Palanca and DeAngelis (2005) found that synchronous spiking showed little dependence on feature grouping and thus concluded that synchrony

does not constitute a 'general mechanism' of visual feature binding. Other research (e.g. Ghose and Freeman, 1997; Shadlen and Movshon, 1999; Blake and Lee, 2000) has also questioned the role of synchrony in binding. Ghose and Freeman (1997) proposed that synchronous firing may be an artefact arising from spiking behaviour unrelated to stimulus-driven phase locking. Shadlen and Movshon (1999) stated that even if neural synchrony were to exist, it would act as a signalling agent for feature clusters, but would not compute which features belong to which object. Despite inconsistent results, several paradigms support the involvement of temporal coding in object feature binding (e.g. Mueller and Elliott, 2001).

4.2 FEATURE BINDING IN VISUAL WORKING MEMORY (VWM)

In the presence of multiple objects, the visual system must integrate features associated with each object and segregate features belonging to others. Despite being synthesised correctly in perception, an object's features sometimes can be erroneously combined in memory. Many mechanisms have been proposed for supporting memory for visual feature conjunctions: verbal (Stefurak and Boynton, 1986); attention to the location of the objects (Treisman et al., 1977; Treisman and Gelade, 1980), and neural synchrony and oscillations (e.g. Raffone and Wolters, 2001).

Walker and Cuthbert (1998) claim that verbal representations can support memory for feature associations regardless of whether they are perceived to belong to the same object or not, whereas visual representations support such memory only when the features are perceived to belong to the same object. However, Robertson et al. (1997) suggested response was too rapid for verbal recoding. Treisman and Gelade (1980) suggested that spatio-temporal position could link visual features in memory, but found that spatial location alone does not support memory for shape-colour conjunctions. Instead Treisman suggested that features of an object are bound by selective attention which is possible when items are separated in time or space. Subsequently Nissen (1985) found that when the cue was location, correct recall of colour and shape features were statistically independent, but when the cue was colour, correct recall of shape depended on correct recall of location. She thus suggested that object features are

represented independently with each feature associated with the object's spatial location. Heathcote et al. (1994) argue that observers may use multiple strategies depending on what cues are available.

Keele et al. (1988) suggest that temporal cues might bind features. In their study, items were presented at or above and below fixation. Findings showed that when items were presented at the same location, more errors were reported for items at -1 or $+1$ (just before or just after) temporal position to the target. When items were presented alternately at two locations (above or below fixation), more errors were reported at -2 or $+2$ (the same spatial location) by more distant temporal positions. Keele et al. argued that spatial contiguity is the dominant requirement for binding, and temporal contiguity is only important when features appear in the same location. In support Ashby et al. (1996) found that conjunction errors are more likely to occur when stimuli are proximal or similar.

Several brain areas have been implicated in object feature binding and in visual short-term memory. The parietal cortex is particularly tuned to explicit or implicit spatial information (Figure 3.3.2). Shafritz et al. (2002) found that specific regions of the parietal cortex were found to be preferentially activated by a spatial attention task relative to an object identity task when the stimulus presentations were identical across the two tasks. Additionally, the parietal cortex was more activated when the objects were presented simultaneously at different locations than when presented sequentially at the same location, even when subjects performed the same conjunction task. This suggests that parietal cortex might be involved in spatial attention in visual feature integration. Shafritz et al. (2002) found that when the visual scene contained multiple objects, the right superior and anterior intraparietal cortex were more active during feature conjunction judgments than during single feature judgments. They concluded that this could reflect the representation or manipulation of spatial information or attention during the conjunction task (e.g. Treisman and Gelade, 1980; Friedman-Hill et al., 1995). However, they added that the finding that a region of parietal cortex responds to both a spatial attention manipulation and a feature integration task does not necessarily establish that the neural substrates involved in spatial attention mediate feature binding.

The parietal cortex has been implicated in several distinct functions, in addition to spatial attention, and these other functions may mediate the binding process. Nevertheless, some theories (e.g. Treisman and Gelade, 1980) of feature integration suggest a role for parietal cortex in binding neural activity in areas of the temporal cortex by means of a location map that specifies the spatial relationships among features. Luck and Ford (1998) proposed that feature binding necessitates selective and successive spatial attention to each object which might be performed by the parietal cortex which has been implicated in spatial shifts of attention (Corbetta et al., 1998). In addition, the right parietal cortex is also activated in tasks that require dissociating a target from distractors in the visual field (Marois et al., 2000). Results from Shafritz et al. (2002) suggest an early perceptual and a later working memory stage involvement for parietal cortex (Shadlen and Movshon, 1999) in feature binding when location cues can be used to resolve scene ambiguity.

Humphreys (2001) proposed an important role in visual binding for the parietal lobes, particularly for linking stimuli to a stable representation of space, and for allocating attention to the localizations of the stimuli (Duncan and Humphreys, 1989; Nobre et al., 2003; Robertson, 2003). The posterior parietal cortex is implicated in orienting spatial attention and visual search for conjunction target, but not a feature in a cluttered display (Corbetta et al., 1995; Friedman-Hill et al., 1995; Milner and Goodale, 1995) and could be involved in directing attention serially to successive locations to integrate object features, as suggested in Feature Integration Theory, reflecting both spatial attention and feature binding mechanisms. Nobre et al. (2003) found that posterior parietal cortex is sensitive to the degree of efficiency in visual search, but is less sensitive to the requirements of feature binding.

Corbetta et al. (1995) found that when observers are asked to search for a conjunction of colour and motion, both posterior temporal and parietal areas are activated, but when asked to search for only one of these features, only the temporal lobes are activated. Temporal lobes register the presence of features, but when the target search involves binding, the parietal cortex is activated.

Neuropsychological and neurophysiological evidence suggest that posterior interactions between dorsal and ventral pathways are necessary for binding surface features. This can be explained by the fact the parietal functions

provide the spatial coordinates that allow attention to localise features and separate objects. Damage to the parietal lobe can result in Balint's syndrome (Balint, 1909), characterized by an inability to simultaneously attend multiple objects. Friedman-Hill et al. (1995) reported a Balint's patient who made large numbers of illusory conjunctions (ICs) under prolonged viewing conditions, whereas normal observers make ICs more frequently when presentations are brief and not fully attended (Treisman and Schmidt, 1982; Donk, 1999). When both parietal lobes are damaged, there is no spatial relation between the perceived object and other objects. The observer is thus unaware of the other objects. Neurological evidence (Ungerleider and Mishkin, 1982) has shown that damage to the parietal lobes produces a variety of spatial, but not temporal disorders (Robertson, 2003) including impairment of visual form to space, shape and surface properties, and impairs multiple shape selection and localization, but not binding of edges to shapes. Further evidence suggests that the parietal cortex may not be involved in all binding processes. For example, Rees et al. (1997) found a lack of parietal activation when stimuli were presented serially at the same location.

Humans can maintain approximately four of visual objects in VWM without interference. However, the number of features for each object is unlimited (Vogel et al., 2001). The brain implements a short-term memory by maintaining neuronal firing during a short period in the absence of a stimulus. In monkeys, firing sustained for hundreds milliseconds in inferior temporal cortex, has been observed even when the animal was not performing a memory task. In ventral temporal cortical areas (e.g. entorhinal cortex), the firing may be maintained for longer periods in delayed match to sample tasks (Suzuki and Miller, 1997), and sustained firing for tens of seconds has been recorded in prefrontal cortex (Fuster, 1997). Areas of prefrontal cortex are specialised for different tasks: dorsolateral and inferior convexity prefrontal cortex may be related to the memory of spatial responses or objects (Goldman-Rakic, 1996; Wilson et al., 1993), or both (Rao et al., 1996).

Evidence shows that prefrontal and inferotemporal cortex are differentially implicated in visual working memory processes. For a detailed description of brain areas involved in these processes, please refer to Section 3.3.

4.2.1 *Synchrony*

Activity must be sustained sufficiently in memory to allow time for awareness of the bound features of an object to occur. Such a process might be achieved through re-entrant connectivity (Edelman and Tononi, 2000) or inhibitory connections throughout the brain. Sporns et al. (1989) and Sporns et al. (1991) found that coherent oscillatory activity in well-defined neuronal groups can be established by re-entrant reciprocal connections and that neurons coding different features of the same object oscillate in phase, while neurons coding properties of different objects oscillate at a different phase, or randomly. These findings support propositions by Edelman (1987; 1989) that cooperative re-entry interactions within and among functionally segregated brain areas are a binding mechanism. However, Tononi et al. (1992b) showed that re-entrant connectivity could bind stimuli consisting of simple geometric shapes, but rather than driving target cells, re-entrant connectivity modulated the amount and precise timing of firing. Schillen and König (1994) found that short range synchronising connections support coherent activity at neighbouring topographic and feature locations, whereas long range desynchronising connections establish asynchronous activity at neighbouring topographic, but different feature locations. The desynchronising connections allow segregation of distributed assemblies across multiple feature domains. Thus synchrony may be mediated by the activity of inhibitory interneurons which constrain the pyramidal cells they innervate within the temporal domain (Buzsáki and Chrobak, 1995). However, as oscillations emerge naturally from networks of inhibitory neurons that suppress and control the firing rate of pyramidal cells (Whittington et al., 1995), a large group of pyramidal cells would exhibit synchronous activity primarily because they fire under identical constraints, rather than being determined by specific sensory information. Large scale dynamics expressing stored contextual influences can influence local processing (Bressler and Kelso, 2001) because top down processing not only feeds back hierarchically, but involves a variety of brain signals conveying information related to past events (Engel et al., 2001). Thus local computations operating on incoming afferents are constantly modified by synchronizing and desynchronizing influences impinging on the local region through long-range projections. Various inputs compete for stable resonant states

to express a successful match which is then amplified and broadcast to other populations.

4.2.2 Spatial and temporal aspects of binding in VWM

We are bombarded with sensory information and it is impossible to attend to it all. We would thus benefit if some encoding processes were automatic and thus required no focused attention or conscious awareness. Hasher and Zacks (1979) hypothesized that some aspects of our environment are encoded in memory regardless of age, education or motivation level, and that this process does not interfere with ongoing cognitive processes. The automatically encoded aspects suggested by Hasher and Zacks include location information, temporal order and frequency of occurrence. However their findings are controversial (e.g. Naveh-Benjamin, 1987; Dayan and Thomas, 1995). In the environment, spatial features seem to be naturally integrated with temporal features. Even if the processing of spatial and temporal order information is not automatic, it might be integrated in memory by a single process in which both features are encoded together in memory. Thus when the temporal order of ongoing events in memory is processed, information of where these events took place and vice versa may be automatically encoded. Hill and Stuckey (1993) studied the effect of spatial cues to temporal serial position on a digit-span task and showed that temporal order judgment was superior when temporal and spatial cues were congruent. Moreover, a similar study (Hill and Maodab, 1995) showed that memory for temporal order improved when spatial cues were added, again suggesting that temporal and spatial features are not independently encoded in memory. The high temporal resolution in the spatially superimposed case may reflect cortical neurons with localised RFs selective to multiple features, whereas reporting spatially separated features probably requires an attention-mediated process (Holcombe and Cavanagh, 1999). Earlier, Healy (1975) had proposed that spatial and temporal order information is encoded hierarchically with temporal information primary and spatial recall secondary. Thus when spatial order information is encoded in memory, it must have been translated into a temporal sequence of events. In contrast, Nairne and Dutta (1992) suggested that neither temporal order nor spatial position provided much information about the position

on the corresponding dimension, thus refuting the hierarchical encoding of spatial and temporal order.

If spatial and temporal order information encoding is neither automatic nor hierarchical, binding spatially separated features and binding spatially superimposed features may be mediated by distinct independent memory processes. Parkin et al. (1995) showed that temporal, but not spatial information memory is impaired in older adults. Kopelman et al. (1997) found that spatial memory deficits were due to temporal lobe damage, while temporal context memory impairments were associated with lesions in the frontal lobes. Kohler et al. (2001) proposed a distinction between encoding and retrieval of general (a range of features, including feature integration within a single memory process) and domain-specific (responsible for the processing of a single type of feature) processes. They showed that focusing on object identity information, led to coding of spatial locations, but not the reverse. Based on the finding that spatial information is not automatically integrated with object identity, they reasoned that object identity processing involves a general process, while spatial location processing is domain specific. They did not however exclude the possibility that spatial information is integrated with temporal order information, or other features. The above discussion illustrates the lack of consensus about the relationship between encoding spatial and temporal order memory.

In order to investigate whether spatial and temporal order information are automatically integrated in memory or whether they rely on independent encoding processes, stimuli were presented sequentially at different locations on a computer screen (Van Asselen et al., 2002) and verbalization was suppressed. They found no evidence for automatic encoding or integration of spatial or temporal order information in memory. They concluded that spatial information was not used to encode temporal information or vice versa as attention helped spatial and temporal order memory to the same extent. Their findings support Kopelman et al. (1997) who suggested separate neuroanatomical systems underlie spatial and temporal order information processing in memory, and Kohler, Moscovitch et al. (2001) who claimed that spatial information processing is domain-specific on finding primacy and recency effects for the temporal, but not the spatial order task. However, the relationship between temporal order

information and other features such as object identity information remains unclear.

In cued recall, a list of n items is presented sequentially for retention. Immediately following the last item in the list, a cue to recall one specific item is presented. Results typically show a strong recency effect because at any moment the activation of the most recently presented item is stronger than the activations of the previously presented items. As more units are activated, average activation decreases linearly. When the number of inputs exceeds the capacity of the network, earlier items have a greater chance of being forgotten because newer items are initially activated at a higher level. However, Usher and Cohen (1999) suggested that system limits are determined by the displacement of old by new items, rather than by passive decay. They found perfect performance was obtained with a 4-item list and as the list length increased, there was a gradual drop in recall probability of earlier items.

Because of the highly complex dynamical systems and operations existing in the brain, which are problematic to understand through experimental methods, computer models are used in neuroscience (Miikkulainen et al., 2005). The following section describes relevant computational models of object feature binding in VSM.

4.3 OBJECT FEATURE BINDING IN VWM: COMPUTATIONAL MODELS

Several computational models have been developed in an attempt to account for object feature binding and for visual working memory (VWM). Studies have suggested that frontal lobe short-term memory is mediated by attractor dynamics that can be sustained in the absence of the original stimulus and during new and irrelevant information processing (Miller et al., 1996).

Many connectionist models propose a unitary model of memory (e.g. Sougne, 2000) in which memory retrieval occurs through pattern reinstantiation and completion in activation spreading (McClelland et al., 1995). In connectionist models that use attractor-based retrieval networks, pattern reinstantiation occurs through convergence to an attractor (Hopfield, 1982; Amit, 1989; 1995) and the system exhibits the plausible properties of graceful

degradation, default assignment, and generalization. However, in biological memory systems, features act as clusters in cluttered environments, hence multiple feature clusters could be simultaneously active to retrieve particular patterns. This is known as the multiple reinstantiation problem which in itself 'constitutes a binding problem' (Raffone and van Leeuwen, 2003).

VWM is characterised by selective persistent activity in prefrontal cortex. Many such models are based on the assumption that persistent activity is sustained by reverberatory excitation within a local recurrent neural network (Hebb, 1949; Amit, 1995), and that frontal lobe short-term memory is mediated by attractor dynamics that can be sustained in the absence of the original stimulus and during new and irrelevant information processing (Miller et al., 1996). In the following subsection models of object feature binding and memory are described.

4.3.1 Rolls (1992) VisNet

VisNet was a model of object recognition, and feature binding. Based on Rolls (1992a) hypotheses, VisNet comprised a series of competitive networks, organised in hierarchical layers. The networks exhibited short-range mutual inhibition in each layer. A convergent series of connections from a localised population of cells in preceding layers to each cell in the following layer allowed the receptive field size of cells to increase layer to layer. A modified Hebb rule incorporating a temporal trace of each cell's previous activity enabled the neurons to learn translation invariance. By incorporating fixed rather than dynamic feature-combination neurons, a hierarchical feature analysis system can implement spatial relations between features. Spatial relations between features is crucial for object recognition and for discrimination between objects with the same features in different spatial locations. VisNet starts with a low-level description of the object and builds features based on earlier layers in the hierarchy so that neurons in 'higher' layers respond to more complex combinations. Hierarchical feature based object recognition systems are fast and biologically plausible (Hubel and Wiesel, 1962). In addition, they can be modelled using competitive networks and local learning rules in which neurons self-organise and can perform partial completion. Such a system is VisNet.

In VisNet, connection was determined by a Gaussian distribution of connection probabilities which extend from the focal point of connections for each neuron such that forward connections came from a radius containing approximately 67% of the connections from the preceding layer. Mutual, lateral inhibition ensured different receiving neurons code for different inputs. A modified Hebbian rule, incorporating a decaying trace of previous activity over time, enabled neurons to learn transform invariance. The biological basis for the decaying trace is that after presentation of only 16ms, persistent firing is exhibited for between 100-400ms. This would provide a time window in which subsequent images can be associated. In addition, the glutamate binding period in the NMDA channels may last for as long as 100ms which may implement a trace rule by producing a time window over which the average activity of each presynaptic site affects learning (Rolls, 1992a). The trace rule achieves translation invariance because it is based on self-organisation. Invariant representations were developed in each successive layer. Because the low order combinations contained sufficient spatial information to identify a unique object, Rolls (1992) claimed the model provided a solution to the spatial binding problem.

4.3.2 Usher and Niebur (1996) model of object oriented attention

Usher and Niebur (1996) presented a parallel dynamic processing model for object oriented attention in which various visual stimuli (shapes, colours, letters, etc.) were represented by sparsely connected, competing, mutually inhibitory, cell assemblies. The model exhibited a response to cue and target sequence which has been found in responses in inferotemporal visual cortex of monkeys performing a visual search task (Chelazzi et al., 1993). During stimulus display, there was an enhanced response which decayed, but remained above a spontaneous rate, after the cue disappeared. Subsequently, when the target plus several distractors was presented, the activity of all stimulus-driven cells is initially enhanced. After a short period of time, the activity of the cell assembly representing the cue stimulus was enhanced while the activity of the distractors decayed because of mutual competition induced by the increasing activation and to a small top-down input induced through 'expectation'. Usher and Niebur (1996) suggested that this process may be used by the visual system for selecting

an expected target appearing at an uncertain location, among distractors. Furthermore, they claimed that objects compete for limited processing capacity at several points between input and response.

4.3.3 Camperi and Wang (1998) visuospatial working memory in PFC

Camperi and Wang (1998) found that the absence of cellular bistability¹⁹, noise or distractors influenced activity profiles and induced systematic drifts over time that resulted in memory loss after a few seconds. They claimed that cellular bistability was a contributing factor in the formation of memory fields of delay period activity which was contingent on the input-output relation of a single cell. Compte et al. (2000) proposed that bistability was implemented by the dominance of recurrent γ -aminobutyrate (GABA)-ergic synaptic inputs, and stabilized by N-methyl-D-aspartate (NMDA) receptors at recurrent synapses, which allow the network to switch between resting and structured activity states as found by Amit and Brunel (1997). Likewise, Compte et al. (2000) reported increasingly less precise memory for cue location in the delay period as a result of random drifts in time.

4.3.4 Usher and Cohen (1999) maintenance and response selection in STM

Typical results from cued recall show a strong recency effect. In neural models of this phenomenon as more units are activated, average activation decreases linearly. When the number of inputs exceeds the capacity of the network, items presented earlier are more likely to be forgotten because newer items are initially activated at a higher level. Usher and Cohen (1999) suggested that frontal load system limits are determined by the displacement of old by new items, rather than by passive decay. Their reverberatory model of active maintenance and response selection in the frontal STM system, showed perfect performance with a 4-item list, and a gradual drop in recall probability of earlier items as the list length increased. The ability of the model to sustain more than one active

¹⁹ Bistability enables neurons to be switched on to a persistent firing pattern by transient cue stimuli and turned off to a resting state of spontaneous activity by another suitable signal. Hence, under equal conditions, a uniform rest or excited state with a non-uniform spatial activity profile can occur. Wilson, H. and J. Cowan (1973). "A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue." *Kybernetik*(13): 55-80, Amari, S. (1977). "Dynamics of pattern formation in lateral inhibition type neural fields." *Biological Cybernetics* 27: 77-87.

representation following stimulus offset depended on the strength of recurrent excitation and lateral inhibition. They suggested that maximal span increased or decreased with the level of excitation or inhibition, and was dependent on the task requirements. Furthermore, they claimed that subjects might have control over these levels.

4.3.5 *Renart et al. (1999); Renart et al. (2001) attractor models of WM*

In order that the WM be maintained during periods in which new stimuli are to be perceived Renart et al. (1999) proposed the need for two networks: one for perceptual functions located in inferotemporal cortex, and one for WM functions in prefrontal cortex. Later, Renart et al. (2001) developed a model comprising two reciprocally connected attractor models which reproduced response properties of inferotemporal cortex and prefrontal cortex cells during delayed-match-to-sample and memory guided attention experiments. Each module contained a large number of sparsely connected excitatory and inhibitory neurons organized in micro-columns (M-Cs) as found in monkey dorsolateral prefrontal cortex²⁰.

Cell connectivity within an M-C was five times denser than between M-Cs. The modules interchanged only excitatory signals, hence synapses to excitatory neurons from excitatory afferents within an M-C were potentiated, and the others were depressed. The baseline efficacies were such that inhibition dominated. A large background excitatory input was also present from outside each module. The two modules exchanged random, but plastic long-range excitatory signals with feed forward connections that were stronger than feedback connections.

This resulted in each M-C in each of the two modules receiving (sending) stronger synapses from (to) an associated pair in the other module. The amount of depression was chosen so that the net afferent synaptic efficacy was the same with or without plasticity. Using leaky integrate-and-fire model neurons, all neurons with the same statistical properties were grouped into a single sub-population, characterized by the mean rate of its neurons. Thus coincident

²⁰ Adjacent excitatory and inhibitory neurons in the monkey dorsolateral PFC have similar response preferences during sensory and delay periods of VSWM tasks, whereas more distant pairs have inverted preferences. In addition, tuned inhibition has been found to contribute to the generation of persistent activity Rao et al. (1999). *Journal of Neurophysiology*.(81): 1903.

patterns of activity in inferotemporal cortex and prefrontal cortex were associated by reciprocally connected Hebbian-like synapses. Bidirectional long range signals between the areas were dependent on the state of the whole network, and were correlated with the representations of the external stimuli in each module.

Findings showed a biologically plausible dynamic excitation-inhibition balance through the M-C architecture which led to temporally irregular activity and tuned inhibitory responses as found in previous studies (Shalden and Newsome, 1994), as well as competition which led to the suppression effects observed by Chelazzi et al. (1998).

4.3.6 Compte et al. (2000) model for spatial working memory

Compte et al.'s (2000) network model for spatial working memory (SWM) incorporated recurrent connectivity consistent with cortical columnar organization of monkey prefrontal cortex (Goldman-Rakic, 1995). The model encoded memory in graded fashion to produce a continuum of activity profiles ('bump' states). Before cue presentation, neurons showed emergent spontaneous asynchronous activity at a few spikes per second, achieved through a combination of suprathreshold external inputs representing background activity in other brain areas, and a strong feedback inhibition in the network. During the cue period, a pattern of increased activity due to the external input to the subpopulation of neurons with preferred cues closest to the cue stimulus, developed around the location of the cue. Throughout the delay period, persistent asynchronous activity achieved through strong excitatory feedback between cells sharing similar tuning properties, remained restricted to a selective neural subpopulation. During the response period, a transient overall increase of external inputs to the network led to a transient increase of firing which stopped the persistent activity (Goldman-Rakic et al., 1990) because of strong inhibitory feedback, and refreshed STM. These inputs increased the firing rates of pyramidal cells and interneurons which resulted in a recurrent inhibitory input that switched off the 'bump' state.

The model showed that strong local recurrent excitatory connections and recurrent inhibition produce persistent activity which create memory fields in individual neurons. They proposed that the emergence of persistent activity required sufficiently strong recurrent synaptic excitation, and that dynamic

stability depended on the predominance of NMDA receptors to the recurrent synaptic excitation, as well as on a strong inhibition that dominated the recurrent circuit. They explained their findings by suggesting that during the stimulus presentation a 'bump' state was selected by the cue, but during the delay period when the cue was withdrawn, no external input constrained the peak's location of the network activity profile, so it could drift. Then, due to continuous random external inputs which marginally affected the location, but not the shape of the bump, it drifted randomly.

4.3.7 Raffone and Wolters (2001) cortical network model of memory and perception

Consistent with Luck and Vogel's (1997) findings, and in support of Courtney et al. (1997), Raffone and Wolters (2001) presented a cortical network model that emphasized continuous and recursive interactions between memory and perception which showed how retention of a limited number of single or bound entities may occur in cortical circuits of VWM. In their model, the frequency of the reverberations depended on conduction and synaptic delays in the inferotemporal cortex-prefrontal circuit in which feedback produced sustained oscillations after stimulus offset. Raffone and Wolters (2001) proposed that this delayed feedback could be mediated by circuitries with stable transmission along diverging and converging synaptic links, known as multisynaptic synfire chains (Abeles, 1982) and argued that the correlated coherence of individual spikes could organise simultaneously active assemblies into internally coherent and mutually incoherent assemblies. In addition, coherence, based on time-averaged spike rates, could reflect the recruitment of new neurons into already active assemblies, and attentive modulation in higher visual areas amplified by increased firing of rate-coherent assemblies and suppressed firing rate of other assemblies. They concluded that for cortical information processing, spike rate and temporal coding were not mutually exclusive, but were complementary.

4.3.8 De Kamps and van der Velde (2001) model of object-based attention

De Kamps and van der Velde (2001) presented a 5-layer (to represent the five hierarchical layers in the ventral stream) perceptron model. Objects were presented at V1 which comprised 4 identical layers representing different

orientations. Four differently shaped objects could be presented at V1 at four locations (a total input of 16 patterns). Presentation of one object at one of the four locations led to distributed network activity, localized in V2 and spreading further in higher visual areas due to the increased size of RFs. The model demonstrated competitive activity in AIT which determined which object remained active. Feedback connections from AIT to lower areas allowed the object's location to be selected. A second ANN with identical architecture, but with reciprocal connectivity was trained with Hebbian learning using the activities in the feed forward network so that AIT was the input layer and activity propagated downwards to V2.

The fan-out structure of the network propagated activity throughout the network. The feedback activity interacted indirectly with the feed forward activity by means of disinhibition (using feedback activity as a gateway for feed forward activity). Because of the network-wide parallel activation, location selection was fast and once the target was selected, information from the lower areas propagated to the dorsal stream so that a saccade could be initiated. Although this model appears to present a solution the binding problem between form and position, position coding is implicit in the lower visual areas.

A second model presented by De Kamps and van de Velde (2001) was trained to identify colour and form of the presented objects in AIT. At lower visual cortex layers, all object information was present. Identification of one feature in higher areas led to the position of the object being selected via feedback connectivity which in turn led to the other features being selected. Hence, the network could correctly identify the colour when presented with the shape of an object, and position information elicited the colour of the object. However, their model did not allow for independent coding of colour and form. This would be problematic if a large number of colour-form combinations was presented at the same location. In such a scenario, the feedback information would lose its discriminatory powers and no single location would dominate. If colour and form were processed in parallel in different streams, colour information would be available after location selection. However, this would not be biologically plausible as the brain does not completely separate colour and form processing. They maintained however that their model was fundamentally correct, if lacking in realism.

4.3.9 Deco and Lee (2002) multimodular model of visual information processing

Deco and Lee (2002) proposed a system with three interacting modules representing the early visual cortex: V1/V2 (EM), ventral stream (VM) and a dorsal stream (PP/PO) (Figure 4.3.9). EM contained orientation-selective complex cells and hypercolumns as evidenced in the primary visual cortex. VM contained neuronal pools encoding specific object classes as evidenced in the inferotemporal cortex. DM contained a map which encoded positions in the retinotopic coordinate. EM and VM were connected with symmetrical connections developed with the Hebbian learning. EM and DM were connected with symmetrically localized connections modelled with Gaussian weights.

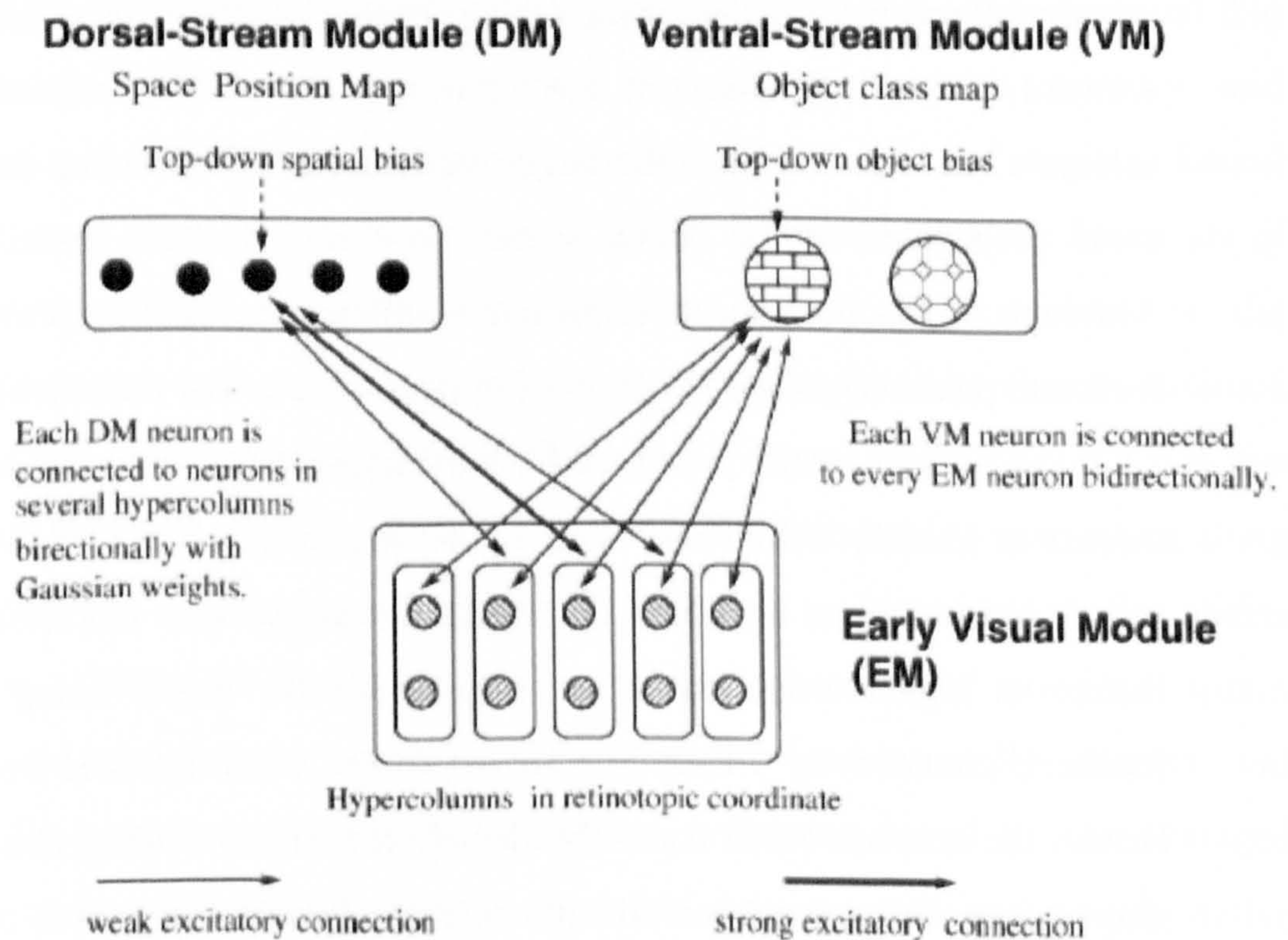


Figure 4.3.9 A schematic diagram of Deco and Lee's (2002) model. The model contains three modules: the early visual module (EM), the ventral-stream module (VM) and the dorsal-stream module (DM). Source: Deco and Lee (2002).

A unit in each module represented a pool of neurons with similar properties characterised by its activation (mean firing rate of pool) and an input current. An inhibitory pool integrated information from, and fed back information to, all the excitatory pools within each module. Connection between modules was excitatory, biasing the competitive dynamics in each module. Concentration of

neural activities to an individual pool in VM corresponded to object recognition. Concentration of neural activities to a unit in DM corresponded to object localization. EM provides a buffer for VM and DM to interact.

Spatial attention was generated by top-down bias input to DM, and object attention to VM. A small enhancement due to object or spatial attention in EM was sufficient to communicate the bias between VM and DM. Because of the mutual constraints among the three modules, there was simultaneous localisation of the target in DM, identification of the target in VM and highlighted features in EM. Thus the object's location, identity and detailed features were bound into a unified percept.

4.3.10 Deco and Lee (2001) models of spatial and object-based attention

Deco and Lee Deco (2002) presented a model of spatial and object-based attention composed of modules structured to represent the dorsal and ventral pathways of the visual system. One particular aspect of visual information processing was modelled in each stream: in the ventral stream, translation-invariant recognition was modelled; and in the dorsal stream encoding of visual space in retinotopic coordinates was modelled. The ventral stream comprised four modules: V1, V2-V4 (to pool and channel responses from V1), inferotemporal (IT) (to achieve translation invariance), and v46 (an area of the prefrontal cortex found to maintain a short-term memory). The dorsal stream comprised 3 modules: V1, posterior parietal (PP) (mediate spatial attention modulation), and d46, corresponding to the dorsal part of area 46 (short-term memory of spatial location and generate attentional bias).

During the learning mode, synaptic connections between V4 and IT were trained using Hebbian learning of a specific object at random positions. In the recognition mode, an object was found by biasing the system with a top-down component to the IT module which enhanced activity of pools in V4 and IT adding to the visual input received by V1. Thus the increased firing in a particular part of V1 led to increased activity in the forward pathway from V1-V2-V4 to PP. Resultant increased firing in PP represented the location of the object. In addition, to IT, PP also received a top-down input from d46 which specified the object's location. PP then drove competition in V2-V4 in favour of the pool associated with that location.

An extension to this model, reported in Rolls and Deco (2002) was a model that bound multiple feature components. They hypothesised that selective attention is the result of independent competition mechanisms within each feature dimension. Thus in their model, each visual feature m was represented by $N(m)$ values, and for each feature there were $N(m)$ layers of neurons representing the presence of each feature value. Mutually inhibitory cell assemblies comprising fully connected integrate-and fire neurons were allocated at every location in each layer resulting in a sparsely distributed representation as the activity of a population represents the presence of different features at a particular location. The PP module, reciprocally coupled with different feature maps, bound the feature dimensions at each item location. IT connections provided top-down information by including an extra excitatory input to the corresponding feature layers.

Deco and Lee (2001) and Rolls and Deco (2002) claim that object and spatial attention can be produced by dynamic interactions between the ventral and dorsal streams. In their models, object or spatial attention emerged as a result of top down bias from short-term memory systems (prefrontal) to the ventral (inferotemporal) stream, or the dorsal (posterior parietal) stream respectively. The result was that the top-down bias guides the dynamics to concentrate at a specific spatial location or object features.

4.3.11 Domijan (2003) model of feature binding and capacity limits in VWM

Several models of feature binding in VWM require biologically implausible explicit connections between features (e.g. Raffone and Wolters, 2001), or activity dependent modulation of synaptic weights to form feature conjunctions (e.g. Ungerleider, 1995; van der Velde and de Kamps, 2003). In contrast, Domijan (2003) developed a competitive two-stage model of sustained neural activity in PFC to simulate feature binding and capacity limits in VWM. The network exhibited sustained activity after the input had ceased. In the first stage, object features were stored in parallel network layers without explicit conjunctions. Features were conjoined into the object in the second stage in a selection layer as proposed by Wheeler and Treisman (2002).

Features of the same object were labelled with the same activity amplitude, characteristic of the competitive network. Thus the model used

neither explicit representation for feature conjunctions except position, nor synaptic modifications of connections between feature representations. Consistent with neuroimaging data which shows no division between object and spatial components of a visual task in PFC (Rao et al., 1997), Domijan proposed that explicit binding existed only between position and the other features. In keeping with physiological evidence, sustained activity was achieved after the input was removed and the network exhibited a limited storage capacity. Despite these results, the model's plausibility is brought into question because the brain region the selection network represents was not specified, and input from a single cell in the selection network was both excitatory and inhibitory. However, Domijan claimed that the key component of the model was its nonlinear dendritic computation (Koch and Segev, 2000).

4.3.12 Van der Voort van der Kleij et al. (2004) model of VWM

Van der Voort van der Kleij et al. (2004) presented a model of VWM in prefrontal cortex to explain the fact that a limited number of objects can be stored in VWM, but the number of features of the objects is not limited (Vogel et al., 2001). The model comprised a 'blackboard' representing one layer of ventral prefrontal cortex, linking to different processors (van der Velde, 1997) (Figure 4.3.12). Activation in ventral prefrontal cortex reverberates in a characteristic way associated with cortical WM activation. Objects, containing representations of partial identity and location information, were represented in the blackboard from PIT. The blackboard bound the information from the processors, which were networks for feature identification.

The bottom layer of V-PFC was connected to higher visual areas: AIT for processing shape, and PPC for processing location. If one feature was selected in AIT, all representations of in the bottom layer of V-PFC consistent with that feature were activated. In addition, an attended location in PPC activated all possible representations for that feature. The bottom layer represented the focus of attention whether it was based on object or location feature information. Thus the interaction between the bottom layer of V-PFC and the blackboard selected the object representation consistent with the current attentional focus, and the resulting activation in the 'select' layer could be used to bind the features.

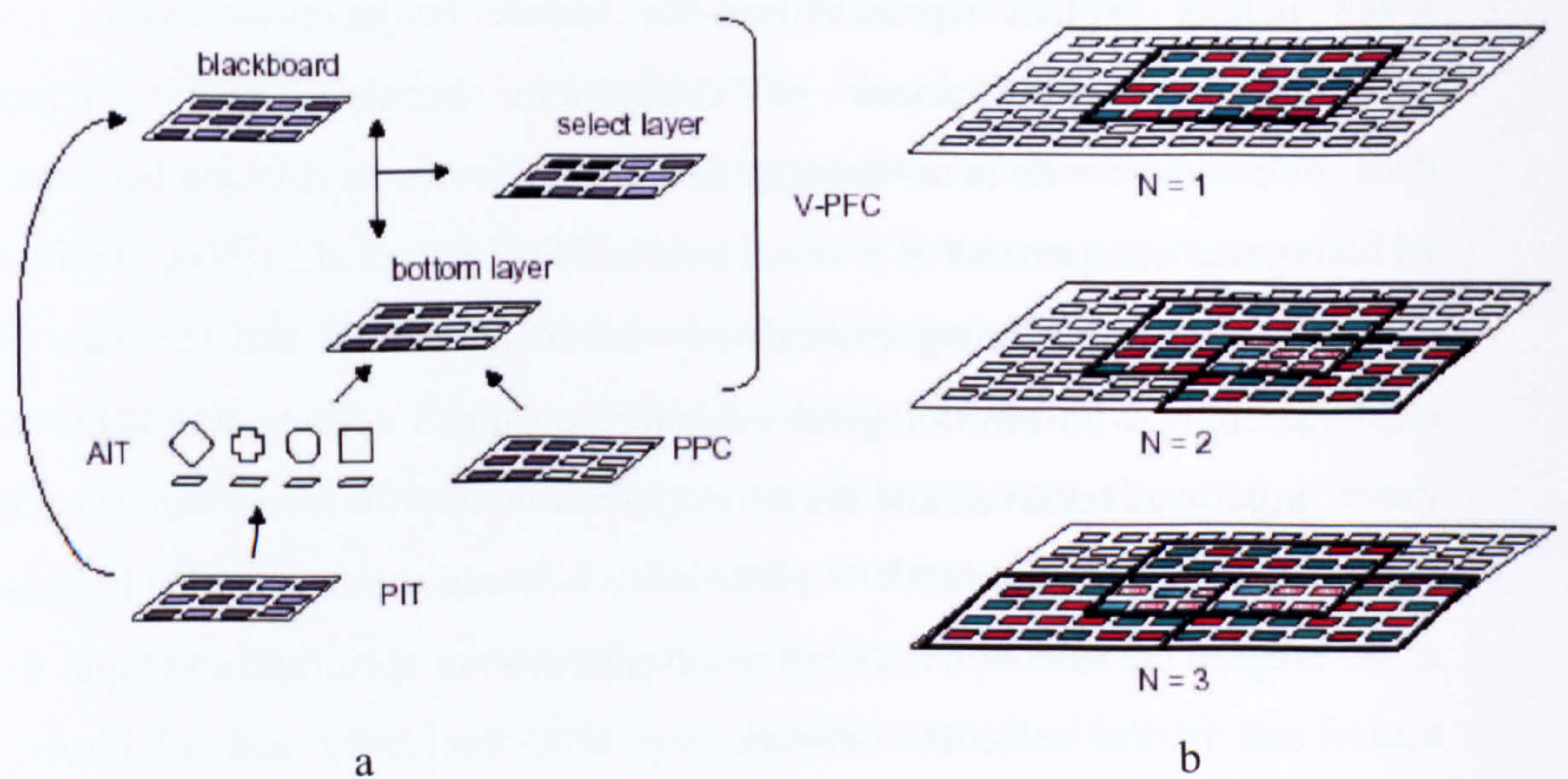


Figure 4.3.12 A blackboard architecture in prefrontal cortex (PFC). PIT = posterior inferotemporal cortex; AIT = anterior inferotemporal cortex; PPC = posterior parietal cortex; V-PFC = ventral prefrontal cortex. Figure 4.3.1b Interference between representations in the blackboard. Source: van der Voort van der Kleij et al. (2003).

They concluded that the limited capacity of VWM was partly dependent on the distance between objects in a display. They also suggested that selection by location information depended on the amount of interference between object representations in the ventral pathway of V1. They suggested that attention may increase the sensitivity for attended features by providing extra input, or by boosting the response strength for attended features without changing the sensitivity.

4.4 SUMMARY

Identifying constituents that together form a coherent whole, is ‘the most significant function’ in visual processing (Miikkulainen, et al., 2005, p.9.). Despite the inherent complexity of visual scenes, and the distributed nature of visual processing, humans rapidly recognise and act on visual information (Roelfsema et al., 1996). In this chapter we considered the most common proposed solutions to the binding problem, and showed that although each goes some way to solve the problem, none alone is sufficient.

Activity must be sustained sufficiently to allow time for awareness of the bound features of an object to occur. We considered the role of visual working memory in object feature binding on the evidence that even when an object’s

features are integrated correctly in perception, they can be erroneously combined in memory. Sustained memories are possibly mediated through re-entrant (Edelman and Tononi, 2000) or inhibitory connections throughout the brain (Sporns et al., 1989; Sporns et al., 1991), and cooperative re-entry interactions may be a binding mechanism (Edelman, 1987) or have a modulatory effect on the amount and precise timing of firing. Thus oscillations emerging naturally from networks of inhibitory neurons would exhibit synchronous activity primarily because they fire under identical constraints, rather than being determined by specific sensory information. We have considered spatial and temporal aspects of binding in VWM and described relevant looked at computational models of object feature binding in VWM.

To sum up, evidence suggests that a single paradigm, albeit hierarchical features, spatial, or temporal mechanisms is inadequate to solve the object feature binding problem. The experiment described in the following chapter attempts to clarify this statement.

BEHAVIOURAL EXPERIMENT**INTRODUCTION**

Object feature binding appears not to be a problem for the nervous system in most situations, but it remains a conceptual problem for neuroscientists. Nevertheless, the finding that binding errors are made in some situations and with some neuropsychological conditions prompts the question: how and where does binding take place? Neuropsychological data demonstrate that object feature binding involves activation of disparate areas of cortex. The problem is how these disparate cortical areas communicate to provide a coherent object representation.

5.1 EXPERIMENTAL GOAL

The goal of the present experiment is to obtain evidence through behavioural investigation to support or refute current object feature binding theories described in Chapter 4. The rationale behind this goal is that no single theory alone provides a satisfactory solution to the binding problem.

Despite the existence of several theories of binding, the two major contenders considered in this thesis are the ‘Spatial Theory’ (e.g. Cave and Zimmerman, 1997), and the ‘Temporal Theory’ (Von der Malsburg, 1981). These, and other theories, are discussed and critiqued in Chapter 4. The rationale for the selection of the ‘Spatial Theory’ and the ‘Temporal Theory’ as components of a theory of object feature binding are presented in Chapter 1.

A corollary of the binding problem is the phenomenon known as ‘illusory conjunction’ (IC) in which features from one object are transferred to another object to lead to the perception of an object which is not present in the scene. Recall from Chapter 4 that if the Spatial Theory of object feature binding is correct, features from objects appearing spatially closer to the target object should be recalled more frequently than features of objects appearing further in space from the target; if the Temporal Theory of object feature binding is correct, features of objects appearing temporally closer to the target object should be

recalled more frequently than features of objects appearing more distant in time from the target. In addition Cohen and Ivry (1991) proposed that ICs are more likely when multiple objects are present (such as in the Spatial condition) than when individual objects are presented sequentially (such as in the Temporal and Spatio-temporal conditions) and with proximity (in all conditions).

In order to further understand how and where object feature binding takes place, the behavioural experiment addressed these questions:

- how does proximity affect object feature binding?
- how does presentation style affect object feature binding?
- how is spatial recall affected by temporal properties of stimuli presentation?
- how is temporal recall affected by spatial properties of stimuli presentation?
- does proximity affect the frequency of IC reports?
- is the occurrence of ICs affected by spatial or temporal properties?
- is the occurrence of ICs affected when stimuli are presented simultaneously or sequentially?

The categorical independent variables of (IVs) are i) position (adjacent, intermediate and distant), ii) condition (spatial, temporal and spatio-temporal), iii) presentation rate (frames per second), and iv) cue colour (turquoise, yellow, green or purple). The dependent variable (DV) is error count for each non-target position for each condition. Therefore DV data are absolute. The observations are dependent in that the same participants take part in each condition.

5.2 METHOD

This section describes the method used in the behavioural experiment.

5.2.1 Stimuli and coding

As erroneous response data might provide insights to the binding problem and correct responses would not, we sought to collect and analyse erroneous response data. Responses were categorised according to their position relative to the 'target' object. The target object was that particular object which was the same

colour as the post-cue in that presentation. Relative positions were adjacent, intermediate or distant. Each presentation featured four rectangular objects, which differed in colour, size and orientation (see Figure 5.2.1). On presentation of the post-cue, observers were required to give the size and orientation of the perceived target object. Therefore for each presentation, two responses, one for size and one for orientation, were recorded. An IC was recorded in the event of both responses related to features corresponding to a single non-target object.

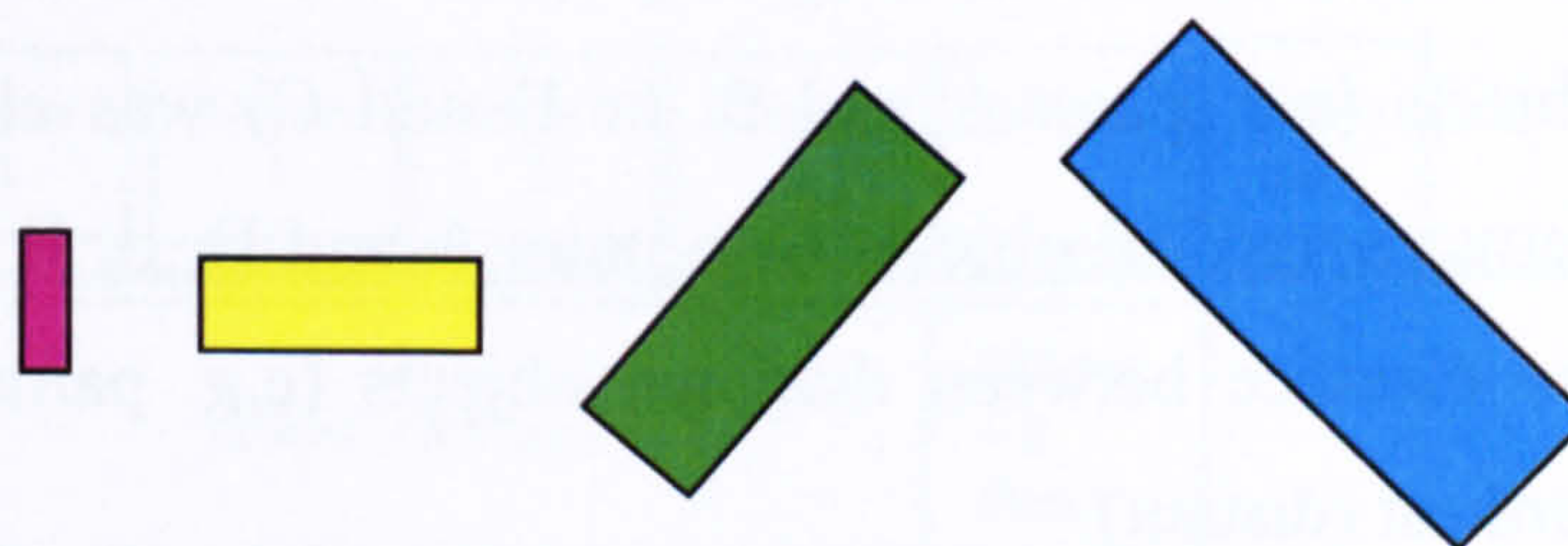


Figure 5.2.1 Example of stimuli showing all possible different colours: purple, yellow, green and turquoise; sizes: small to large; and orientations: vertical, horizontal, right diagonal, left diagonal.

5.2.2 Stimulus onset asynchrony (SOA)

In order to collect sufficient response data for meaningful analysis, errors were ‘forced’ by manipulating presentation rate and stimulus onset asynchrony (SOA) in frames per second (f/sec.) by using a variation of the ‘staircase method’ (Cornsweet, 1962). The staircase method is a modification of the Method of Limits in which stimulus intensity is systematically decreased until it becomes too weak to be detected, then it is systematically increased. In this study, time intervals of each presentation were sequentially increased and then decreased as follows: presentations were shown at 2 f/sec., 3 f/sec., 4 f/sec., 5 f/sec., 6 f/sec., 7 f/sec., 8 f/sec., and 9 f/sec. then the presentation rate was reversed (9 f/sec. to 2 f/sec.). The aim was to achieve approximately 50% error rate on both object features (size and orientation). This procedure was repeated for each observer, for each condition, so that presentation rates for each observer for each condition were set *a priori* to achieve approximately 50% error on each feature. Therefore, the rate could be different for each condition for a single observer (Table 5.3.1). Presentation rate (f/sec.) and stimulus time on screen were the same. That is, if a presentation refresh rate was 5 f/sec. then that frame would be on screen for 200 ms.

5.2.3 The experimental conditions

In order to fully exploit spatial and temporal features, three conditions: spatial, temporal and their combination: spatio-temporal, were designed. These are described below.

In the Spatial condition, the four composite objects (Figure 5.2) were presented simultaneously in four pre-defined positions (see Figure 5.2.3a) on the circumference of an imaginary circle such that the distance between horizontal objects (e.g. pairs A and B, or D and C) was closest (adjacent), the distance between vertical objects (e.g. pairs A and D, or B and C) was intermediate, and the distance between diagonal objects (e.g. pairs A and C, or B and D) was furthest (distant).

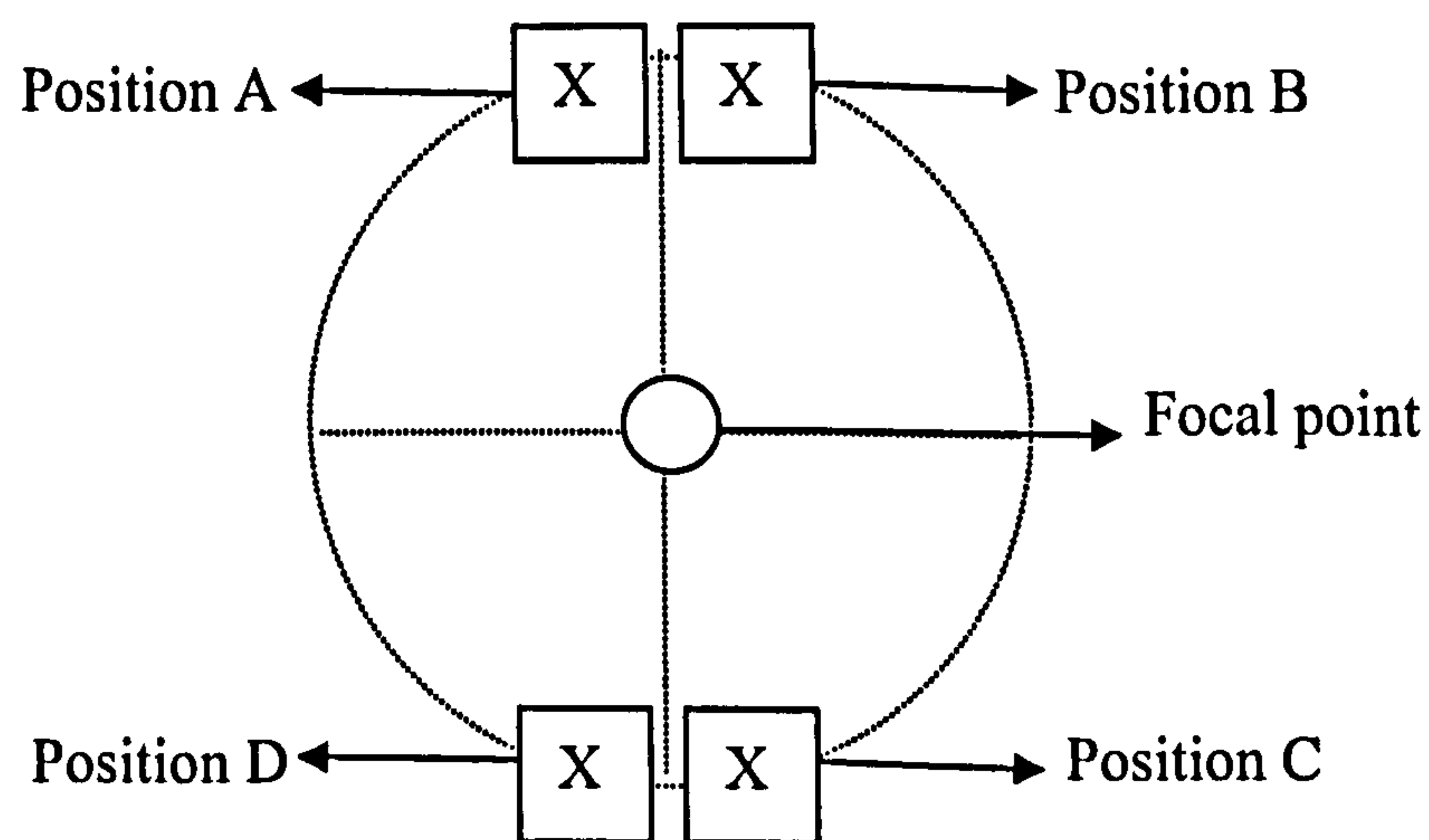


Figure 5.2.3a. The spatial positions of stimuli in the Spatial and Spatio-temporal conditions. The stimuli were presented simultaneously at the spatial positions A, B, C and D in the Spatial condition, and sequentially at the spatial positions A, B, C and D in the Spatio-temporal condition

Each individual presentation was preceded and followed by a checkerboard mask (Frames 1 and 3, Figure 5.2.3b) in an attempt to eliminate memory traces of the preceding screen. Hence presentation in the Spatial condition was: Mask → Objects → Mask → Cue. The presentation time on and off screen (illustrated with an arrow (→)) was equal (the SOA) determined when ~50% error was achieved for both features.

The first mask was presented for a period (in Figure 5.2.3b, the period was 125ms or 8 frames per second) when the participant indicated he or she was ready to start. The four objects were simultaneously presented at the locations shown in Figure 5.2.3a for the same time period, then another mask was presented for the same period, followed by the post-cue, presented for the same period ((Figures 5.2.3b and 5.2.3e).

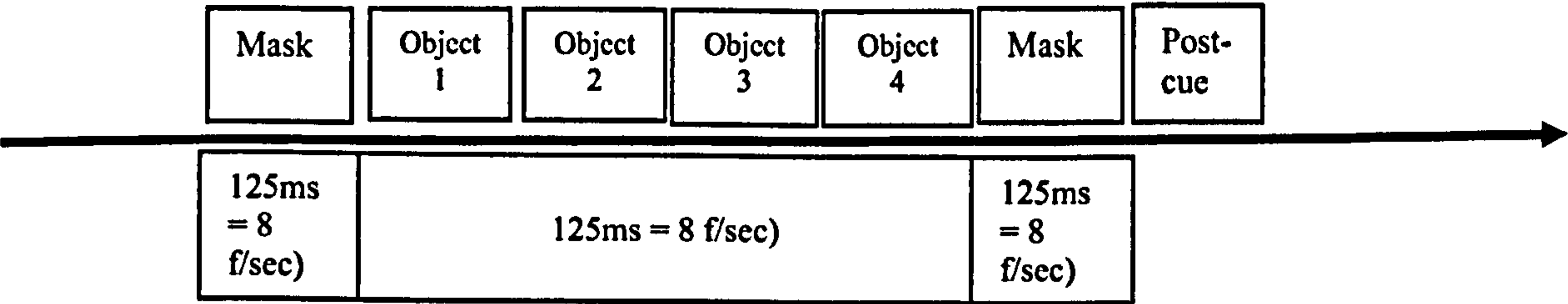


Figure 5.2.3b Sequence of stimuli presentation in the Spatial condition. In this illustration, the SOA is 125ms which = 8f frames per second. The entire presentation would thus take 375ms.

The response time was not recorded. Thus the participants could take as long as they needed to respond. This sequence, described above, was repeated 64 times for each presentation rate (2-9 frames per second) to determine which achieved closest to 50% error on each feature. The data resulting from observations at this presentation rate was used in the subsequent analyses.

In the Spatio-temporal condition, the four objects were presented sequentially in the same four pre-defined positions as in the Spatial condition (Figures 5.2.3a), and responses for both spatial position and temporal position were recorded. Presentation in the Spatio-Temporal condition was: Mask→Object1→Mask→Object2→Mask→Object3→Mask→Object4→Mask→Cue (Figures 5.2.3c and 5.2.3f). As in the Spatial condition, this sequence was repeated 64 times for each presentation rate (2-9 frames per second) to determine which achieved closest to 50% error on each feature. The data resulting from observations at this presentation rate was used in the subsequent analyses.

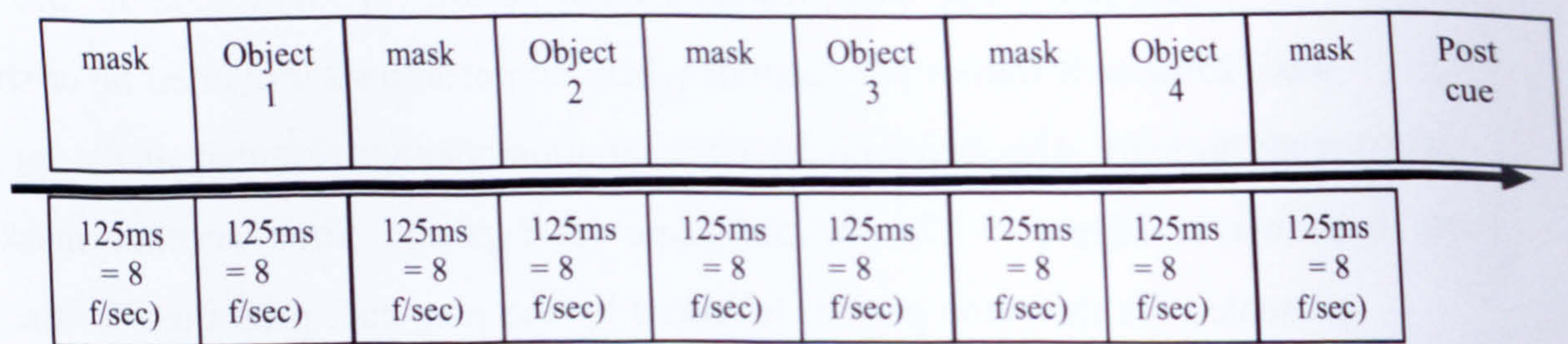


Figure 5.2.3c Sequence of stimuli presentation in the Spatio-temporal and Temporal conditions. In this illustration, the SOA is 125ms which = 8f frames per second. The entire presentation would thus take 1125ms.

In the Temporal condition the four objects were presented sequentially at the central fixation point (see Figure 5.2.3d). Temporal distance was defined by temporal sequence. Hence an object appearing either +1 or -1 temporally relative to the target was coded adjacent; an object appearing either +2 or -2 temporally relative to the target was coded intermediate; and an object appearing either +3 or -3 relative to the target object was coded distant.

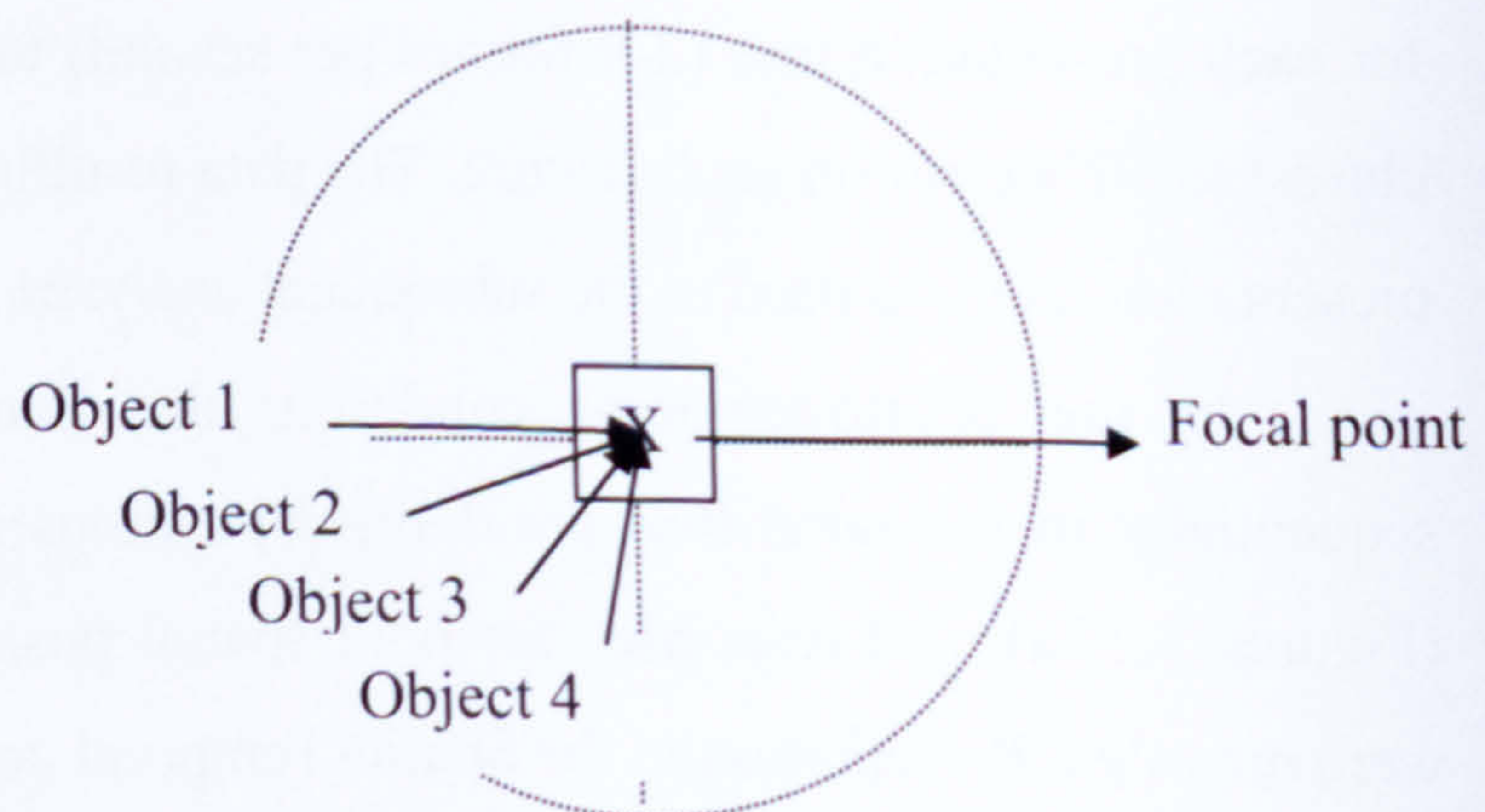


Figure 5.2.3d. The spatial position of stimuli in the Temporal condition. The stimuli were presented sequentially at the central focal position.

Presentation in the Temporal condition was:

Mask → Object1 → Mask → Object2 → Mask → Object3 → Mask → Object4 → Mask → Cue (Figures 5.2.3c, 5.2.3d and 5.2.3g). As in the Spatial and Spatio-temporal conditions, this sequence was repeated 64 times for each presentation rate (2-9 frames per second) to determine which achieved closest to 50% error on each feature. The data resulting from observations at this presentation rate was used in the subsequent analyses.

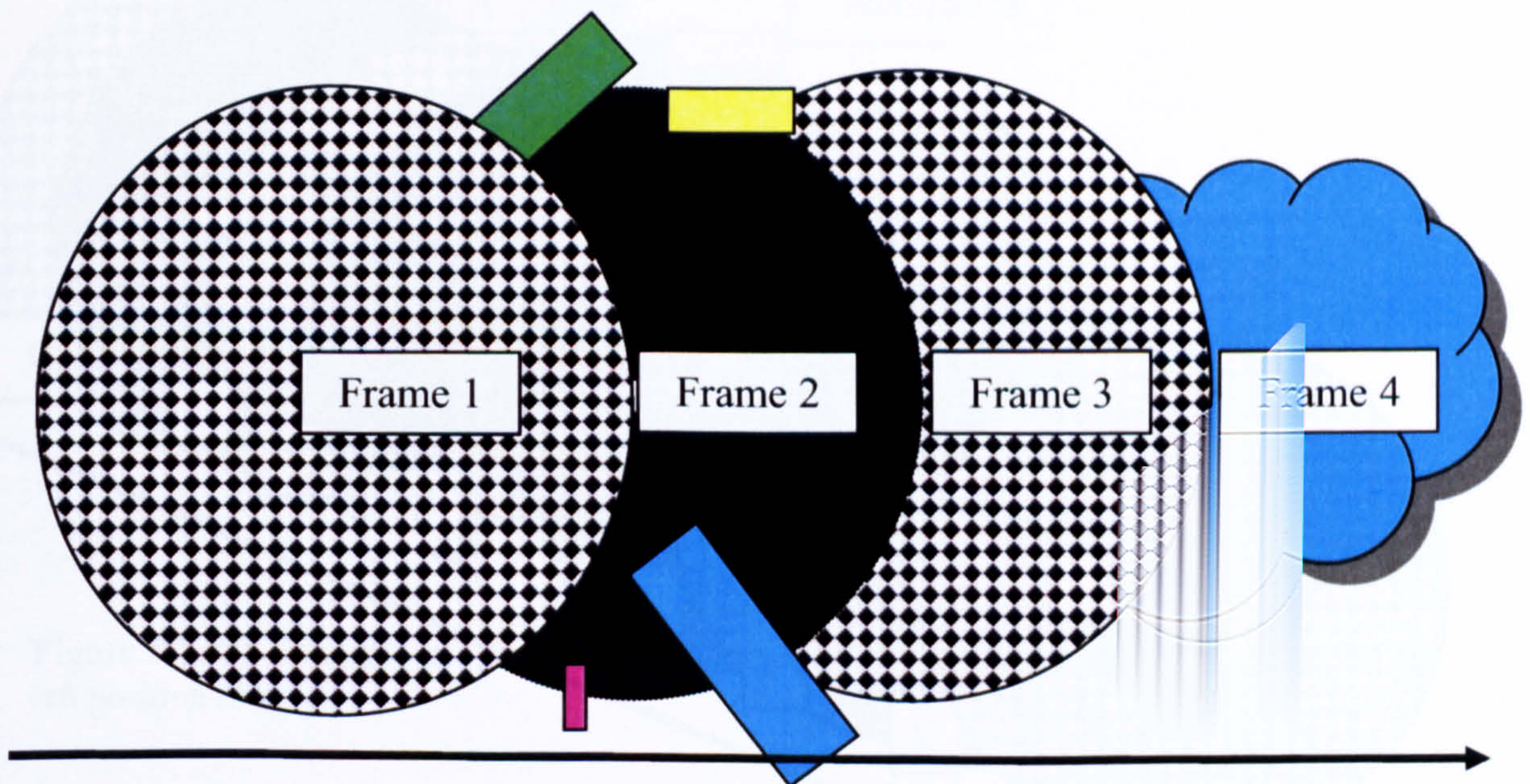


Figure 5.2.3e Example presentation in the Spatial condition. Frame 1, checkerboard mask; frame 2, the four objects; frame 3, checkerboard mask; frame 4, the post-cue. In this example, the correct response would be size 1 (largest), left diagonal. Time is indicated by the arrow.

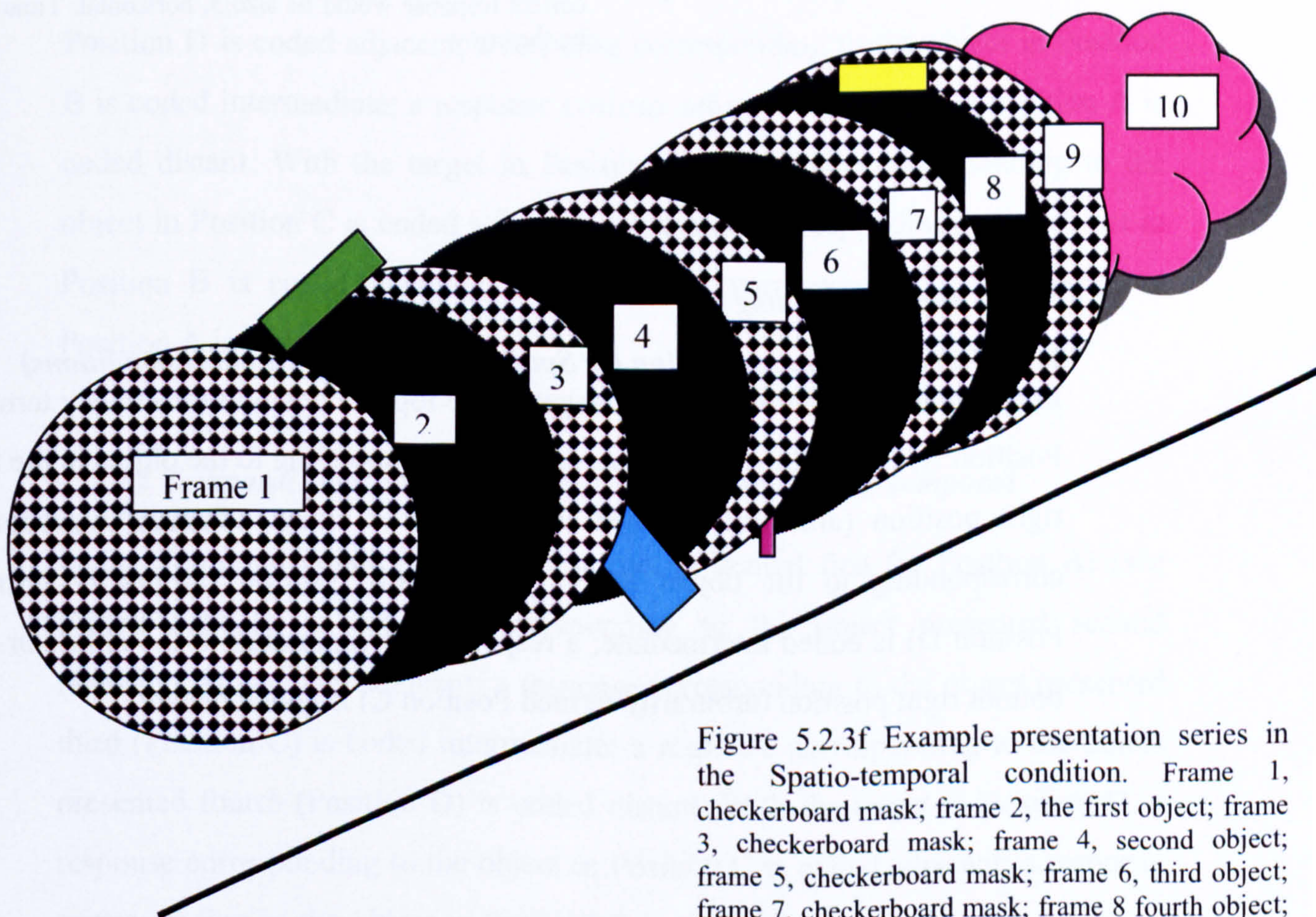


Figure 5.2.3f Example presentation series in the Spatio-temporal condition. Frame 1, checkerboard mask; frame 2, the first object; frame 3, checkerboard mask; frame 4, second object; frame 5, checkerboard mask; frame 6, third object; frame 7, checkerboard mask; frame 8 fourth object; frame 9, checkerboard mask; frame 10, post-cue. In this example, the correct response would be size 4 (smallest), vertical. Time is indicated by the arrow.

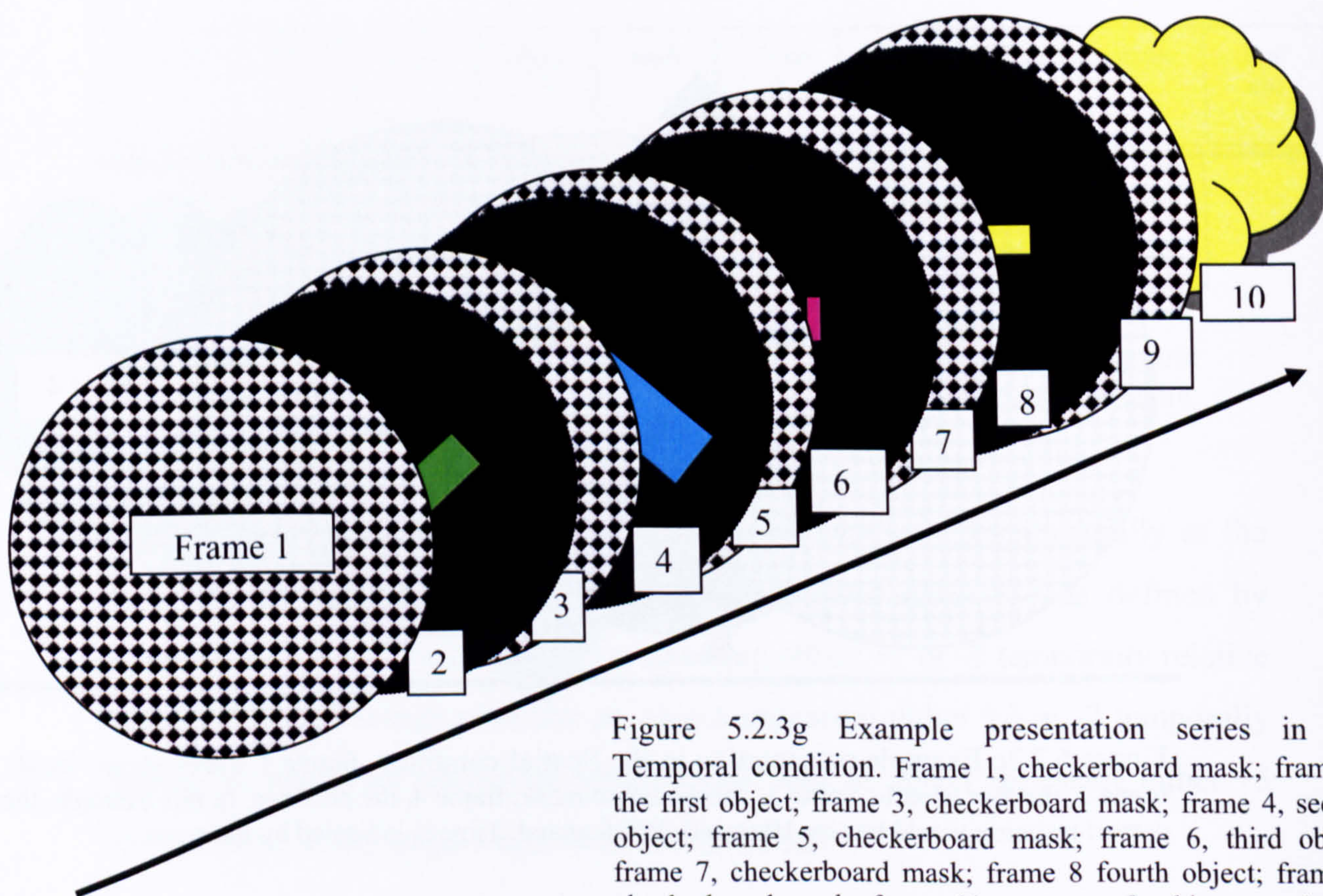


Figure 5.2.3g Example presentation series in the Temporal condition. Frame 1, checkerboard mask; frame 2, the first object; frame 3, checkerboard mask; frame 4, second object; frame 5, checkerboard mask; frame 6, third object; frame 7, checkerboard mask; frame 8 fourth object; frame 9, checkerboard mask; frame 10, post-cue. In this example, the correct response would be size 3, horizontal. Time is indicated by the arrow

5.2.4 Response coding

5.2.4.1 Spatial response coding (in Spatial and Spatio-temporal conditions)

In the Spatial condition with the target in the top left position (arbitrarily termed Position A) (see Figure 5.2.4.1), a response corresponding to the object in the top right position (arbitrarily termed Position B) is coded adjacent; a response corresponding to the object in the bottom left position (arbitrarily termed Position D) is coded intermediate; a response corresponding to the object in the bottom right position (arbitrarily termed Position C) is coded distant.

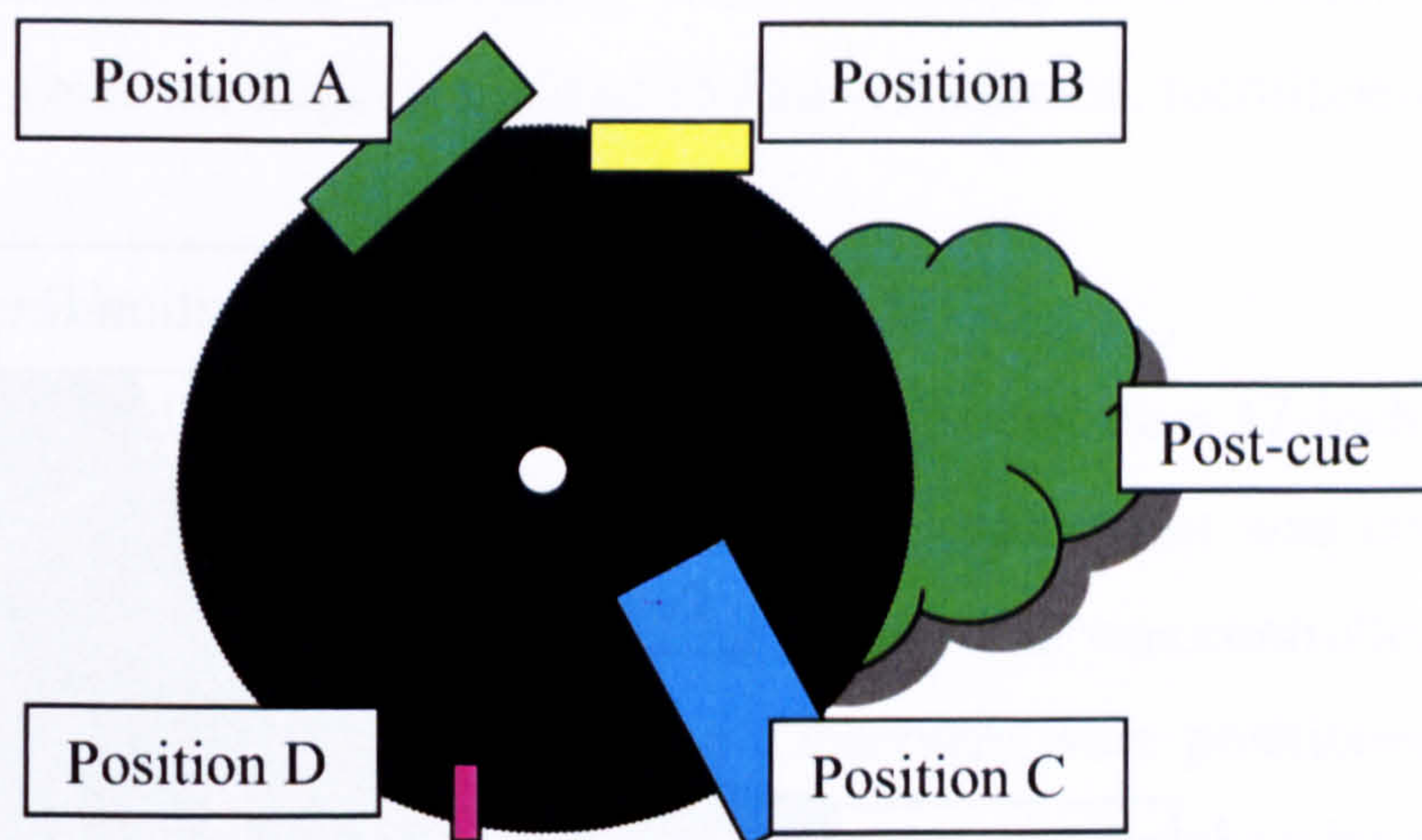


Figure 5.2.4.1 Example of response coding the Spatial condition with the target in the top left position (A).

With the target in Position B a response corresponding to the object in Position A is coded adjacent; a response corresponding to the object in Position C is coded intermediate; a response corresponding to the object in Position D is coded distant. With the target in Position C, a response corresponding to the object in Position D is coded adjacent; a response corresponding to the object in Position B is coded intermediate; a response corresponding to the object in Position A is coded distant. With the target in Position D, a response corresponding to the object in Position C is coded adjacent; a response corresponding to the object in Position B is coded intermediate; a response corresponding to the object in Position A is coded distant.

5.2.4.2 Temporal response coding (in Temporal and Spatio-temporal conditions)

In the Temporal condition, with the target presented first (in Position A) (see Figure 5.2.4.2), a response corresponding to the object presented second (Position B) is coded adjacent; a response corresponding to the object presented third (Position C) is coded intermediate; a response corresponding to the object presented fourth (Position D) is coded distant. With the target in Position D, a response corresponding to the object in Position C is coded adjacent; a response corresponding to the object in Position B is coded intermediate; and a response corresponding to the object in Position A is coded distant.

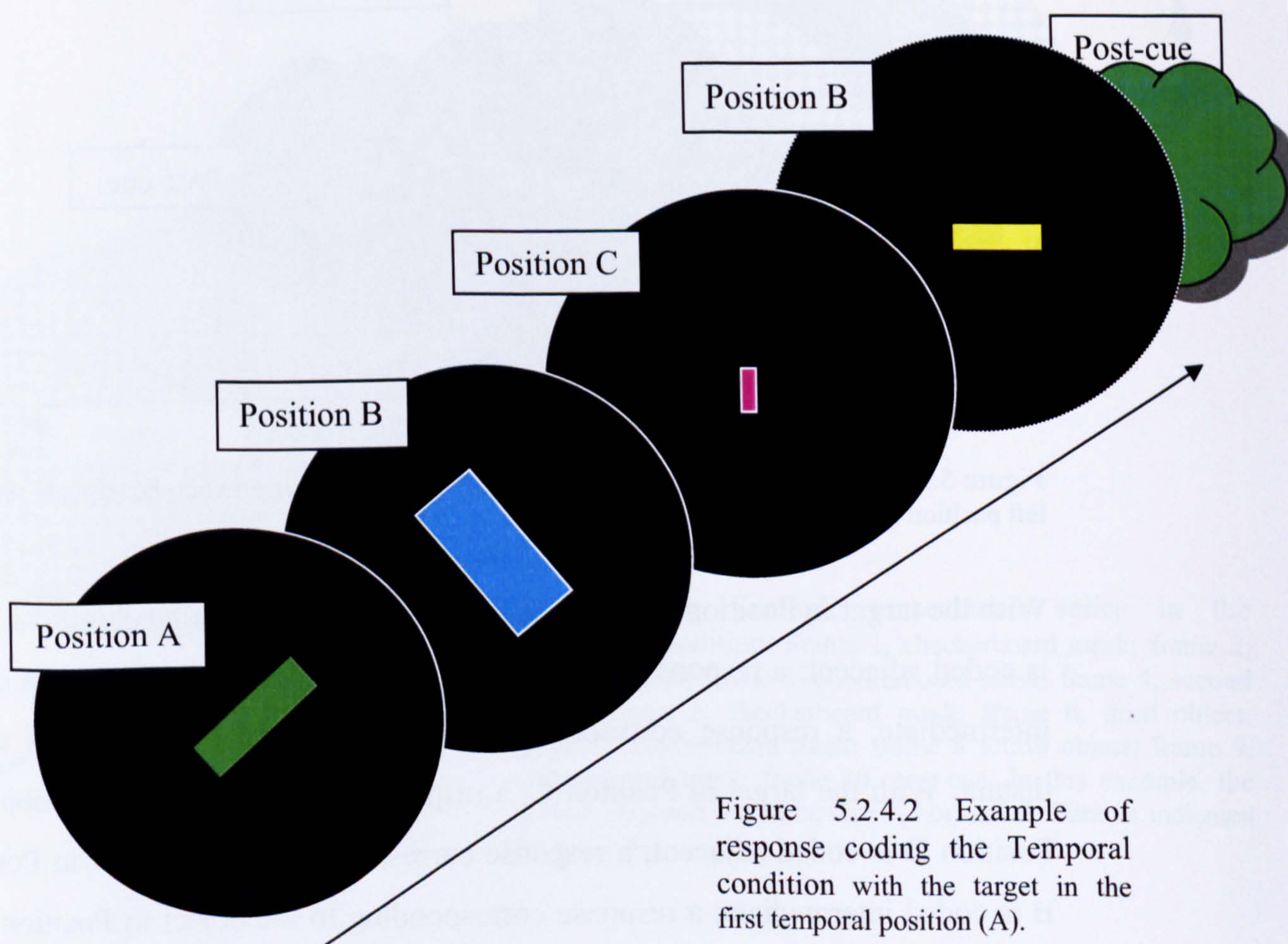


Figure 5.2.4.2 Example of response coding the Temporal condition with the target in the first temporal position (A).

With the target in Position B or Position C, no distant error can be recorded. Thus with the target in Position B a response corresponding to the object in Positions A and C are coded adjacent; a response corresponding to the object in Position D is coded intermediate. With the target in Position C, a response corresponding to the object in Positions B and D are coded adjacent; a response corresponding to the object in Position A is coded intermediate. Thus there exists an unequal chance of making each type of error: this is determined as a 50% chance of an adjacent error, a 33.3% chance of an intermediate error and a 16.7% of a distant error.

5.2.3 Participants

Fifteen adult observers with normal or corrected-to-normal vision volunteered to take part in this study. This opportunity sample was drawn from the population

of staff at Bournemouth University who responded to an email request for participants. Once the target sample of 15 had volunteered, recruitment ceased.

5.2.4 Materials

The stimuli were displayed against a black background on a 17-inch CRT, Dell PC monitor with a refresh rate of 75 Hz. The experiment was controlled by Macromedia Director 7 software. Lighting in the room was controlled by closed Venetian blinds and fluorescent lighting. Observers were positioned such that their faces were 15 inches from the monitor screen, and the chair height adjusted so that the focal point on the display was at the centre of their visual field.

5.2.5 Protocol

Individual participants sat facing the monitor with their chair adjusted so that the display focal point (marked with a white dot) was at the centre of their visual field. Participants were told that they were free to leave the study at any time and that their results would be anonymous. They were asked if they had any visual problems such as colour blindness or if they experienced adverse reaction to flashing lights. They were also told that response time was not being measured. One volunteer reported colour blindness and consequently he was excluded from participating in the study and another participant was recruited. Instructions were read aloud to each participant whilst simultaneously appearing on screen. When the participants indicated that they were ready, the trial began.

In order to establish a 50% error rate for each participant in each condition, the staircase procedure (Cornsweet, 1962) was adopted (Section 5.2.1). The order of condition was randomized in an attempt to eliminate confounding variables such as maturity or fatigue. The first 4 presentations were displayed at 2 f/sec., the next 4 presentations were presented at 3 f/sec., the next at 4 f/sec. and so on until the maximum 9 f/sec. was reached at 33rd-36th presentations. The presentations were then displayed at a decreasing rate of presentation, until the 61st-64th presentation at 2 f/sec. This procedure was repeated for the remaining two conditions. Responses were collected on a check sheet in order to establish which frame rate would best achieve a 50% error rate for that condition. Please see Section 5.4.1 which provides details of individual participants' preferred SOAs. Having established the f/sec. presentation rate for

each condition, the experiment proper began. Again the order of conditions was randomized.

5.3 HYPOTHESES

Twelve main hypotheses (and 12 sub-hypotheses) are tested by means of this experimental study. The first four null hypotheses are concerned with target proximity: that is target proximity has no effect on object feature binding. The fifth null hypothesis is concerned with the temporal effect: that is temporal properties have no effect on object feature binding. The sixth null hypothesis is concerned with the spatial effect: that is spatial properties have no effect on temporal object feature binding. The seventh to tenth null hypotheses are concerned with proximity and illusory conjunctions: that is proximity has no effect on illusory conjunctions (ICs). The eleventh null hypothesis is concerned with the temporal effect on illusory conjunctions: that is temporal properties have no effect on illusory conjunctions (ICs). The final null hypothesis is concerned with the spatial effect on illusory conjunctions: that is spatial properties have no effect on illusory conjunctions (ICs).

H₀1i: The proportion of errors in each position reported by observers will not differ from a chance proportion (33.3%) in the Spatial condition.

H1i: In support of the Spatial Theory, the alternative hypotheses state that, in the Spatial condition the observed errors will differ from expected.

H₀1ii: The proportion of adjacent errors in each position reported by observers will exceed the proportions of intermediate or distant errors in the Spatial condition.

H1ii: In support of the Spatial Theory, the alternative hypotheses state that, in the Spatial condition there will be more adjacent errors than intermediate or distant.

H₀2i: The proportion of spatial errors in each position reported by observers will not differ from a chance proportion (33.3%) in the Spatio-temporal condition.

H2i: In support of the Spatial Theory, the alternative hypotheses state that, in the Spatio-temporal condition the observed errors will differ from expected.

H₀2ii: The proportion of adjacent errors in reported by observers will exceed the proportions of intermediate or distant errors in the Spatio-temporal condition.

H2ii: In support of the Spatial Theory, the alternative hypotheses state that, in the Spatio-temporal condition there will be more adjacent errors than intermediate or distant.

H₀3i: The proportion of temporal errors reported by observers in the Spatio-temporal condition will not differ from chance proportion (50% for adjacent errors, 33.3% for intermediate errors, and 16.7% for distant errors).

H3i: In support of the Temporal Theory, the alternative hypothesis states that in the Spatio-temporal condition, the observed errors will differ from expected.

H₀3ii: The proportion of temporally adjacent errors in reported by observers will exceed the proportions of intermediate or distant errors in the Spatio-temporal condition.

H3ii: In support of the Temporal Theory, the alternative hypotheses state that, in the Spatio-temporal condition there will be proportionately more adjacent errors than intermediate or distant.

H₀4i: The proportion of temporal errors reported by observers in the Temporal condition will not differ from chance proportion for each position (50% for adjacent errors, 33.3% for intermediate errors, and 16.7% for distant errors).

H4i: In support of the Temporal Theory, the alternative hypothesis states that in the Temporal condition, the observed errors will differ from expected.

H₀4ii: The proportion of temporally adjacent errors in reported by observers will exceed the proportions of intermediate or distant errors in the Temporal condition.

H4ii: In support of the Temporal Theory, the alternative hypotheses state that, in the Temporal condition there will be proportionately more adjacent errors than intermediate or distant.

H₀5i: There will be no differences in the proportion of spatial and temporal errors in the Spatio-temporal conditions.

H₀5i: In support of the Spatial Theory, there will be differences in the proportion of spatial and temporal errors in the Spatio-temporal condition.

H5ii: There will be differences in the proportion of spatial and temporal errors in the Spatio-temporal condition.

H5ii: In support of the Spatial Theory, because of the effect of spatial and temporal properties on temporal and spatial feature binding, the alternative hypothesis states that there will be more spatially than temporally adjacent errors in the Spatio-temporal condition.

H₀6i: There will be no differences in the proportion of temporal errors in the Temporal and Spatio-temporal conditions.

H6i: In support of the Spatial Theory, there will not be more adjacent errors in the Spatio-temporal than the Temporal condition.

H₀6ii: There will be a difference in the proportion of temporal errors in the Temporal and Spatio-temporal conditions.

H6ii: In support of the Spatial Theory, because of the effect of spatial and temporal properties on temporal and spatial feature binding, the alternative hypothesis states that there will be more adjacent errors in the Spatio-temporal than the Temporal condition.

H₀7i: The proportion of ICs in each position reported by observers will not differ from a chance proportion (33.3%) in the Spatial condition.

H7i: In support of the Spatial Theory, the alternative hypothesis states that the largest proportion of ICs will be adjacent in the Spatial condition.

H₀7ii: The proportion of adjacent ICs in reported by observers will exceed the proportions of intermediate or distant ICs in the Spatial condition.

H7ii: In support of the Spatial Theory, the alternative hypotheses state that, in the Spatial condition there will be more adjacent ICs than intermediate or distant.

H₀8i: The proportion of spatial ICs in each position reported by observers will not differ from a chance proportion (33.3%) in the Spatio-temporal condition.

H8i: In support of the Spatial Theory, the alternative hypotheses state that, in the Spatio-temporal condition the observed ICs will differ from expected.

H₀8ii: The proportion of adjacent ICs in reported by observers will exceed the proportions of intermediate or distant ICs in the Spatio-temporal condition.

H8ii: In support of the Spatial Theory, the alternative hypotheses state that, in the Spatio-temporal condition there will be more adjacent ICs than intermediate or distant.

H₀9i: The proportion of temporal ICs reported by observers in the Spatio-temporal condition will not differ from chance proportion for each position. The chance proportion of temporal ICs for each position in this condition is 50% for adjacent ICs, 33.3% for intermediate ICs, and 16.7% for distant ICs.

H9i: In support of the Temporal Theory, the alternative hypothesis states that in the Spatio-temporal condition, the observed ICs will differ from expected.

H₀9ii: The proportion of temporally adjacent ICs in reported by observers will exceed the proportions of intermediate or distant ICs in the Spatio-temporal condition.

H9ii: In support of the Temporal Theory, the alternative hypotheses state that, in the Spatio-temporal condition there will be proportionately more adjacent ICs than intermediate or distant.

H₀10i: The proportion of temporal ICs reported by observers in the Temporal condition will not differ from chance proportion for each position. The chance proportion of temporal ICs for each position in this condition is 50% for adjacent ICs, 33.3% for intermediate ICs, and 16.7% for distant ICs.

H10i: In support of the Temporal Theory, the alternative hypothesis states that in the Temporal condition, the observed ICs will differ from expected.

H₀10ii: The proportion of temporally adjacent ICs in reported by observers will exceed the proportions of intermediate or distant ICs in the Temporal condition.

H10ii: In support of the Temporal Theory, the alternative hypotheses state that, in the Temporal condition there will be proportionately more adjacent ICs than intermediate or distant.

H₀11i: There will be no differences in the proportion of spatial and temporal ICs in the Spatio-temporal condition.

H11i: In support of the Spatial Theory, because of the effect of temporal properties on spatial binding, the alternative hypothesis states that the proportion

of spatial ICs will be greater than the proportion of temporal ICs in the Spatio-temporal condition.

H_{011ii}: There will be no differences in the proportion of spatial and temporal adjacent ICs in the Spatio-temporal condition.

H11ii: In support of the Spatial Theory, because of the effect of temporal properties on spatial binding, the alternative hypothesis states that there will be more spatially, than temporally adjacent ICs in the Spatio-temporal condition.

H_{012i}: There will be no differences in the proportion of temporal ICs between the Temporal and Spatio-temporal conditions.

H12i: In support of the Spatial Theory, the proportion of adjacent ICs in the Spatio-temporal condition will be greater than the proportion of ICs in the Temporal conditions.

H_{012ii}: There will be no difference between the count of adjacent ICs in the Spatio-temporal condition and in the Temporal condition.

H12ii: In support of the Spatial Theory, because of the effect of spatial and temporal properties on temporal and spatial feature binding, the alternative hypothesis states that there will be more adjacent ICs in the Spatio-temporal condition than in the Temporal condition.

In all cases H_A : The null hypothesis will be rejected at $p = 0.05$.

If H_0 is correct, for $H_1 - H_4$, the number of observations that appear in any cell of the contingency table is determined by chance. If the alternative hypothesis is correct, the observed frequency will differ from this expected frequency. These hypotheses are tested by means of χ^2 . Mean counts are compared using one-tailed *t*-tests. If H_0 is correct, for $H_5 - H_6$, the number of observations that appear in any cell of the contingency table is determined by chance. If the alternative hypothesis is correct, the observed frequency will differ from this expected frequency. These hypotheses are tested by means of χ^2 . Mean counts are compared using one-tailed *t*-tests.

If H_0 is correct, for $H_7 - H_{10}$ the number of observations that appear in any cell of the contingency table is determined by chance. If the alternative hypothesis is correct, the observed frequency will differ from this expected

frequency. These hypotheses are tested by means of χ^2 . Mean counts for ICs are compared using one-tailed *t*-tests. If H_0 is correct for $H_{11} - H_{12}$ the number of observations that appear in any cell of the contingency table is determined by chance. If the alternative hypothesis is correct, the observed frequency will differ from this expected frequency. These hypotheses are tested by means of the Chi square statistic. Mean counts are compared using one-tailed *t*-tests.

5.4 RESULTS

This section is in two parts. Part 1 is concerned with error types for each condition and describes the spatial and temporal effect (evidenced by comparing spatial and temporal errors in the spatio-temporal condition, and by comparing temporal errors in the Spatio-temporal condition with errors in the Temporal condition). Part 2 is concerned with the incidence and characteristics of illusory conjunctions (ICs). Following the presentation of results (Section 5.3), they are discussed in Section 5.5.

Section 5.3.1 provides an overview of presentation rates (Table 5.3.1). Results from the Spatial condition are given in Section 5.3.2. Spatial and temporal errors recorded in the Spatio-temporal condition are compared and described in Sections 5.4.3, 5.4.4 and 5.4.5. Results from the Temporal condition are described in Section 5.4.6 and compared with temporal errors in the Spatio-temporal condition in Section 5.4.7. The spatial condition comprised 3 frames, whereas the spatio-temporal and temporal conditions comprised 9 frames. Hence it is not possible to meaningfully compare results from the spatial with the other two conditions.

5.4.1 Presentation rates

Details of the presentation rate for each condition for each participant is given in Table 5.4.1. Presentation rates varied across conditions: mode 6 f/sec, mean 7.2 f/sec in the Spatial condition, mode 8 f/sec, mean 7.33 f/sec in the Spatio-temporal condition, and mode 7 f/sec, mean 5.8 f/sec. in the Temporal condition. Mean presentation rate (in f/sec.) was similar for the Spatial and Spatio-temporal conditions, but slower in the Temporal condition (5.8f/sec.) (Table 5.4.1b). However, the modes were not similar.

P	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
S	9	7	6	6	8	8	8	9	7	6	6	7	6	7	8
ST	9	9	7	4	5	8	7	6	8	8	7	8	8	7	9
T	4	5	2	3	8	7	5	7	7	7	7	6	7	5	7

Table 5.4.1a Individual mean frame/sec for each participant (1-15) in each condition (P=participant; S=Spatial; ST=Spatio-temporal; T=Temporal)

Although mean presentation rate was similar for Spatial and Spatio-temporal presentations, modes showed that a shorter time (8 f/sec = 125ms) was required by more participants to achieve 50% error rate in the Spatio-temporal condition, and longest time (6 f/sec = 166.66ms) in the Spatial condition.

Condition	SD	Mode f/sec	Mean f/sec	Mean ms.	frames/ presentation	Mean duration
Spatial	1.08	6	7.2	138.88	3	416.64
Spatio-temporal	1.45	8	7.33	136.25	9	1226.25
Temporal	1.74	7	5.8	172.41	9	1551.69

Table 5.4.1b Standard deviation, mode and mean presentation rate in frame/sec and ms, number of frames/ presentation and mean duration in milliseconds.

In order to achieve the desired 50% error rate, 66.7% observers (10/15) in the Spatio-temporal condition, and 46.7% (7/15) in the Spatial condition required the fastest presentation rate. Conversely, 60% (9/15) of observers required the slowest rate for Temporal presentations (Table 5.4.1a).

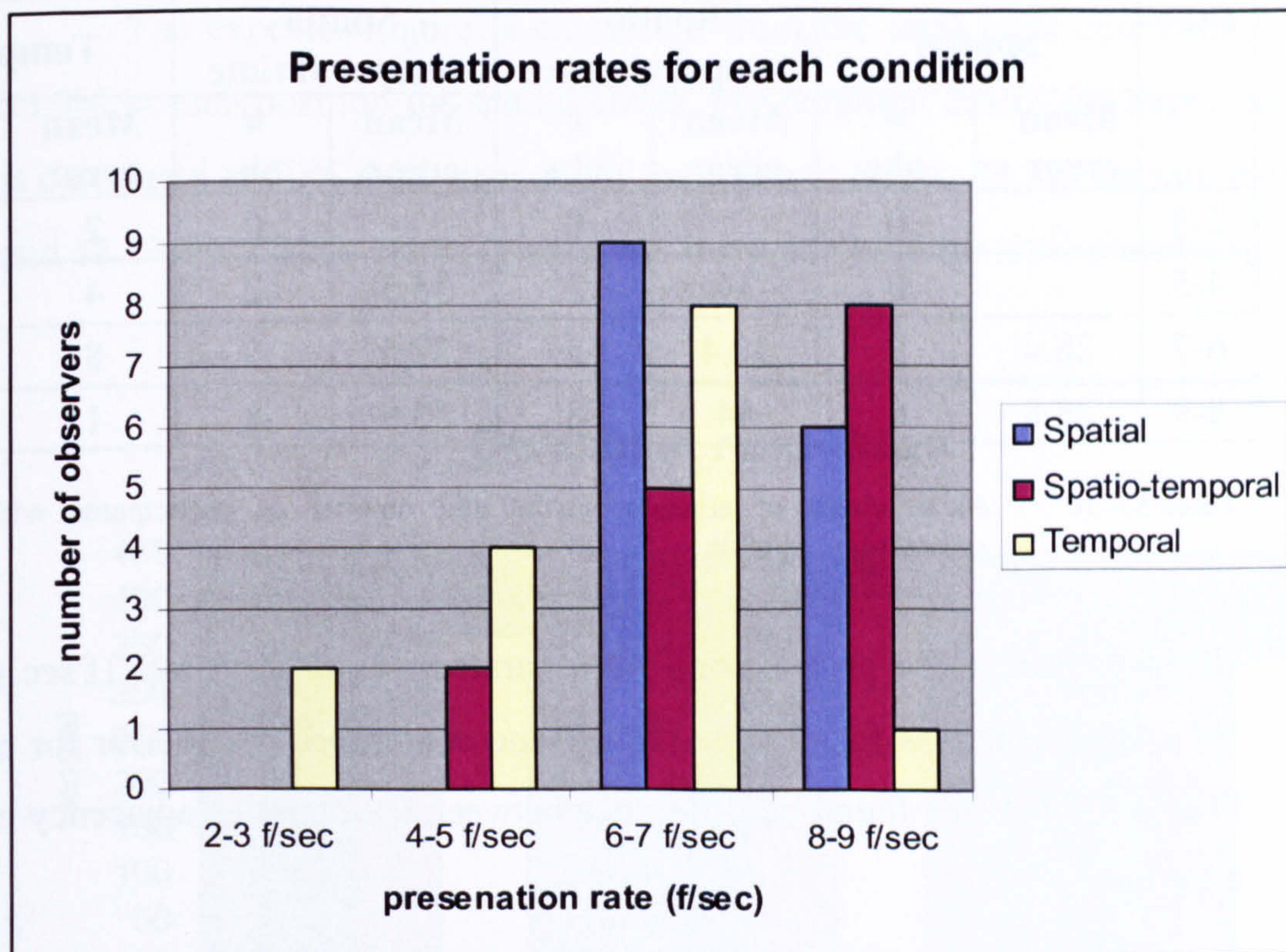


Figure 5.4.1 Presentation rate bins for each condition

The presentation rates suggest that more time was required to achieve a 50% error rate in the Spatial than other conditions. However, SD was lowest in the Spatial condition and greatest in the temporal condition where one participant required 500ms per presentation, and another required 3333ms. These slow rates were not required in either other condition. However, these rates correspond to responses in which 50% were incorrect. It may well be that these participants achieved a greater correct response with faster presentation rates as individual observers on occasion required a faster or lower presentation rate to achieve their individual required error rate.

We questioned whether if more adjacency errors are reported with a longer presentation (that is at fewer frames per second), are binding errors are taking place in memory? In addition, if more adjacency errors are reported with shorter presentation, (more frames per second), are binding errors perceptual?

f/sec	Spatial		Spatio-temporal space		Spatio-temporal time		Temporal	
	Mean error	# obs	Mean error	# obs	Mean error	# obs	Mean error	# obs
2-3	-	0	-	0	-	0	2	31.5
4-5	-	0	34	2	35.5	2	4	32.3
6-7	28.4	9	32.4	5	32.5	5	8	32.5
8-9	28.8	6	34	8	33.9	8	1	32

Table 5.4.1c Mean count of adjacent errors and number of participants within each presentation rate bin

To investigate these propositions we separated presentation rate (f/sec.) (Table 5.4.1) into four bins for each condition, and considered mean error for each bin (Table 5.4.1c). We found no difference between the count of adjacency errors in each bin.

5.4.2 Overall errors

This section provides an overview of results in all conditions. Subsequent subsections consider each condition separately.

For each condition, 1920 responses (15 participants, 64 presentations, 2 features) were recorded. With error rate preset to ~50%, one would expect 3840 errors (960 in each condition). This target was achieved almost exactly in all conditions and suggests that the ‘staircase method’ was successful. Figure 5.4.2a shows total combined responses in each condition. Most double correct responses (351) were recorded in the Spatial condition, least in the Temporal condition (304). The Spatio-temporal condition resulted in more responses in which both features were erroneously reported (359), again least in the Temporal condition (331). In the Temporal condition, observers reported more errors in which response for one feature was correct and one was wrong (325). This compares with 250 in the Spatial condition and 231 in the Spatio-temporal condition.

Erroneous responses could be adjacent, intermediate or distant in relation to the target. Table 5.4.2a shows the total observed and expected errors (rounded to nearest whole number) in all positions in all conditions.

The expected figure is calculated from the total error count divided by 33.3% in each position for spatial errors. For temporal errors, the expected error is calculated as 50% total error count for adjacent errors, 33.3% for intermediate, and 16.7% for distant errors (please see Section 5.2.4.2 for an explanation).

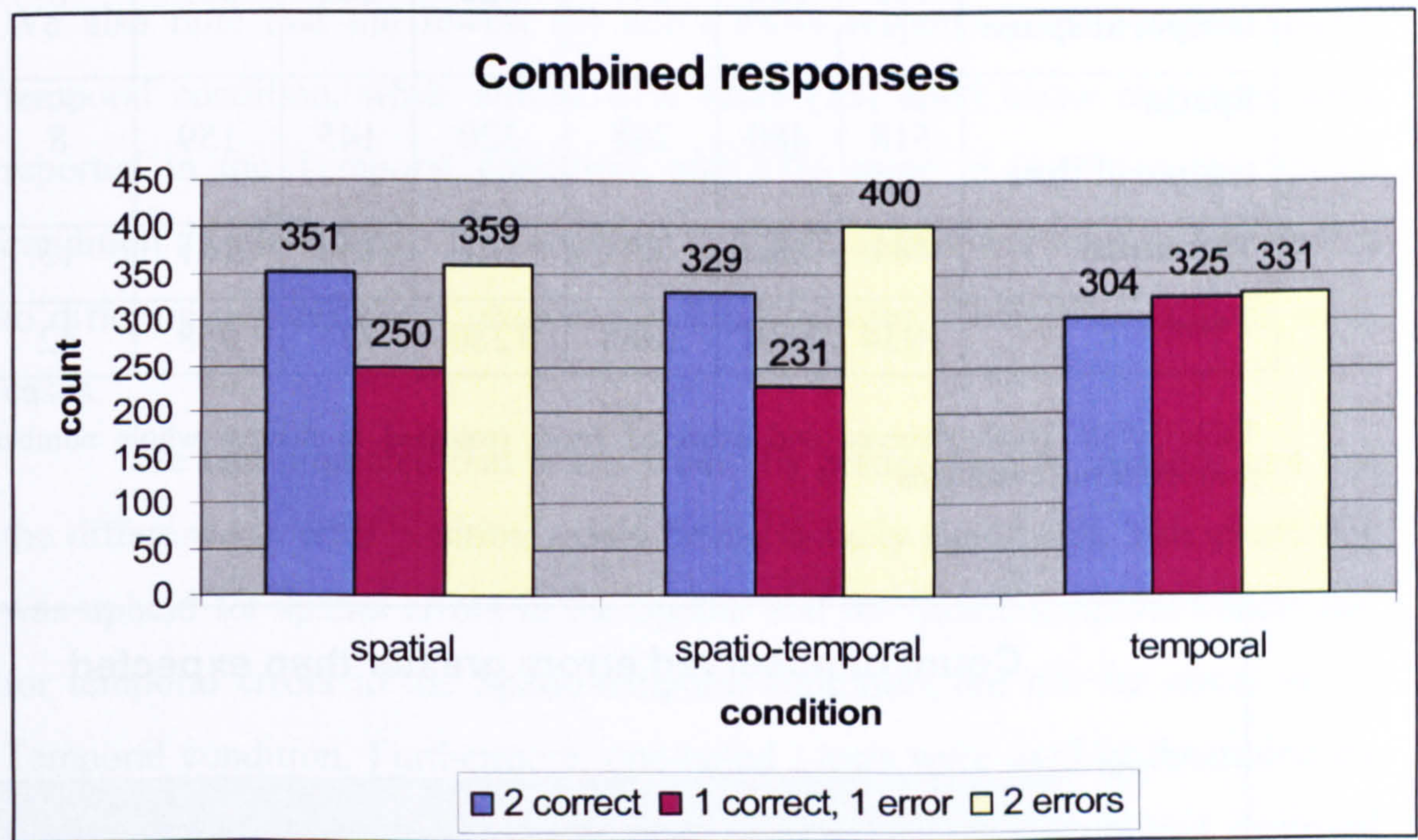


Figure 5.4.2a Total combined responses in each condition (2 correct, a combination of 1 correct and 1 error, and 2 errors)

Although the ‘staircase’ method was successful in achieving the desired error rate, difference in position of errors in each condition is apparent. For example, observed spatially adjacent errors exceeded expected adjacent errors by 111 in the Spatial condition, and by 178 in the Spatio-temporal condition, whereas observed temporally adjacent errors exceeded expected adjacent errors by 38 in the Spatio-temporal condition, and by 2 in the Temporal condition (Fig. 5.4.2a).

Because the same number of errors was not expected in each position in spatial and temporal conditions, we are not interested in the count of errors *per se*, but in the count of errors greater than expected (Figure 5.4.2b). It is interesting to note that the greatest differences were in the spatial conditions: Spatial and Spatio-temporal space (spatial errors in the Spatio-temporal condition), and the least in the temporal conditions: Temporal and Spatio-temporal time (temporal errors in the Spatio-temporal condition).

	Adjacent		Intermediate		Distant		Don't know	Total
	Obs	Exp	Obs	Exp	Obs	Exp		
Spatial	430	319	231	319	285	319	11	957
Spatio-temporal space	502	320	249	320	200	319	8	959
Spatio-temporal time	518	480	288	320	145	159	8	959
Temporal	484	482	320	321	144	161	15	963
Total	1934	1601	1088	1280	774	958	42	3838

Table 5.4.2a Total observed and expected errors (rounded to nearest whole number) in all positions in all conditions

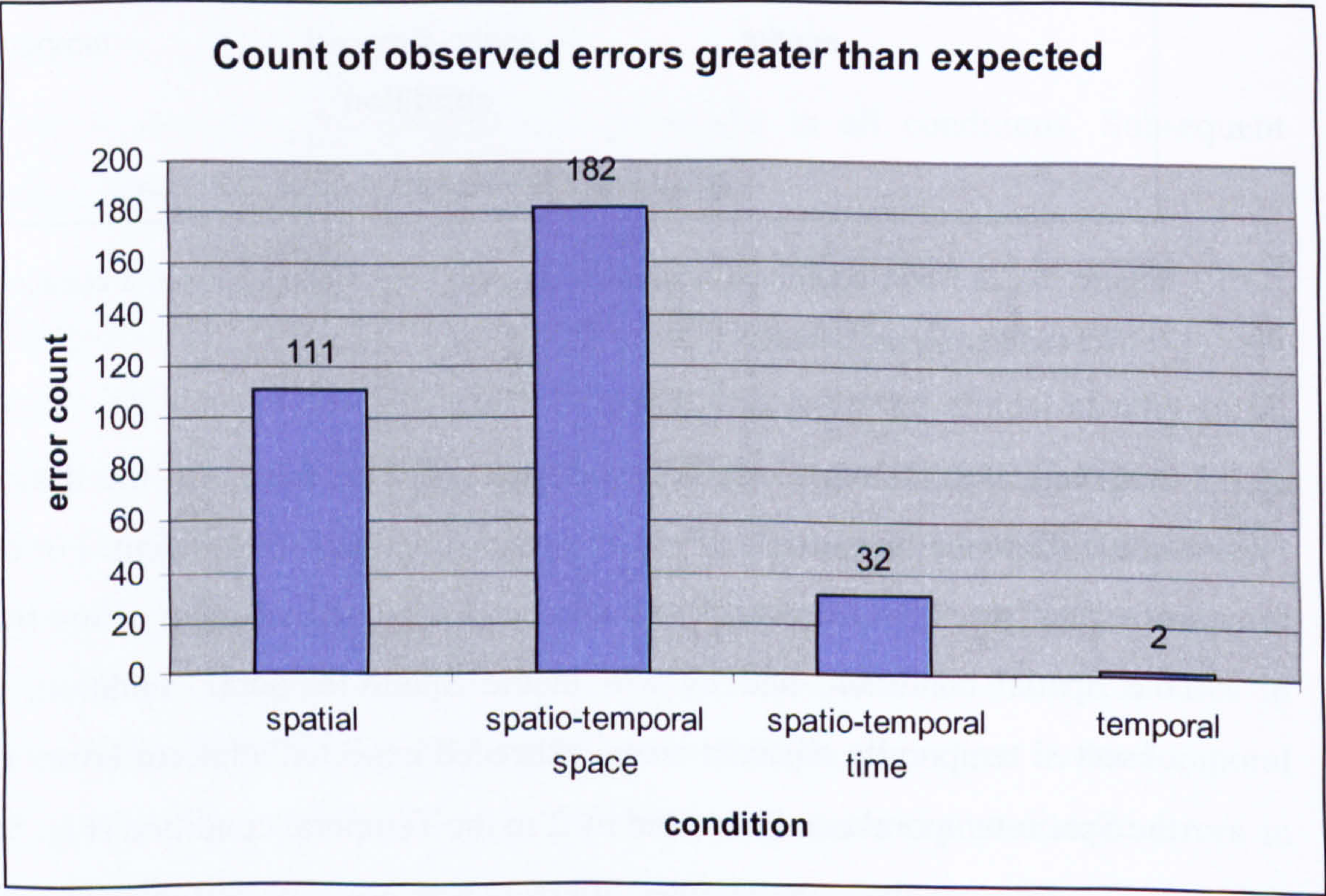


Figure 5.4.2b Count of observed errors greater than expected in each condition

We cannot compare errors in the Spatial condition with those in other conditions because, as stated earlier, there were fewer frames (3) in each complete presentation in the Spatial condition and this might have affected recall. However, we can compare spatial and temporal errors in the Spatio-temporal condition as these data were collected simultaneously, and temporal errors in

Temporal and Spatio-temporal conditions as these data resulted from the same number of frames (9) in each condition. Consequently we note the differences between intermediate and distant errors in the Spatio-temporal condition (discussed in Section 5.4.7); and the differences between temporally adjacent and intermediate errors in Spatio-temporal and Temporal conditions (Section 5.4.8). We also note that the fewest (8) don't know responses were in the Spatio-temporal condition, while almost 47% more (15) don't know responses were reported in the Temporal condition, and 27% more in condition the Spatial condition (Table 5.4.2a). This suggests that both spatial and temporal properties to differing degrees allow observers to bind features, albeit incorrectly in some cases.

We had predicted that errors would be predominantly adjacent, and that the difference in error position would be statistically significant. This prediction was upheld for spatial errors in the Spatial and the Spatio-temporal conditions; for temporal errors in the Spatio-temporal condition; but not for errors in the Temporal condition. Furthermore, one-tailed *t-tests* were used to determine the significance of differences between means of participants' errors and predicted errors (Table 5.4.2b.).

CONDITION	Adjacent	Intermediate	Distant
Spatial	<0.001*	<0.001*	0.065
Spatio-temporal space	<0.001*	0.002*	0.001*
Spatio-temporal time	0.003*	0.002*	0.067
Temporal	0.312	0.465	0.123
Spatio-temporal space--- Spatio-temporal time	0.510	0.147	0.05*
Spatio-temporal--- Temporal	0.008*	0.051	0.949

Table 5.4.2b P values from one-tailed *t-tests* for differences between means of observed and expected errors in each position in each condition. Significant differences are marked *. Shaded cells indicate observed exceeded expected.

Having given an overview of results in general, the following subsections detail results for each condition. Firstly we consider errors for each feature.

Because of the successful implementation of the staircase method in achieving 50% error rate for each feature, there was no significant difference between the proportion of errors for pairs of features for each position in any condition except between size and orientation adjacent errors in the Temporal condition ($p = 0.012$; $t = 2.902$). Looking at proportions of errors on each feature, we find that adjacent errors exceeded expected for each feature in all conditions except for size in the Temporal condition. This could be due to the fact that there were no spatial properties in this condition and relative size was more difficult to discriminate when there is no unique spatial position. Unique spatial locations may have allowed observers to ‘place’ the objects spatially and discriminate more easily in recall. Observed intermediate and distant errors exceeded expected only in the case of intermediate orientation errors (Table 5.3.2c). The greatest difference between features was for adjacent temporal errors. The proportion of orientation errors was 4.5% greater than size errors (the proportion of size errors was 2% fewer than expected, while the proportion of orientation errors was 2.5% greater than expected). These results suggest that without unique spatial properties, the feature ‘orientation’ is more likely than the feature ‘size’, to migrate to an adjacent object.

Proportions of observed errors above (+) and below (-) the expected proportions are tabulated in Table 5.4.2c. The greatest difference between features was for adjacent temporal errors.

POSITION	FEATURE	SPATIAL		TEMPORAL	
		spatial	spatio-temporal	spatio-temporal	temporal
Adjacent	size	+11.3	+20.1	+2.9	-2
	orientation	+12	+18	+5.2	+2.5
Intermediate	size	-7.6	-8	-2.2	-0.5
	orientation	-10.8	-6.6	-4.3	+0.3
Distant	size	-4.8	-12.8	-1.5	-1.5
	orientation	-2	-11.9	-1.7	-4.5

Table 5.4.2c Feature errors in each position for each condition in percentage difference from expected.

The greatest proportional difference for orientation and size features was in the Temporal condition (4.5%). These results suggest that without unique spatial properties, the feature 'orientation' is more likely than the feature 'size', to migrate to an adjacent object.

5.4.3 *The Spatial condition*

Presentation speed was set to achieve 50% errors on each feature (480 for each, 960 in total). Observers had an equal chance (33%) of erroneously selecting each position. There were 957 observed errors in the Spatial condition. Of these 250 combined errors²¹ contained a correct response to 1 feature, and 359 were double errors. Adjacent errors exceeded expected for both features. A goodness-of-fit test, the Chi-square statistic, was used to test if errors were recorded at random giving 33% in each position. 'Don't know' responses have been removed from the analysis and from the total error count in order to calculate the statistic. Therefore, with 2 degrees of freedom, and $\alpha = 0.05$, the critical value for Chi-square is 5.99. Because the obtained value is 77.79, H_{II} is accepted: features were not equally reported in all positions in the Spatial condition.

A comparison of observed and expected errors of the feature 'size' in each position reveals that the proportion of adjacent errors was greater than expected (+11.3%), but less than expected for intermediate (-7.6%) and distant errors (-4.8%). When comparing observed and expected errors of the feature 'orientation' in each position, the proportion of adjacent errors was greater (+12%), and less for intermediate (-10.8%) and distant (-2%) errors than expected. Overall, the proportion of observed adjacent errors was 11.6% greater than expected, while there were 9.2% fewer observed intermediate errors and 3.5% fewer observed distant errors than expected (see Table 5.4.3). H_{III} , which predicts that there will be more adjacent than distant errors is tested by means of one-tailed *t*-tests. Results were statistically highly significant ($p < 0.001$; $t = 5.638$). Hence H_{III} is accepted.

²¹ A combined response comprises 2 single responses: 1 for each feature (size and orientation).

Feature	Position							
	Adjacent		Intermediate		Distant		DK	Total
	Obs	Exp	Obs	Exp	Obs	Exp	Obs	-
Size count	213	159.3	123	159.3	136	159.3%	6	478
Size proportion	44.6%	33.3%	25.7%	33.3%	28.5%	33.3%	1.2%	-
Orientation count	217	159.67	108	159.67	149	159.67	5	479
Orientation proportion	45.3%	33.3%	22.5%	33.3%	31.1%	33.3%	1.1%	-
Total count	430	319	231	319	285	319	11	957
Total proportion	44.9%	33.3%	24.1%	33.3%	29.8%	33.3%	1.2%	-
Mean	28.33	21.33	15.4	21.33	19.7	21.33	0.8	-
Standard deviation	4.91	-	4.79	-	3.01	-	1.78	-

Table 5.4.3 Total observed (obs), expected (exp) and proportion of errors in the Spatial condition for each feature: size and orientation. Mean error and standard deviation are shown below totals.

These findings support the Spatial Theory which suggests that if spatial properties are primary a greater number of errors would be adjacent as features of one object would migrate to a spatially adjacent object rather than to a spatially distant object.

5.4.4. Spatio-temporal (spatial errors) condition

There were 959 observed errors in the Spatio-temporal condition. Of these, 231 combined errors contained a correct response to 1 feature, and 400 were double errors. Adjacent errors exceeded expected for both features. Observers had an equal chance of erroneously selecting each position. Hence the expected proportion of errors was 33% in each position. Errors on both features greatly exceeded expected in the adjacent position. If errors were recorded at random 33% errors would be in each position. The procedure for calculating the Chi-square statistic is as described in Section 5.4.3. The obtained value is 164.8, thus

H_{02i} can be rejected: features were not reported in all spatial positions equally in the Spatio-temporal condition.

Feature	Position							
	Adjacent		Intermediate		Distant		DK	Total
	Obs	Exp	Obs	Exp	Obs	Exp	Obs	-
Size count	258	161	122	161	99	161	4	483
Size proportion	53.4%	33.3%	25.3%	33.3%	20.5%	33.3%	0.8%	-
Orientation count	244	161	127	161	101	159.67	4	476
Orientation proportion	51.3%	33.3%	26.7%	33.3%	21.4%	33.3%	0.6%	-
Total count	502	159.67	249	159.67	200	159.67	8	959
Total proportion	52.3%	33.3%	26%	33.3%	21%	33.3%	0.8%	-
Mean	33.47	21.3	16.6	21.3	13.4	21.3	0.47	-
Standard deviation	4.85	-	5.21	-	3.4	-	0.99	-

Table 5.4.4 Total observed (obs), expected (exp) and proportion of spatial errors in the Spatio-temporal condition for each feature: size and orientation. Mean error and standard deviation are also shown.

A comparison of observed and expected errors of the feature ‘size’ in each position reveals that the proportion of adjacent errors was greater (+20.1%) than expected, but less for intermediate (-8%) and distant (-12.8%) errors of the feature ‘size’ (33%).

When comparing observed and expected errors of the feature ‘orientation’ in each position, the proportion of adjacent errors was greater (+18%), but less for and fewer intermediate (-6.6%) and distant (-11.9%) errors of the feature ‘orientation’ than expected (33%). Overall, the proportion of

observed adjacent errors was 19% greater than expected, while there were 7.3% fewer observed intermediate errors and 12.3% fewer observed distant errors than expected (see Table 5.3.4). H_{2ii} , which states that in the Spatio-temporal condition, there will be more spatial errors in the adjacent position is tested by means of one-tailed *t-tests*. Results are statistically highly significant ($p < 0.00$; $t = 11.14$). Hence H_{2ii} is accepted.

The Spatial Theory suggests that if spatial properties are important a greater number of spatial errors would be adjacent as features of one object would migrate to a spatially adjacent object rather than to a spatially distant object. Findings in this condition, as in the Spatial condition (Section 5.4.3), support the Spatial Theory. However, in contrast to the presentations in the Spatial condition, observers were exposed to temporal properties also. The result was an increased difference between observed and expected spatially adjacent errors in the Spatio-temporal condition in than in the Spatial condition in which only spatial properties were available. Therefore the Spatial Theory is upheld.

5.4.5 *Spatio-temporal (temporal errors) condition*

This section describes results for temporal errors in the Spatio-temporal condition. 959 errors were recorded, of these, 231 combined errors contained a correct response to 1 feature, and 400 were double errors.

Because of the characteristics of the sequential presentation (see Section 5.2.4), observers had an unequal chance of erroneously selecting each position. Hence the expected proportion of errors was 50% in the adjacent position, 33.3% in the intermediate position, and 16.7% in the distant position. Adjacent errors exceeded expected for both features (Table 5.4.5). The procedure used to calculate the Chi-square statistic is described in Section 5.4.3. The obtained value is 7.6, thus H_{03i} can be rejected: temporal errors were not equally reported in the Spatio-temporal condition. In order to perform a one-tailed *t-test*, data were normalized.

A comparison of observed and expected errors of the feature 'size' in each position reveals that the proportion of adjacent errors was greater (+2.9%) and fewer intermediate (-2.2%) and distant (-1.5%) errors of the feature 'size' than expected. When comparing observed and expected errors of the feature 'orientation' in each position, the proportion of adjacent errors was greater

(+5.2%) and fewer intermediate (-4.3%) and distant (-1.7%) errors of the feature ‘orientation’ than expected.

Feature	Position							
	Adjacent		Intermediate		Distant		DK	Total
	Obs	Exp	Obs	Exp	Obs	Exp	Obs	-
Size count	257	243	151	162	74	81	4	486
Size proportion	52.9%	50%	31.1%	33.3%	15.2%	16.7%	0.8%	-
Orientation count	261	243	137	162	71	81	4	473
Orientation proportion	55.2%	50%	29%	33.3%	15%	16.7%	0.8%	-
Total count	518	479.5	288	319.7	145	159.83	8	959
Total proportion	54%	50%	30%	33.3%	15.1%	16.7%	0.8%	-
Mean	34.53	32	19.2	21.3	9.67	10.6	0.53	-
Standard deviation	2.92	-	3.49	-	2.55	-	1.19	-

Table 5.4.5 Total observed (obs), expected (exp) and proportion of temporal errors in the Spatio-temporal condition for each feature: size and orientation. Mean error and standard deviation are also shown.

Overall, the proportion of observed adjacent errors was 4% greater than expected, while there were 3.3% fewer observed intermediate errors and 1.6% fewer observed distant errors than expected (see Table 5.4.4). H_{3ii} , which states that in the Spatio-temporal condition, there will be more temporal errors is tested by means of one-tailed *t*-tests. Results were highly statistically significant ($p<0.001$; $t = 10.301$). Hence H_{3ii} is accepted.

The Temporal Theory suggests that temporal properties are important in object feature binding. Hence, support for the Temporal Theory would predict that a greater number of feature errors would be reported adjacent to the target object as features of one object would migrate to a temporally adjacent object

rather than to a temporally distant object. Findings in this condition support the Temporal Theory.

5.4.6 The Temporal condition

There were 963 reported errors in this condition. Of these, 325 combined errors contained 1 correct response, and there were 331 double errors. The procedure for calculating the Chi-square statistic is described in Section 5.4.3. The obtained value is 1.5, hence H_{04i} which states that in the Temporal condition, the proportions of observed errors will not differ from expected, is accepted.

A comparison of observed and expected errors of the feature ‘size’ in each position reveals that the proportion of adjacent, intermediate and distant errors was fewer (-2%, -0.5%, and -1.5% respectively) than expected.

Feature	Position							
	Adjacent		Intermediate		Distant		DK	Total
	Obs	Exp	Obs	Exp	Obs	Exp	Obs	-
Size count	231	240.5	158	160.3	85	80.2	7	481
Size proportion	48%	50%	32.8%	33.3%	15.2%	16.7%	1.5%	-
Orientation count	253	241	162	160.7	59	80.3	8	482
Orientation proportion	52.5%	50%	33.6%	33.3%	12.2%	16.7%	1.7%	-
Total count	484	481.5	320	321	144	160.5	15	963
Total proportion	50.2%	50%	33.2%	33.3%	15%	16.7%	1.6%	-
Mean	32.27	32.1	21.33	21.4	9.6	10.7	1	-
Standard deviation	2.05	-	3.54	-	3.68	-	1.56	-

Table 5.4.6 Total observed (obs), expected (exp) and proportion of errors in the Temporal condition for each feature: size and orientation. Mean error and standard deviation are also shown.

When comparing observed and expected errors of the feature ‘orientation’ in each position, the proportion of adjacent and intermediate errors was greater (+2.5% and +0.3% respectively), but less for distant (-4.5%) errors than expected.

Overall, the proportion of observed adjacent errors and intermediate errors was 0.2% greater than expected, while there were 0.1% fewer and 1.7% fewer observed distant errors than expected (see Table 5.3.6). H_{4ii} , which states that in the Temporal condition, there will be more temporally adjacent errors is tested by means of one-tailed t -tests. Results were not significant. Hence H_{04ii} is accepted. In contrast to the findings for temporal errors in the Spatio-temporal condition, these findings do not support the Temporal Theory which suggests that if temporal properties are important a greater number of errors would be adjacent as features of one object would migrate to a temporally adjacent object rather than to a temporally distant object as errors were equally distributed among the 3 non-target positions.

The previous subsections have reported individual sets of results for each condition: Spatial, Spatio-temporal (space), Spatio-temporal (time) and Temporal. The following subsections compare sets of results. Therefore, spatial and temporal errors in the Spatio-temporal condition are compared (Section 5.4.7); as are temporal errors in the Spatio-temporal and Temporal conditions (Section 5.4.8). Results from the Spatial condition cannot be compared with any other set as there were 3 frames in each complete Spatial display, whereas there were 9 frames in the other 2 conditions.

5.4.7 Comparison of spatial and temporal errors in the Spatio-temporal condition

This section compares results for spatial and temporal errors in the Spatio-temporal condition. In order to determine the effect of spatial properties on temporal feature binding, and temporal properties on spatial feature binding, results for spatial and temporal errors in the Spatio-temporal condition are compared.

In order to perform t -tests on these data, they were normalized to account for the variance in probability distribution, and ‘don’t know’ responses were

removed. Percentages have been calculated on total error (all errors plus don't know responses) 959. H_{5j} which states that there will be differences between the proportions of spatial and temporal errors in the Spatio-temporal condition, is supported. Spatially adjacent errors exceeded expected by 19%, whereas temporally adjacent errors exceeded expected by 4%. Spatially intermediate errors were 7.3% less than expected, whereas temporally intermediate errors were 3.3% less than expected. Spatially distant errors were 12.3% less than expected, whereas temporally intermediate errors were 1.6% less than expected.

Property	Adjacent		Intermediate		Distant		Total
	Obs	Exp	Obs	Exp	Obs	Exp	-
Spatial observed	502	319.7	249	319.7	201	319.7	952
Spatial proportion	52.3%	33.3%	26%	33.3%	21%	33.3%	-
Temporal observed	518	479.5	288	319.7	145	159.83	951
Temporal proportion	54%	50%	30%	33.3%	15.1%	16.7%	-

Table 5.4.7 Observed and expected spatial and temporal errors and error proportion in the Spatio-temporal condition. NB 7 spatial and 8 temporal don't know responses have been removed.

H_{5ii} which states that there will be more spatially than temporally adjacent errors in the Spatio-temporal condition is tested by means of one-tailed *t-tests* on the normalised data. Results showed highly significant differences between means for spatially and temporally adjacent errors ($p < 0.001$, $t = 7.28$). Thus H_{5ii} is accepted. The greater proportion of spatially adjacent than temporally adjacent errors reported in this condition can only have been the result of temporal properties as the data for both spatial and temporal errors were recorded simultaneously. Hence we can conclude that temporal properties increase the likelihood of spatially adjacent errors.

5.4.8 Comparison of temporal errors in the Spatio-temporal condition and the Temporal condition

This section compares results for temporal errors in the Spatio-temporal condition and the Temporal condition. Table 5.3.8 shows the total number of observed and expected errors in each position. H_{0i} states that the proportions of temporal errors in each position will be different between in the Spatio-temporal condition when compared with errors in the Temporal condition. of spatial errors will exceed temporal errors. However, differences are weak: in the Spatio-temporal condition, the proportion of temporally adjacent errors exceeded expected by 4%, whereas the proportion of adjacent errors in the Temporal condition exceeded expected by 0.2%.

In order to determine the significance of this difference, the data were normalized to account for the variance in probability distribution, and 'don't know' responses were removed.

Condition	Adjacent		Intermediate		Distant		Total
	Obs	Exp	Obs	Exp	Obs	Exp	-
Spatio-temporal observed	518	479.5	288	319.7	145	159.83	959
Spatio-temporal proportion	54%	50%	30%	33.3%	15.1%	16.7%	-
Temporal observed	484	481.5	320	321	144	160.5	963
Temporal proportion	50.2%	50%	33.2%	33.3%	15%	16.7%	-

Table 5.4.8 Observed and expected temporal errors in the Spatio-temporal and Temporal condition.

The two-tailed t -tests on the resultant data showed no significant differences between means for temporal errors in any position. Hence H_{00i} is accepted: there is no spatial effect on temporal errors.

H_{0ii} states that there will be more temporally adjacent errors in the Spatio-temporal than the Temporal condition. The data were normalized, and ‘don’t know’ responses removed in order to perform *t-tests*. The one-tailed *t-tests* on normalized data showed no significant differences. Hence H_{0ii} is accepted.

If we compare this result with results from the Spatio-temporal condition in which there was a significant difference between spatially and temporal adjacent and distant errors, we can conclude that spatial properties *per se* have no effect on temporal object feature binding, but temporal properties enhance spatial object feature binding.

We now consider responses that resulted in illusory conjunctions (ICs).

5.5 ILLUSORY CONJUNCTIONS

An illusory conjunction (IC) is the phenomenon in which features from one object are transferred to another object to create the illusion of a composite object which is not present (see Chapter 4). The following subsections consider the proportion of ICs within each condition. For each response there were 16 possible combined²² responses: AC, AD, AI, AA, IC, ID, II, IA, DC, DD, DI, DA, CA, CI, CD, CC (where A = adjacent, C = correct, I = intermediate and D = distant; the combination represents responses for both features: size and orientation). Recall that the responses are recorded from responses to presentations which would achieve a 50% error rate. From the 16 possible combinations of errors, one is correct (CC). There is therefore a 15/16 chance of recording at least 1 error for each response. For the purpose of this study, an illusory conjunction (IC) must contain errors for both features, thus the 6 responses which contain one erroneous response and one correct response for each feature (AC, IC, DC, CA, CI, CD) are termed single errors. The remaining 9 combined errors (AD, AI, AA, ID, II, IA, DD, DI, DA), have a 9/15 (60%) chance being recorded. The expected total number of ICs is calculated from the overall total of combined errors (AD, AI, AA, ID, II, IA, DD, DI, DA).

²² A combined response comprises 2 single responses: 1 for each feature (size and orientation). Thus a combined error could comprise 1 right, 1 wrong response, or 2 wrong responses. A double error need not be an IC.

We discount correct responses although the erroneous response in a combined error comprising one correct, one erroneous response counts towards the total error. For ICs we are interested in responses where both features are wrongly recalled (these are termed double errors). Of combined errors, 3 are ICs (AA, II and DD). Hence the overall probability of recording a IC = 3/9 (33.3%), and of all ICs, the probability of recording a spatial IC = 33.3% in any position, for temporally adjacent ICs = 50%, temporally intermediate ICs = 33.3%, and temporally distant ICs = 16.7%.

5.5.1 Overall ICs

Overall, there were 3838 errors. Of these 1033 were combined errors, 877 were double errors and 613 were ICs. Thus 1590 responses contained 2 erroneous reports. Of this we would expect 33.3% to be ICs. We observed 38.6% ICs overall. In the Spatial condition 359 double errors (181 double + 178 ICs) were recorded. We would expect 120 ICs, but observed 178 (49.6%) ICs. For spatial ICs in the Spatio-temporal condition, 400 double errors (227 double + 173 ICs) were recorded. We would expect 133 ICs, but observed 173 (43.2%) ICs. For temporal ICs in the Spatio-temporal condition, 400 double errors (255 double + 145 ICs) were recorded. We would expect 133 ICs, but observed 145 (36.3%) were ICs. For temporal ICs in the Temporal condition, 331 double errors (214 double + 117 ICs) were recorded. We would expect 110 ICs, but observed 117 (35.3%) were ICs. (Table 5.4.1a).

Condition	Spatial	Spatio-temporal space	Spatio-temporal time	Temporal	Total
Double correct	351	329	329	304	1313
Combined (1 correct, 1 wrong)	250	231	231	321	1033
Double wrong	181	227	255	214	877
Total IC	178	173	145	117	613

Table 5.5.1a Responses: correct, combined, double wrong and IC. Please note that each single response comprises 2 features.

We note that in each condition, the observed count of adjacent ICs exceeded expected. The greatest difference was in the Spatial condition, and the least in the Temporal condition. We now consider the distribution of ICs which had an equal chance (33.3%) of being recorded in each position in the spatial conditions, and a 50% chance of being adjacent, 33.3% chance of being intermediate and a 16.7% chance of being distant in the temporal conditions.

	Spatial	Spatio-temporal space	Spatio-temporal time	Temporal
Adjacent	89 (50%)	96 (55.5%)	92 (63.5%)	61 (52.1%)
Intermediate	42 (23.6%)	49 (28.3%)	34 (23.4%)	45 (38.5%)
Distant	47 (26.4%)	28 (16.2%)	19 (13.1%)	11 (9.4%)
Total	178	173	145	117

Table 5.5.1b Position count and percentage of ICs in each condition

Observed adjacent ICs exceeded expected in each condition (Table 5.4.1b), but mostly for spatially adjacent errors in the Spatio-temporal condition where they exceeded chance by 30.2%. We can calculate the expected errors in each position by dividing the total number of ICs by 3. Thus for the Spatial position we expect 66 ICs in each position; for spatial ICs in the Spatio-temporal condition we expect 61 ICs; for temporal ICs in the Spatio-temporal condition we expect 48 ICs; and for ICs in the Temporal condition we expect 39 IC (Figure 5.5.1).

It is apparent that both spatial and temporal properties have an effect on erroneous binding which results in ICs. When both are present the effect is greater. When either alone is present, adjacent ICs were more than expected, but far less than when both were present (in the Spatio-temporal condition (Figure 5.5.1). Accordingly, intermediate and distant ICs were less than expected and are not included in Figure 5.5.1.

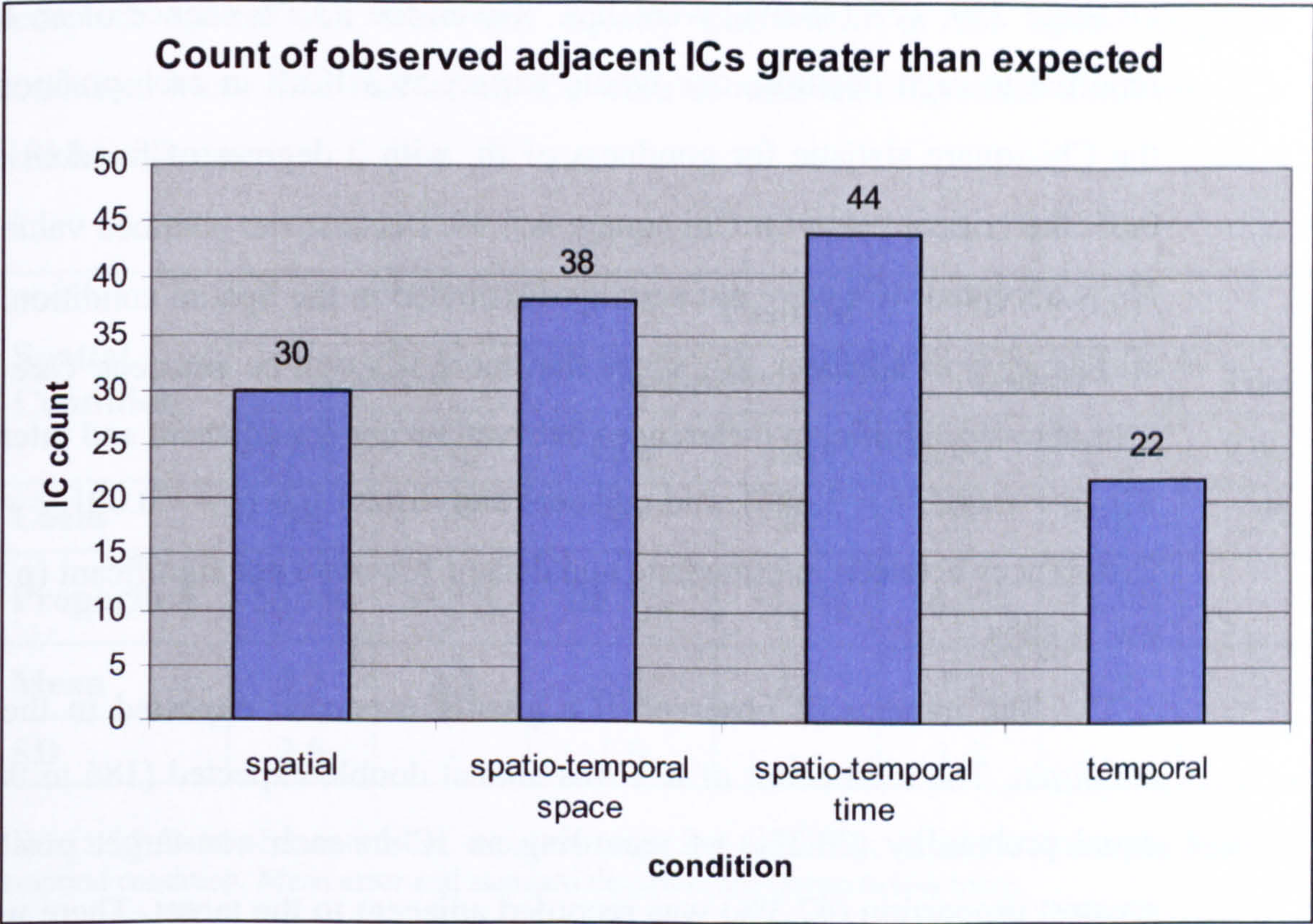


Figure 5.5.1 Count of observed adjacent ICs greater than chance in all conditions

We now describe the results for ICs in each condition individually.

5.5.2 *Illusory conjunctions in the Spatial condition*

In the Spatial condition, there were 957 errors in total for each feature. Of these, 250 were combined responses with 1 error and 1 correct, 359 were combined errors in which both features were incorrect (double errors).

Spatial Condition	Position						
	Adjacent		Intermediate		Distant		Total
	Obs	Exp	Obs	Exp	Obs	Exp	Obs
Count	89	59	42	59	47	59	178
Proportion	50%	33.3%	23.6%	33.3%	26.4%	33.3%	-
Mean	5.9	2.1	2.8	2.1	3.1	2.1	-
SD	2.6	-	1.9	-	1.3	-	-

Table 5.5.2 Total observed (obs), expected (exp) and proportion of ICs in the Spatial condition. Mean error and standard deviation are shown below totals.

Of these 359, 178 (49.6%) were ICs. Given the 33.3% chance of an IC being reported in each position, one would expect 59.3 ICs²³ in each position. Using the Chi-square statistic for goodness of fit, with 2 degrees of freedom and $\alpha = 0.05$, the critical value for Chi-square is 5.99. Because the obtained value is 22.6, H_{7i} is accepted: ICs were not equally distributed in the Spatial condition. Half of all ICs were of adjacent. H_{7ii} states that more ICs will be adjacent. One-tailed *t*-tests showed significant differences between means for adjacent and intermediate ICs ($p = 0.001$, $t = 3.686$), and adjacent and distant ICs ($p = <0.001$, $t = 3.934$). Differences between intermediate and distant ICs were not significant ($p = 0.194$, $t = -0.892$).

The number of observed ICs greatly exceeded expected in the Spatial condition. The total count of ICs was almost double expected (186 to 96). With equal probability (33.3%) of recording an IC in each non-target position, the greatest proportion (47.3%) was recorded adjacent to the target. There were thus fewer than expected ICs in the other positions (23.7% intermediate and 29.8% distant). These proportions are similar to feature errors in the Spatial condition (44.9%, 24.1%, 29.8%). One-tailed *t*-tests showed significant differences between means for adjacent and distant ICs ($p = 0.002$, $t = 3.523$), intermediate and distant ICs ($p = 0.05$, $t = -1.784$), and adjacent and intermediate ICs ($p = 0.002$, $t = 3.460$) errors. Again, the findings provide support the Spatial Theory which suggests that if spatial properties are important a greater number of ICs would be adjacent as features of one object would migrate to a spatially adjacent object rather than to a spatially distant object.

5.5.3 *Spatial illusory conjunctions in the Spatio-temporal condition*

There were 959 spatial errors in total for both features in the Spatio-temporal condition. Of these, 231 were combined responses with 1 error and 1 correct, 400 were combined errors in which both features were incorrect (double errors). Of these 400, 173 (43%) were ICs. Given the 33.3% chance of an IC being reported on would expect 133.3 ICs to be recorded, and a 33% chance of a spatial IC being reported in each position, one would expect 57.7 ICs in each position. Using the Chi-square statistic the obtained value is 41.8. Therefore, H_{8i} is

²³ As an IC comprises 2 errors, the proportion of ICs is rounded to the nearest whole number.

accepted: spatial ICs were not equally distributed in the Spatio-temporal condition. More than half of all ICs were of adjacent. H_{8ii} states that more ICs will be adjacent.

Spatial Condition	Position						
	Adjacent		Intermediate		Distant		Total
	Obs	Exp	Obs	Exp	Obs	Exp	Obs
Count	96	58	49	58	28	58	173
Proportion	55.5%	33.3%	28.3%	33.3%	16.2%	33.3%	-
Mean	5.9	2.3	2.8	2.3	3.1	2.3	-
SD	2.6	-	1.9	-	1.3	-	-

Table 5.5.3 Total observed (obs), expected (exp) and proportion of spatial ICs in the Spatio-temporal condition. Mean error and standard deviation are shown below totals.

One-tailed *t*-tests showed significant differences between means for adjacent and intermediate ICs ($p = <0.001$, $t = 4.150$), and adjacent and distant ICs ($p = <0.001$, $t = 5.768$). Differences between intermediate and distant ICs were not significant ($p = 0.011$, $t = 2.585$). The number of observed spatial ICs greatly exceeded expected in the Spatio-temporal condition. The total count of spatial ICs was almost double expected (170 to 96). With equal probability (33.3%) of recording an IC in each non-target position, the greatest proportion (55.9%) was recorded adjacent to the target. There were thus fewer than expected ICs in the other positions (27.6% intermediate and 16.5% distant). The pattern for these proportions is similar to single or unrelated spatial errors in the Spatio-temporal condition (52.3%, 26%, 21%). One-tailed *t*-tests showed significant differences between means for adjacent and distant ICs ($p <0.001$, $t = 5.819$), intermediate and distant ICs ($p = 0.016$, $t = 2.391$), and between adjacent and intermediate ICs ($p <0.001$, $t = 4.061$). These findings support the Spatial Theory.

5.5.4 Temporal illusory conjunctions in the Spatio-temporal condition

There were 959 temporal errors in total for both features in the Spatio-temporal condition. Of these, 231 were combined responses with 1 error and 1 correct, 400 were combined errors in which both features were incorrect (double errors). Of

these 400, 145 (36%) were ICs. Given the 50% chance of an adjacent IC being reported one would expect 72.5 adjacent ICs; the 33.3% chance of an intermediate IC being reported one would expect 48.3 intermediate ICs to be recorded; and the 16.7% chance of a distant IC being reported one would expect 24.2 distant ICs to be recorded. A comparison of the proportion of observed and expected ICs in each position reveals that there were more adjacent ICs (13.4%), but fewer intermediate (9.9%), and distant (4.6%) ICs than expected (see Table 5.5.5). Using the Chi-square statistic the obtained value is 11.9, therefore H_{9i} is accepted: temporal ICs were not equally distributed in the Spatio-temporal condition.

Spatial Condition	Position						
	Adjacent		Intermediate		Distant		Total
	Obs	Exp	Obs	Exp	Obs	Exp	Obs
Count	92	72.5	34	48.3	19	24.2	145
Proportion	63.4%	50%	23.4%	33.3%	13.1%	16.7%	-
Mean	6.1	2.1	2.3	2.3	1.3	2.3	-
SD	2.1	-	1.4	-	1.1	-	-

Table 5.5.4 Total observed (obs), expected (exp) and proportion of temporal ICs in the Spatio-temporal condition. Mean error and standard deviation are shown below totals.

One-tailed *t*-tests showed significant differences between means for adjacent and distant ICs ($p = 0.036$, $t = 1.956$) and between adjacent and intermediate ICs ($p < 0.001$, $t = 3.872$). These findings support the Temporal Theory. The greatest difference between proportions of observed and expected ICs was adjacent as predicted in. H_{9ij} . In order to statistically compare means, the data were normalised. One-tailed *t*-tests showed significant differences between means for adjacent and intermediate ICs ($p = < 0.001$, $t = 3.885$), and significance between adjacent and distant ICs ($p = 0.035$, $t = 5.865$). There was no significant difference between intermediate and distant ICs ($p = 0.362$, $t = -0.360$). These findings support the Temporal Theory which suggests that if temporal properties are important a greater number of ICs would be adjacent as features of one object would migrate to a temporally adjacent object rather than to a spatially distant object. Hence, H_{9ij} is supported.

5.5.5 Illusory conjunctions in the Temporal condition

There were 959 temporal errors in total for both features in the Temporal condition. Of these, 325 were combined responses with 1 error and 1 correct, 331 were combined errors in which both features were incorrect (double errors). Of these 331, 117 (35.3%) were ICs. Given the 50% chance of an adjacent IC being reported one would expect 58.5 adjacent ICs; the 33.3% chance of an intermediate IC being reported one would expect 39 intermediate ICs to be recorded; and the 16.7% chance of a distant IC being reported one would expect 19.5 distant ICs to be recorded. Using the Chi-square statistic the obtained value is 11.9, therefore H_{9i} is accepted: temporal ICs were not equally distributed in the Spatio-temporal condition. The critical value for Chi-square is 5.99. Because the obtained value is 11.75, H_{10i} is supported: ICs were not equally distributed in the Temporal condition. A comparison of the proportion of observed and expected errors in each position reveals that there were more adjacent (2.1%), and intermediate (5.2%), but fewer distant (7.3%) ICs than expected (see Table 5.5.5).

Spatial Condition	Position						
	Adjacent		Intermediate		Distant		Total
	Obs	Exp	Obs	Exp	Obs	Exp	Obs
Count	61	58.5	45	39	11	19.5	117
Proportion	52.1%	50%	38.5%	33.3%	9.4%	16.7%	-
Mean	4.1	2.3	3.2	2.3	0.9	2.3	-
SD	1.5	-	2.5	-	1.2	-	-

Table 5.5.5 Total observed (obs), expected (exp) and proportion of ICs in the Temporal condition. Mean error and standard deviation are shown below totals.

One-tailed *t*-tests showed no significant differences between adjacent and intermediate ICs ($p = 0.25$, $t = -0.692$), but highly significant differences between adjacent and distant ICs ($p = 5.865$, $t = <0.001$). Hence, H_{10ii} is supported.

5.5.6 Comparison of spatial and temporal illusory conjunctions in the Spatio-temporal condition

This section compares results for spatial and temporal ICs within the Spatio-temporal condition. The comparison of temporal ICs in the Spatio-temporal and temporal conditions is given in Section 5.5.7.

Because of the experimental design, comparison of ICs within the Spatio-temporal condition can be made by considering proportions and by normalising both sets of data. Observed adjacent ICs exceeded expected both spatially and temporally. The proportion of spatially adjacent ICs was 22.2% greater than expected, whereas the proportion of temporally adjacent ICs exceeded expected by 13.4%. This suggests that temporal properties have a stronger migrating effect in that when both spatial and temporal properties are available, more adjacent ICs are spatial, then temporal (Table 5.4.5). H_{11i} which states that there will be differences between the proportions of spatial and temporal ICs in the Spatio-temporal condition is supported.

Property	Adjacent		Intermediate		Distant	
	Obs	Exp	Obs	Exp	Obs	Exp
Spatial proportion	55.5%	33.3%	28.3%	33.3%	16.2%	33.3%
Temporal proportion	63.4%	50%	23.4%	33.3%	13.1%	16.7%

Table 5.5.6 Observed and expected spatial and temporal proportions of ICs in the Spatio-temporal condition.

Spatially intermediate ICs were 5% less than expected, whereas temporally intermediate ICs were 9.9% less than expected. Spatially distant ICs were 17.1% less than expected, whereas temporally distant ICs were 3.6% less than expected. H_{11ii} which states that there will be more spatial than temporal ICs in the Spatio-temporal condition is tested by means of one-tailed *t-tests* on the normalised data.

The tests showed significant differences between means for spatial and temporal ICs in each position: adjacent ($p = 0.001$, $t = 3.795$), intermediate ($p = 0.002$, $t = -3.423$), and distant ($p = 0.004$, $t = -3.147$). H_{11iii} is accepted and the Spatial Theory supported.

5.5.7 Comparison of temporal illusory conjunctions in the Spatio-temporal condition and Temporal condition

In order to determine the spatial effect, results for ICs in the Temporal condition are compared with temporal ICs in the Spatio-temporal condition. In order to make this comparison, data are normalised. The proportions of adjacent and intermediate ICs in the Temporal condition slightly exceeded chance, but the proportion of temporally adjacent ICs in the Spatio-temporal condition significantly exceeded chance (13.4%), and by 2.1% in the Temporal condition (Table 5.4.6).

Condition	Adjacent		Intermediate		Distant	
	Obs	Exp	Obs	Exp	Obs	Exp
Spatio-temporal proportion	63.4%	50%	23.4%	33.3%	13.1%	16.7%
Temporal proportion	52.1%	50%	38.5%	33.3%	9.4%	16.7%

Table 5.5.7 Observed and expected temporal ICs in the Spatio-temporal and Temporal condition.

This suggests that when both spatial and temporal properties are available, features are more likely to migrate to an adjacent object resulting in a greater number of ICs. When only temporal properties exist, the likelihood of making a distant IC is low (Table 5.4.5). H_{12i} which states that there will be differences between the proportions of temporal ICs in the Spatio-temporal and Temporal conditions, is supported. H_{12ii} which states that there will be more temporally adjacent ICs in the Spatio-temporal condition than in the Temporal condition, is tested by means of one-tailed t -tests on the normalised data. The tests showed significant differences between means for adjacent ICs in both conditions ($p < 0.001$, $t = 4.372$). Thus H_{12iii} is accepted and the Spatial Theory supported.

5.6 DISCUSSION OF RESULTS

This section discusses results from the behavioural experiment and is organised as follows: firstly we summarise the hypotheses in Table 5.5.

Hypothesis	Condition	Statement	Theory	Supported
<i>H1i</i>	Spatial	The proportions of observed errors will differ from expected	Spatial	Yes
<i>H1ii</i>	Spatial	There will be more adjacent than distant errors	Spatial	Yes
<i>H2i</i>	Spatio-temporal (space)	The proportions of observed spatial errors will differ from expected	Spatial	Yes
<i>H2ii</i>	Spatio-temporal (space)	There will be more spatially adjacent than distant errors	Spatial	Yes
<i>H3i</i>	Spatio-temporal (time)	The proportions of observed temporal errors will differ from expected	Temporal	Yes
<i>H3ii</i>	Spatio-temporal (time)	There will be more temporally adjacent than distant errors	Temporal	Yes
<i>H4i</i>	Temporal	The proportions of observed temporal errors will differ from expected	Temporal	No
<i>H4ii</i>	Temporal	There will be more temporally adjacent than distant errors	Temporal	No
<i>H5i</i>	Spatio-temporal (space and time)	The proportions of observed spatial and temporal errors will differ	Spatial	Yes
<i>H5ii</i>	Spatio-temporal (space and time)	There will be more spatially adjacent than temporally adjacent errors	Spatial	Yes
<i>H6i</i>	Spatio-temporal and Temporal	The proportions of observed errors in each condition will differ	Temporal	No
<i>H6ii</i>	Spatio-temporal and Temporal	There will be more adjacent errors in the Spatio-temporal than in the Temporal condition	Temporal	No
<i>H7i</i>	Spatial	The proportions of observed ICs will differ from expected	Spatial	Yes
<i>H7ii</i>	Spatial	There will be more adjacent than distant ICs	Spatial	Yes
<i>H8i</i>	Spatio-temporal (space)	The proportions of observed spatial ICs will differ from expected	Spatial	Yes
<i>H8ii</i>	Spatio-temporal (space)	There will be more spatially adjacent than distant ICs	Spatial	Yes

<i>H9i</i>	Spatio-temporal (time)	The proportions of observed temporal ICs will differ from expected	Temporal	Yes
<i>H9ii</i>	Spatio-temporal (time)	There will be more temporally adjacent than distant ICs	Temporal	Yes
<i>H10i</i>	Temporal	The proportions of observed temporal ICs will differ from expected	Temporal	No
<i>H10ii</i>	Temporal	There will be more temporally adjacent than distant ICs	Temporal	No
<i>H11i</i>	Spatio-temporal	The proportions of observed spatial and temporal ICs will differ	Spatial	Yes
<i>H11ii</i>	Spatio-temporal	There will be more spatially adjacent than temporally adjacent ICs	Spatial	Yes
<i>H12i</i>	Spatio-temporal and Temporal	The proportions of observed ICs in each condition will differ	Temporal	Yes
<i>H12ii</i>	Spatio-temporal and Temporal	There will be more adjacent in the Spatio- temporal than the Temporal condition	Temporal	Yes

Table 5.6 results of hypotheses testing for each condition.

All hypotheses, except for those concerning temporal-only properties, were upheld. These findings suggest that spatial properties are primary as features from an object adjacent to the target migrated more frequently than those from objects more distant from the target. In all conditions, in which there was a spatial component, proportions of observed adjacent errors and ICs significantly exceeded expected proportions. Furthermore, in all conditions, in which there was a spatial component, observers reported statistically significantly more adjacent errors and ICs than either intermediate or distant errors or ICs. When comparing spatial and temporal errors within the same (Spatio-temporal) condition, findings showed significantly greater differences between observed and expected spatial errors and ICs than between observed and expected temporal errors and ICs, and significantly more spatially than temporally adjacent errors and ICs. In addition, when comparing temporal errors and ICs in different (Spatio-temporal and Temporal) conditions, findings showed significantly greater differences between observed and expected errors and ICs in

the Spatio-temporal condition than the Temporal condition, and significantly more adjacent errors and ICs in the Spatio-temporal condition than the Temporal condition. The findings within and between conditions show the primacy of spatial properties.

5.6.1 Presentation rates

Details of presentation rate for each condition, for each participant, are given in Table 5.3.1a. Mean presentation rates (7.2 f/sec = 138.88ms in the Spatial condition, 7.3 f/sec = 136.25ms in the Spatio-temporal condition, and 5.8 f/sec = 172.41ms in the Temporal condition) resulted in mean presentation duration of 416.6ms in the Spatial condition, 1226.25ms in the Spatio-temporal condition, and 1551.69ms in the temporal condition (Table 5.3.1b). The primacy of spatial properties is again indicated by the similarity of mean presentation rates in the Spatial and the Spatio-temporal conditions, and reinforced by the difference in mean rates for the Spatio-temporal and the Temporal conditions. However, the modes suggest less variance among the conditions. We conclude that the noticeable difference in the mean presentation rate for the Temporal condition was due to individual differences of participants #3 and #4 specifically (see Table 5.4.1b).

We questioned whether adjacency errors occurring with longer presentation, that is fewer frames per second, would provide evidence for binding in memory, and conversely whether more adjacency errors reported with shorter presentation (more frames per second) would define perceptual binding (Figure 5.4.1a). To investigate these propositions we separated presentation rate (f/sec.) into four bins for each condition (Figure 5.4.1a), and considered mean error for each bin (Table 5.4.1c). However, we found no difference between the count of adjacency errors in each bin.

We also questioned whether presentation rate affects the position of errors, and suggested that increased adjacent error reports when presentation rates are longer (fewer frames per second) might indicate that binding errors are memory rather than perceptual errors. In addition, we proposed that if adjacent errors are more frequently reported with shorter presentation (more frames per second), binding errors are perceptual. To investigate these propositions we separated presentation rate (f/sec.) into four bins for each condition (Figure

5.4.1b), and considered mean error for each bin (Table 5.4.1c). Again, our analysis found almost no difference between the count of adjacency errors in each bin. However, this might have been due to the short time differences between presentation rates. For example, the slowest 2 f/sec. = 500ms whereas the fastest, 9 f/sec = 111ms. Thus the difference between 6 f/sec. and 9 f/sec = 55.56ms. Further work could extend the experiment to determine whether binding is a perceptual or memory process. In doing so, differences between presentation rates would need to be more pronounced. For example, rather than ranging from 111.11ms to 500ms (= 9-2 f/sec.), the range could be from 100ms to >1000ms. Another possibility is that the stimuli offset, the time when the stimuli are not on screen, could be similarly varied and thus be different from the SOA which could be maintained as in the present study.

The following section discusses overall errors in terms of their position across and within conditions, and compares the findings for errors with ICs across and within conditions.

5.6.2 General findings

As previously stated, spatial errors had equal probability of being distributed in all 3 non-target positions, and temporal errors had the probability of being distributed 50% adjacent, 33.3% intermediate and 16.7% distant. Therefore in order to compare spatial and temporal errors in the Spatio-temporal condition and temporal errors in the Spatio-temporal and Temporal conditions, we consider the proportion rather than count of errors. Proportions of observed errors above (+) and below (-) the expected proportions are tabulated in Table 5.6.3a.

POSITION	SPATIAL errors		TEMPORAL errors	
	Spatial	Spatio-temporal	Spatio-temporal	Temporal
Adjacent	+11.6	+19	+4	+0.2
Intermediate	-9.2	-7.3	-3.3	-0.1
Distant	-3.5	-12.3	-1.6	-1.7

Table 5.6.2a Percent proportion of errors greater or less than expected in each position in each condition

In all conditions, the proportion of observed adjacent errors significantly exceeded the proportion of expected errors (not significantly for temporally adjacent errors). Accordingly, for observed intermediate and distant errors the proportion of errors was less than the proportion of expected errors (not significantly for intermediate errors in the Temporal condition, nor for temporally distant errors in either the Spatio-temporal or the Temporal conditions) (see Table 5.4.3b).

The greatest difference in the proportion of observed and expected adjacent errors was for spatial errors in the Spatio-temporal condition. This suggests that when both spatial and temporal properties are available, features migrate to the object closest in space to the target. When only spatial properties are available, as in the Spatial condition, the difference between the proportion of observed and expected adjacent errors was significant, but not as pronounced as in the Spatio-temporal condition.

Although the 'staircase' method was successful in achieving the desired 50% error rate, differences in the position of errors in each condition is apparent. For example, observed spatially adjacent errors exceeded expected adjacent errors by 34.8% in the Spatial condition, and by 55.6% in the Spatio-temporal condition, whereas observed temporally adjacent errors exceeded expected adjacent errors by 7.9% in the Spatio-temporal condition, and by 0.4% in the Temporal condition. What might be the reason for these differences?

Representations of a limited number of composite objects and their location can be maintained in prefrontal cortex. However, remembering the temporal order of the presentation imposes further demands on visual information processing and memory. We propose that composite memories in prefrontal cortex are uniquely identified by their features and their spatial location. In this way, a scene comprising a moving object or objects (or in the case of the experiment, rapid sequential presentations) is represented by synchronous activity for the object at the different spatial, but not temporal positions. In real life, object movement usually follows a predictable trajectory. However, in the present experiment, objects appeared pseudorandomly in any one of four spatial locations, thus recalling the temporal sequence of spatial position from the last would be impossible. This could also account for the

finding in both conditions with temporal properties (Spatio-temporal time and Temporal), responses were close to chance.

While it is not possible to directly compare spatial errors between conditions as the number of frames in a single display in the Spatial condition was less than in the Spatio-temporal condition. It is however possible to compare spatial and temporal errors in the Spatio-temporal condition and temporal errors in the Spatio-temporal and Temporal conditions. In the Spatio-temporal condition, responses were collected simultaneously, spatially adjacent errors exceeded expected by 55.6% and temporal errors exceeded expected by 7.9%. This difference can only be a result of the effect of spatial and temporal properties. That is, when both properties are available, features from spatially adjacent objects are highly likely to migrate, but features from temporally adjacent objects are much less likely to migrate. This finding is upheld when we compare temporal errors in the Spatio-temporal and Temporal conditions. Although fewer temporally adjacent features were less likely to migrate, they did so more when spatial properties were available (in the Spatio-temporal condition), than in the Temporal condition. In fact, errors in the Temporal condition were close to expected. This suggests observers might have been guessing because with no unique spatial properties, they were unable to map features to objects. As there was no spatial indicator to help match location in this condition, one can assume that spatial properties are helpful in object feature binding. In the other conditions, findings also support the primacy of spatial properties (absent in the Temporal condition) in that observed adjacent errors exceeded expected errors (Table 5.6.3a). These findings suggest that when spatial properties are present, features migrate to the nearest spatial object. We would expect these results given the primacy of spatial properties in object feature binding (Hopf, Scheoenfeld et al., 2005).

When objects are presented simultaneously, observers try to recall the stimuli in the positions in which they appeared; the features appear relative to each other. In the Spatio-temporal condition, observers can recall the unique spatial positions of each object, but less so their temporal position. However, in the Temporal condition, the proportion of observed errors was close to expected which suggests that, possibly due to the lack of spatial properties, observers were guessing which feature belonged to each object. This finding clearly shows that

when objects are presented with unique spatial properties, more correct responses are recorded than when objects are presented with unique temporal properties.

Given the 50% error rate, out of 1920 responses, 960 errors (480 combined errors) should be reported (for all 15 observers). A combined error could be right (or wrong) (termed combined correct, and double error respectively) on both features, or correct on one feature and wrong on the other. For example, orientation might be correctly reported, but size given wrongly (Figure 5.4.2a). When both spatial and temporal properties are available (as in the Spatio-temporal condition), fewer combined correct responses were reported than when only spatial properties are available, but more combined correct responses were reported than when only temporal properties are available. This suggests the ideal arrangement for correctly identifying the features of an object is when objects have unique individual spatial properties and are presented simultaneously. However, this finding could be confounded by the shorter duration of the presentation in the Spatial condition compared with the others, and the findings could thus be a result of reduced memory load. Nevertheless, when comparing results for temporal errors from the Spatio-temporal and Temporal conditions, more combined correct responses were recorded when both properties were available (Spatio-temporal condition) than when only temporal properties were available (Temporal condition). This suggests that having both properties available enhances object feature binding. This proposition can be further investigated by comparing results for spatial and temporal errors within the same condition (Spatio-temporal condition).

In the Spatio-temporal condition, the count for combined correct, double error, and one correct one error responses are the necessarily same for spatial and temporal errors. However, the distribution of these errors is different. For example, the proportion of spatially adjacent errors accounted for 52.3% of all errors (including don't know responses) which is 19.3% more than the expected proportion (33.3%); whereas the proportion of temporally adjacent errors accounted for 54% of all errors (including don't know responses) which is 4% more than the expected proportion (50%). Despite the expected proportions being different for spatial and temporal errors, their observed distribution was similar and is in fact, close to the predicted distribution for temporal errors given

the experimental design (50% adjacent, 33.3% intermediate, and 16.7% distant) (Tables 5.4.9 and 5.6.2b).

Spatio-temporal	Adjacent		Intermediate		Distant		DK		Total	
	Count	%	Count	%	Count	%	Count	%	Count	%
Spatial	502	52.3	249	26	201	21	7	0.7	959	100
Temporal	518	54	288	30	145	15.2	8	0.8	959	100

Table 5.6.2b Spatial and temporal error count and percentage of total error in the Spatio-temporal condition.

These findings show that there are more adjacent and intermediate errors when temporal, as well as spatial properties, are available. In order to investigate to effect of spatial properties on temporal feature binding, we compare temporal results from the Spatio-temporal condition with those from the Temporal condition (Tables 5.3.10 and 5.6.2b). The percentage of errors was close to expected in each position, but spatial properties influenced temporal responses towards an adjacency effect.

A greater number of ‘don’t know’ responses were recorded in the Temporal condition which may suggest that observers found it more difficult to respond when no spatial properties were available, and supports our proposal that observers were guessing. We now investigate how these effects manifest in ICs.

	Adjacent		Intermediate		Distant		DK		Total	
	Count	%	Count	%	Count	%	Count	%	Count	%
Spatio-temporal	518	54	288	30	145	15.2	8	0.8	959	100
Temporal	484	50.3	320	33.3	144	15	15	1.4	963	100

Table 5.6.2c Temporal error count and percentage of total error in the Spatio-temporal and Temporal conditions.

When we compare ICs across and within conditions, we see that spatially adjacent ICs exceeded predicted by 22.2%, and temporally adjacent ICs exceeded expected by 13.5%. Interestingly, as was the case for individual errors described above, spatial ICs were closer than temporal ICs to the predicted

proportions of temporal ICs (Table 5.6.2c). These findings again suggest that when both are available, temporal properties exert a force over spatial properties and vice versa in terms of ICs. Whereas we noted above that temporal properties were a stronger influence than spatial properties in terms of individual errors.

Spatio-temporal	Adjacent IC		Intermediate IC		Distant IC		Total IC	
	Count	%	Count	%	Count	%	Count	%
Spatial	96	55.5	45.7	26	28	16.2	173	100
Temporal	92	63.5	34	23.5	19	13.1	145	100

Table 5.6.2d Spatial and temporal IC count and percentage of total ICs in the Spatio-temporal condition.

In order to investigate to effect of spatial properties on temporal ICs, we compare temporal results from the Spatio-temporal condition with those from the Temporal condition (Table 5.6.2d). These findings show that the percentage of observed temporal ICs was closer than observed spatial ICs to the expected percentage in each position, but that spatial properties influenced temporal responses significantly towards an adjacency effect. In addition, ICs appear to be more likely when both spatial and temporal properties are available.

	Adjacent IC		Intermediate IC		Distant IC		Total IC	
	Count	%	Count	%	Count	%	Count	%
Spatio-temporal	92	63.5	34	23.5	19	13.1	145	100
Temporal	61	52.1	45	38.5	11	9.4	117	100

Table 5.6.2e Temporal IC count and percentage of total IC in the Spatio-temporal and Temporal conditions.

Treisman and Schmidt (1982) claimed that a substantial proportion of conjunction errors were ICs. The present study found more adjacent ICs than expected which supports Prinzmetal et al. (1995) who claimed that ICs were due to a misperception of feature location. In addition, Cohen and Ivry (1990) found that ICs were more likely to be reported between adjacent objects, and when

multiple objects were present. We suggest that observers in this experiment might have perceived multiple items in the Spatio-temporal condition as a result of the brief presentation. In this condition, the proportions of both spatial and temporal significantly ICs exceeded chance.

As Maunsell and Gibson (1992) and Blake and Yang (1997) suggest, both spatial and temporal coherence may be involved in object feature binding. The findings discussed above could be due to the existence of distinct mechanisms responsible for binding spatially separated features (mediated by attention) and spatially superimposed features (through localised RFs selective to multiple features). The latter might be more prone to interference than the former as suggested by Holcombe and Cavanagh (1999). Finally, despite evidence for the primacy of spatial information, binding is possible with superimposed features (Frackowiak, 1997; Holcombe and Cavanagh, 1999); without spatial information (Duncan, 1984), and without attention (Houck and Hoffman, 1986). Displays in this study were likely to rapid the deployment of attention prior to recall, but Desimone and Duncan (1995) proposed that attention is an emergent property of recall mechanisms to help object feature binding. Prior to modelling these claims and proposals in the following chapter, we outline the threats to validity inherent in the present work.

5.7 THREATS TO VALIDITY AND LIMITATIONS

Validity is important in determining whether or not the findings can be generalized. Threats to validity can be categorised as internal and external. External validity refers to how applicable the findings are empirically. These threats are discussed in subsection 5.7.1. Internal validity refers to the relationship among variables: whether or not a cause and effect relationship can be established. This type of validity is discussed subsection 5.7.2. In order that results have value, is important to maintain both internal and external validity.

5.7.1 External validity

External validity can be compromised by several factors, for example, improper sampling, the research setting, experimental design and the researcher's mere presence.

Selection: The validity of the sample used is questionable when participants are not chosen at random. In this study, participants were recruited through an on-line request on the Bournemouth University intranet. This is available to all members of staff. However, how many actually access the site is unknown. Response was slow and therefore when 15 participants had volunteered, recruitment ceased. One participant self-reported colour blindness and thus was excluded from the study. Following further requests, another participant came forward and was included in the study. Furthermore, a greater number of participants might increase generalizability of results, however given the slow recruitment process, time pressures, and the considerable number of presentations that each participant would complete, we considered fifteen an adequate number. The sample is not a random sample, but an opportunity sample. However, because it can be generalized to an adult working population in the UK as there is no reason to suspect those who took part in this study differ markedly from the adult population of the UK as a whole. Therefore, this study has external validity in terms of selection procedures.

Research setting: The environment in which the studies took place was identical for each observer. However, the time of day varied as appointments had to be set to suit the observers, not the reverse. The observers' circadian rhythms could have affected their responses. However, no observer reported tiredness either before or after the experiment.

Experimenter bias: Experimenter bias is a threat when the experimenter evaluates the subject, but is unable to do so objectively and either overestimates or underestimates the participant in some way. This was not the case in this experiment as evaluations were made on responses to structured design.

Demand Characteristics: When participants become aware of anticipated results they can exhibit performance that they believe is expected of them. Making sure that anticipated outcomes remain undisclosed reduces the possibility of this threat. In the present study, participants were told that they would not be judged on their results and that the time they took to respond was not being monitored. They could not have anticipated the aims of the experiment beforehand. However, they were fully debriefed afterwards and any questions raised were answered.

Placebo Effects: Research (e.g. Zajonc, 1965) has found that the mere presence being watched causes a change in performance. Therefore, how can one assess whether or not these changes will also occur when no one is watching? Results in this study were recorded according to a structured protocol and other behaviours were not recorded. However, some observers enquired about the purpose of the study and some tried to justify their perceived good or poor performance. They were reassured that other behavioural factors were not being observed.

Order Effects: Order effects refer to the order in which treatment is administered and can be a major threat to external validity if multiple treatments are used. In order to eliminate this threat, conditions were randomized in this study.

Treatment Interaction Effects: Potential threats to external validity include the interaction between treatment and any of the following: selection, history, and testing. Because this repeated measures study comprised only 3 conditions which were randomized for each participant at one sitting, interaction effects cannot be considered a threat.

5.7.2 *Construct validity*

A construct is an attribute under study which is assumed to be reflected in test performance. In the present study, the construct is the effect on performance of space and time on object feature binding. The researcher believes the participants to be representative of the adult population in general. The test was based on spatial and temporal theories of binding and hypotheses were stated in support of either theory. Interpretation of the data was designed to support or refute the

hypotheses by means of statistical tests. Discussion of the results draws on parallels or differences with previous related work described in Chapter 4. The researcher believes the interpretation to be adequate and therefore valid.

External validity depends on attention to threats to internal validity.

5.7.3 *Internal validity*

Campbell and Stanley (1963) describe threats to internal validity in the following terms:

History: history refers to the measurement of behaviour at different points in time which could result in differences in the impact of the independent variable or other effects occurring as a result of change over which the experimenter has no control. The greater the period of time that elapses between measurements, the greater the risk of a history effect. In this study the time of day at which the experiment was conducted varied among participants as appointments had to be set to suit them, not the researcher. Thus the observers' responses could have been influenced by their circadian rhythms. In an attempt to eliminate this threat, in all cases, the experiment proper in which the conditions were randomised immediately followed the trials. Furthermore, environmental events were controlled to eliminate any impact they might have on the study.

Maturation: Maturation may produce changes across time and is evidenced when participants become fatigued. This can produce behavioural changes unrelated to the observer's experience or the impact of an independent variable. Maturation is not a threat to short studies such as the present one and no observer reported tiredness either before or after the experiment. We therefore assume maturation was not a threat to validity in this study.

Selection: Selection refers to the manner in which subjects are selected to participate in a study and the manner in which they are assigned to groups. These differences can be addressed through subject matching or randomization. In this repeated measures design, participants volunteered from a normally distributed population in that all members are employed at the same company and are aged between 21 and 65.

Testing: People tend to improve performance on a given activity the more they are exposed to it. Testing effects can be determined by the degree of change as a result practice or learning from exposure to repeated testing. In this study, participants would have been unable to learn or practice because both the conditions and the stimuli sets were pseudo-randomized.

Statistical regression: Where repeated measures are used, and when participants are selected because they are extreme on the classification variable of interest statistical regression to the mean may become apparent. Repeated measures were used in this study, but the sample characteristics were not assessed prior to inclusion.

Instrumentation: Threats due to instrumentation can occur when long periods of time lapse within a study and testing instruments become invalid because of cultural change. Furthermore, if the measurement device or the experimenters change during the course of the study, changes in scores may be related to those changes rather than to the independent variable. In the present study, neither the instrumentation nor the experimenter changed.

Mortality or attrition: This threat is a threat to a longitudinal study since the sample remaining at the end of the study is unlikely to be comparable to the initial sample. However, it was not a threat to the present study.

Experimenter bias: Researchers may be biased toward their desired results. This bias can affect observations and result in research errors that skew the study in particular direction. Experimenter bias can be removed by using an experimenter who is unaware of the anticipated results. The present experimenter does not consider she was biased toward any particular experimental results.

5.7.4 General limitations

The design is complex and attempts to measure a number of factors simultaneously. In retrospect, a series of simpler experiments might have provided data that would be more transparent. Nevertheless, this design is parsimonious in that it measures a number of factors in one experiment.

The present design has equal probability of spatial, but not temporal, errors being reported. If the study were to be redesigned, a consideration would be to vary the time distances between presentations in the Temporal and Spatio-temporal conditions so that the temporal properties more closely matched the spatial properties. If this were done, comparisons could be made directly between spatial and temporal errors in the Spatio-temporal condition. In order to achieve this comparison, in the present study, data were normalised.

The Spatial condition comprised 3 frames, whereas the other conditions comprised 9 frames. Thus results from this condition could not be compared with results from the other conditions as they would be more greatly affected by memory demands. This could be solved by inserting 3 blank screens before and after the stimuli so that the entire presentation in the Spatial condition also comprises 9 frames.

The stimuli were designed following a series of pilot studies using a variety of shape of stimuli. However, in order to minimize the number of feature variables, a decision was taken to make the object rectangular, and vary the size and orientation. Initially we used pastel shades, however, in trials observers reported the stimuli were indistinguishable. To solve this issue, we used bright colours (basic printing colours): yellow, green magenta and cyan (we termed the latter two purple and turquoise for ease of understanding). In retrospect, magenta was more outstanding than the other colours. It would be interesting to investigate results using a different set of colours.

Stimuli occupied four positions (in the spatial conditions) and four dimensions of each feature. In respect of the size feature, the largest and smallest were easily distinguishable, however, the other two could be confused particularly when presented sequentially (as in the Spatio-temporal and Temporal condition), although no observers reported this to be the case.

The design attempted to test a number of hypotheses using one experiment. While this is beneficial in some aspects, such as the simultaneous collection of spatial and temporal data in one condition, it results in a large data set of results which is complex to analyse and consequently to interpret. In retrospect, a series of experiments, each testing a single hypothesis, building on the previous results might result in more manageable data sets and more easily interpretable results. However, a corollary would be that more participants and

more time would be required. Nevertheless, the author considers that the data have been thoroughly analysed and insightfully interpreted in the light of existing theories.

5.8 SUMMARY

All hypotheses, except for those concerning temporal-only properties, were upheld, suggesting that spatial properties are primary. Our analysis found almost no difference between the count of adjacency errors when the variance in SOA was taken into account.

In all conditions, in which there was a spatial component, proportions of observed adjacent errors and ICs significantly exceeded expected proportions, and observers reported statistically significantly more adjacent errors and ICs than either intermediate or distant errors or ICs. In the Spatio-temporal condition, findings showed significantly greater differences between observed and expected spatial errors and ICs than between observed and expected temporal errors and ICs, and significantly more spatially than temporally adjacent errors and ICs. In addition, the primacy of spatial properties was apparent when comparing temporal errors and ICs in different (Spatio-temporal and Temporal) conditions, in which findings showed significantly greater differences between observed and expected errors and ICs in the Spatio-temporal condition than the Temporal condition, and significantly more adjacent errors and ICs in the Spatio-temporal condition than the Temporal condition.

Findings for ICs suggest that when both are available, temporal properties increase the likelihood of spatial properties and vice versa; whereas for individual errors, temporal properties are a stronger influence than spatial properties (i.e. temporal errors increase spatial errors, but spatial properties have less affect temporal errors. We found that features from spatially adjacent objects are highly likely to migrate, whereas temporally adjacent features are much less likely to migrate. Errors in the Temporal condition were close to expected (chance) which suggests observers were guessing. We would expect these results if, as suggested, spatial location is key in object feature binding (e.g. Hopf et al., 2005).

It is understood that spatial properties have a dedicated processing stream whereas temporal properties rely on memory systems which have limited capacity and are subject to interference from several sources. For example, inferotemporal cortex, implicated in maintaining a stimulus memory in the short term until it is suppressed by another, is involved in transposing temporal to spatial properties. Thus memories for sequential information would be more difficult to accurately recall.

We propose that composite memories in prefrontal cortex are uniquely identified by their features and their spatial location. In this way, a scene comprising a moving object or objects (or in the case of the experiment, rapid sequential presentations) is represented by synchronous activity for the object at the different spatial, but not temporal positions. In real life, object movement usually follows a predictable trajectory. However, in the present experiment, objects appeared pseudorandomly in any one of four spatial locations, thus recalling the temporal sequence of spatial position from the last would be impossible. This could also account for the finding in both conditions with temporal properties (Spatio-temporal time and Temporal), responses were close to chance.

Our study supports findings of Treisman and Schmidt (1982), Prinzmetal et al. (1995) and Cohen and Ivry (1990). We propose as did Maunsell and Gibson (1992) and Blake and Yang (1997) that both spatial and temporal coherence may be involved in object feature binding, perhaps driven by distinct mechanisms as suggested by Holcombe and Cavanaugh (1999).

OBJECT FEATURE BINDING IN VISUAL WORKING MEMORY: COGNITIVE MODELLING

INTRODUCTION

Neuroscience aims to understand the nervous system, particularly the brain. Sub-branches include cognitive and computational neuroscience. The former attempts to understand the relation between mental processes and brain processes; the latter aims to develop and test hypotheses about brain functions through the development and evaluation of mathematical models which are measured against experimental data from neurobiology (e.g. synaptic efficiencies), neurophysiology (e.g. single neuron behaviour), and psychology (behavioural effects via psychophysical experiments). Experimental measurements try to explain *what* the brain does; cognitive neuroscience uses these measures to build theories, while computational neuroscience tries to explain *how* and *where* brain processes are carried out. Models are used to understand complex behaviour, thus in neuroscience models are designed to contribute to a further understanding of the complexity of human information processing (Trappenberg, 2002).

Despite considerable progress in neural modelling of binding and memory, there remain several challenges. One such challenge is understanding which mechanisms could facilitate dynamic feature binding (Wersing and Ritter, 1999); another is how these mechanisms operate in short-term memory. We attempt to further understand these phenomena through a cognitive model developed upon the experimental results of the previous chapter which is designed to inform a proposed computational model for later implementation.

6.1 RATIONALE FOR THE COGNITIVE MODEL

No satisfactory, nor complete solution to object feature binding exists to date. In addition, theories of binding in memory are conflicting and incomplete. The novel approach adopted in this thesis integrates the three leading approaches and suggests that all three theories are involved in object feature binding in VSTM. The three most influential theories of object feature binding as described in

Chapter 4, are i) hierarchical feature analysis described in theories of coincidence (Hubel and Wiesel, 1962) and convergence (Barlow, 1972); ii) attention to spatial location (Treisman and Gelade, 1980; Reynolds and Desimone, 1999); and iii) temporal synchrony (Von der Malsburg, 1981; Singer and Gray, 1995; Gray, 1999).

Beyond explaining the experimental data in this thesis, the model should lead towards a further understanding of object feature binding. Page: 184 The cognitive model aims to describe behaviour and understand the neural correlates from data derived in the experiment described in Chapter 5. Our novel model thus combines, relates and adapts existing evidence with existing theories on the neural mechanisms involved in object feature binding and on VSTM. Furthermore, based on this cognitive model, we propose parameters for a future implementation of a computational model. This thesis therefore endeavours to extend our understanding object feature binding in VSTM.

The model is evaluated by means of its ability to predict the experimental data described in Chapter 5 in addition to known binding error phenomena evidenced in the literature (Chapter 4). In addition, it adapts and integrates in a novel way, existing cognitive and computational models. The novel model is therefore more concise, has better predictive power, and combines more phenomena than do existing models. Attempting to maintain biological plausibility has been paramount throughout the modelling process. For example, the model demonstrates a functional and anatomical division of visual information processing, dynamic interactions and connectivity. Specifications of how the model meets these requirements are given below in sub-section 6.2.

6.2 DESIGN

Evidence supports a functional and anatomical division of visual information processing from the LGN to the extrastriate cortex (Chapter 3). Many specialised brain areas collaborate in order to carry out cognitive functions (Chapters 3 and 4). Thus cognitive models need to take into account intra- and inter-connectivity of brain areas (see Figure 3.3.2). Previously, the dorsal stream had been implicated solely in processing spatial information and controlling visually guided action, while the ventral stream was implicated in processing object

information and mediating conscious visual perception (Ungerleider and Mishkin, 1982; Milner and Goodale, 1995). However, more recently, Deco and Lee (2002) claimed that dynamic interactions between these pathways leads to both object and spatial attention which guides the dynamics to concentrate on a specific spatial location or object features (Section 4.3.9). We take Deco and Lee's approach in that these two processing streams are modelled with dynamic intra- and inter-connectivity (Figure 6.2.1).

Evidence that prefrontal cortex plays an eminent role in WM (Fuster and Alexander, 1971; Goldman-Rakic, 1987; Fuster, 1988), PET scans (Haxby et al., 1995), and fMRI studies (Cohen et al., 1997). Prefrontal cortex has been found to have differentially selective areas specialised for the memory of spatial or object responses (Wilson et al., 1993; Goldman-Rakic, 1996) or both (Rao et al., 1997). A top-down bias from ventral or dorsal prefrontal cortex (depending on the input) would lead to object or spatial bias (respectively), possibly for attention. Thus, attentional bias could be implemented by STM, which must therefore be a separate system from perceptual systems as proposed by Renart et al. (1999) and Renart et al. (2001). Our cognitive model represents prefrontal cortex as a competitive attractor network with reciprocal connections to both ventral and dorsal visual processing streams.

In addition to prefrontal cortex, inferotemporal cortex has been implicated in memory processing. In their model, Renart et al. (1999) proposed the need for two networks: one for perceptual functions located in inferotemporal (IT) cortex, and one for WM functions in prefrontal (PF) cortex (Section 4.3.5). Thus VSTM likely depends on interaction between areas of prefrontal and inferotemporal cortex (Williams and Goldman-Rakic, 1995) in which firing may be maintained by associatively modified recurrent collateral connections between nearby pyramidal cells producing attractor states in autoassociative networks (Rolls and Treves, 1998) (Chapter 2). Furthermore, systems in addition to those in IT and PF cortices have been implicated in object feature binding. For example, in the dorsal stream, areas of posterior parietal (PP) cortex have been implicated in a variety of visual information processing tasks including spatial attention (see Section 4.2). Specific regions of parietal cortex are preferentially activated by a spatial attention task relative to an object identity task, and when

objects are presented simultaneously at different locations relative to sequentially at the same location (Shafritz et al., 2002).

This suggests a role for parietal cortex in early perceptual and a later working memory stages (Shadlen and Movshon, 1999) in feature binding when (Blake and Yang, 1997) location cues can be used to resolve scene ambiguity. This finding of fewer resources available when there are no spatial location cues, might account for our results in the Temporal condition and for temporal ICs in the Spatio-temporal condition.

Because of the multimodal nature of biological systems and our attempt to maintain plausibility, our novel model builds on these aforementioned propositions employs a multimodal structure. Accordingly the modules are a ventral stream module (predominantly for object feature processing), dorsal stream module (predominantly for processing the object's spatial properties) and prefrontal cortex module (for STM).

In summary, the multimodal attractor network model of object feature binding is a high level abstraction of many cortical systems comprising three main networks: one representing the ventral stream (VS) proceeding from V1 to inferior temporal (IT) visual cortex, via V2-V4; one representing the dorsal stream (DS) proceeding from V1 to the posterior parietal (PP) cortex, via V2-V3, MT/MST (V2/MT); the third representing prefrontal (PF) cortex with reciprocal connections to and from ventral prefrontal (V-PF) cortex via IT, and dorsal prefrontal (D-PF) cortex via PP. VS and DS are concerned with processing perceptual functions, specifically object and spatial representations respectively. Activity in these areas is interrupted by intervening stimuli, whereas in PF, a cortical area concerned with STM, activity is sustained during the delay period and when intervening stimuli are presented. The three modules themselves comprise submodules which are detailed below.

6.2.1 The dorsal stream module (DS)

To recap briefly on the visual and memory systems involved (for full details see Chapter 3), the magnocellular pathway from the occipital cortex projects from retinal and LGN magno cells through V1 (layers 4C α and 4B) to the thick stripes

in V2, to MT, MST and posterior parietal (PP) area. This is modelled as DS (Figures 6.2.1 and 6.2.2).

6.2.2 The ventral stream module (VS)

Form and colour information is relayed from retinal parvo cells, through LGN parvo layers to V1 (interblobs and blobs respectively), V2, V4 and IT. V4 is particularly sensitive to colour and orientation processing (Zeki, 1993). This pathway is represented in the model as VS (Figures 6.2.1 and 6.2.2).

6.2.3 The prefrontal cortex module (PF)

Responses to information remain separated in the pathways until representations activate pools of neurons in areas of prefrontal cortex (for object information V-PF, representing 46v and D-PF, representing 46d, for spatial information). This is represented in the model as PF (Figures 6.2.1 and 6.2.2). In this scenario, the coupled prefrontal cortex (PF) attractor network, triggered by feed forward activity driven by stimulus features in inferotemporal (IT) cortex and/or by stimulus spatial location in PP, would sustain activity in the event of an intervening stimulus. (see Figures 6.2.1 and 6.2.2).

6.2.4 Operation of the multimodular cognitive model

In the real world, visual information occurs over both space and time. Temporal aspects of VSTM have been less studied than spatial aspects and conflicting results have been found. For example, Brockmole et al.'s (2002) 'image-percept hypothesis' proposed that a mental image of a first array is integrated with a second on presentation. However, correct recall would depend on a long (~1300ms) SOA. In contrast, the 'limited integration hypothesis' (Jiang et al., 2005) suggested that events are represented and retained by spatio-temporal continuity.

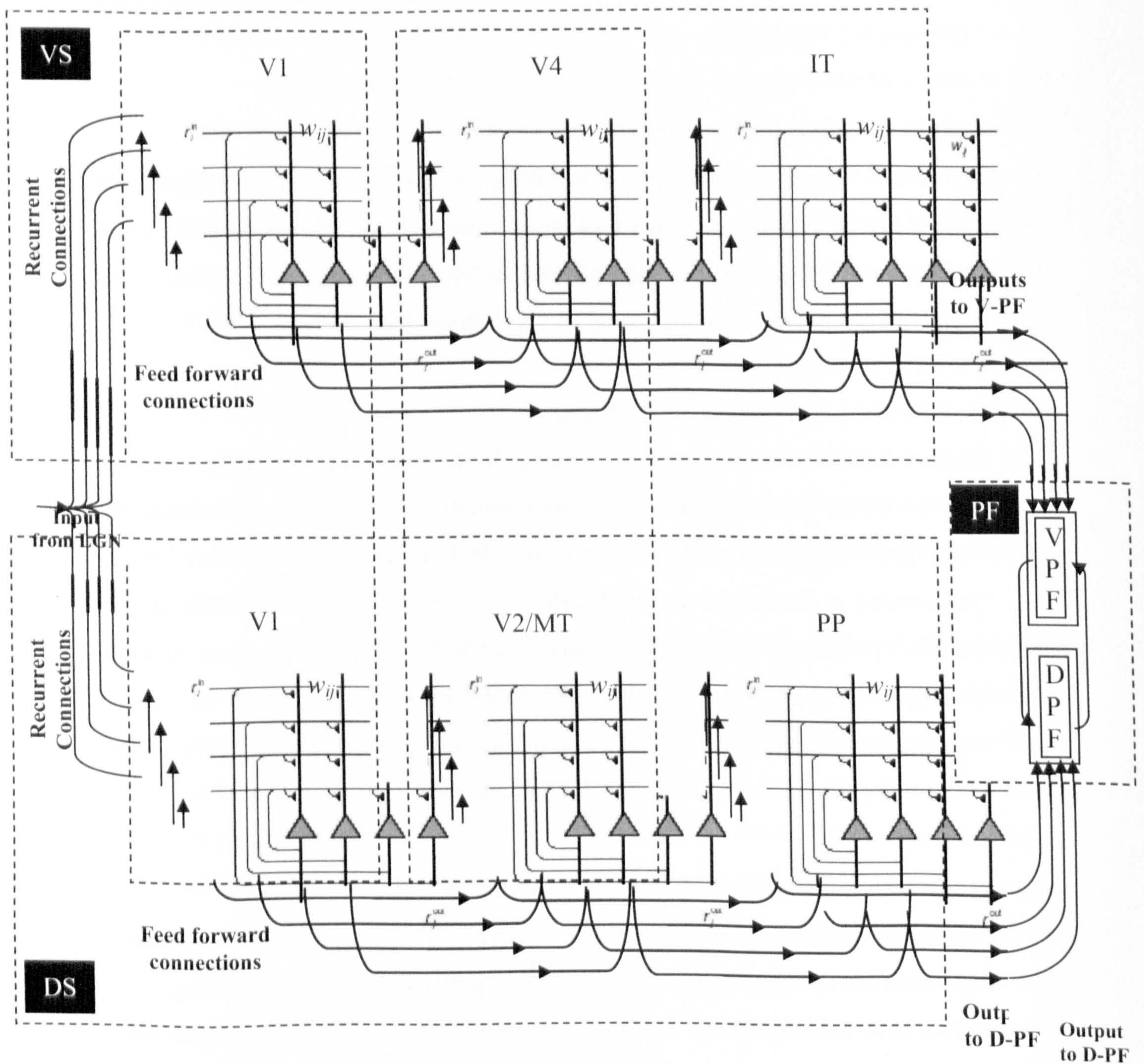


Figure 6 Computational model architecture. A schematic of the loosely coupled networks showing feed forward and recurrent connections between submodules representing the VS module (V1, V4 and IT), the DS (V1, V2-V3, MT/MST (V2/MT) and PP), and the PF module (V-PF and D-PF).

Because attractor neural networks are able to store and retrieve memories as patterns of neural activity via synaptic modifications (Hebb, 1949), and can form memory states (Hopfield, 1982) manifested as self-sustaining selective neural activity (e.g. Miyashita, 1988; Fuster, 1995), the neural assembly is able to sustain an active representation of the stimulus in its absence, and recall a complete pattern on presentation of a similar or partial one (Kesner and Rolls, 2001). Thus, on presentation of the post-stimulus cue, neurons in IT would increase their activation because they are being stimulated both by bottom-up connectivity from V1 from the external visual cue, as well as top-down connectivity from PF where the memory of the integrated object features has been sustained. However, when the subsequent object (or mask) is presented, activity in IT for the previous object (or mask) would be suppressed, but activity for the previous object (or mask) in PF would continue alongside activity for the subsequent object (or mask).

6.2.5 *Predictions*

The proposed model suggests that object feature binding takes place as a result of external or internal activation which drives competition in prefrontal cortex (PF) where memories of the features and their spatial positions are retained as attractors. In this way the memory of the correctly bound features that comprise the target is more highly activated than other sustained memories active in prefrontal cortex and becomes the ‘winner’. In doing so, top-down bias is relayed to the appropriate visual processing stream (depending on whether the input is an object feature (processed in VS), or a spatial reference (processed in DS) so that the composite object can be recalled.

This would be the case in most natural situations. However, we tuned the experiment so that we achieved a ~50% error rate on each feature in order to have sufficient data to analyse. Within these errors, observed single feature errors and ICs were not distributed according to chance in each condition. For example, in the Spatial condition and for spatial errors in the Spatio-temporal condition, chance single errors would be distributed 33.3% in each non-target position. In the Temporal condition and for temporal errors in the Spatio-temporal condition, chance single errors would be distributed 50% in the adjacent position, 33.3% in the intermediate position, and 16.7% in the distant position (see Chapter 5). ICs

were taken to be two errors relating to the same non-target object and were calculated to be expected equally (33.3%) in each on-target position in each condition (see Sub-section 5.5). To recap, observed and expected results are given in Table 6.2.5 below.

Adjacent errors	SPATIAL				TEMPORAL			
	Spatial		Spatio-temporal		Spatio-temporal		Temporal	
	obs	exp	obs	exp	obs	exp	obs	exp
Single	44.9%	33.3%	52.3%	33.3%	54%	50%	50.2%	50%
IC	50%	33.3%	55.5%	33.3%	63.4%	50%	52.1%	50%

Table 6.2.5 Observed and expected adjacent single errors and ICs from behavioural data (Chapter 5). (NB obs=observed; exp=expected)

In all conditions, proportions of adjacent single errors and ICs exceeded chance. This proportional difference was significant for all spatial errors and ICs and for temporal ICs in the Spatio-temporal condition, but not for single temporal errors or ICs in the Temporal condition. The proportional difference was greater for ICs than for single feature errors in all cases.

Because reports for temporal single errors and ICs were close to the proportions expected by chance in the Temporal condition, we speculate that observers were guessing. That is, they were unable to recall individual target features when objects were presented sequentially with no unique spatial property. However, although the same proposition would hold for single temporal errors in the Spatio-temporal condition, there was a significant difference between proportions of observed and expected ICs in this condition.

This can only be accounted for by the effect of spatial properties as data for both temporal and spatial errors were recorded simultaneously in the Spatio-temporal condition. Our findings support findings of Prinzmetal et al. (1995) who reported ICs were a result of an incorrect perception of feature location. In addition, Nissen (1985) found that when the cue was colour, correct recall of shape depended on correct recall of location. In our experiment, the cue was colour and observers were asked to recall size and orientation. These features could be interpreted as shape. Hence in the Temporal condition, with no unique

spatial location available, we conclude that observers guessed which features belonged to the target. For temporal ICs in the Spatio-temporal condition, the simultaneous availability of spatial properties resulted in the migration of both features to a non-target object.

In comparison to spatial ICs in the Spatio-temporal condition, the proportional difference for temporal ICs was less. Nevertheless, the difference was significant. We have speculated in Chapter 5 that when spatial properties are available, observers maintain the spatial correlates of the object and link to these the object's features. This is likely through attention to the spatial location activated through the dorsal prefrontal cortex. The finding that features migrate to an adjacent object rather than an intermediate or distant object, suggests a merging of memories for spatially adjacent objects, perhaps in inferotemporal cortex which has been found to maintain a memory during stimulus presentation, but not following intervening presentations. We speculate that temporal feature migration could take place in this cortical area. In addition, memory fields in inferotemporal cortex have been found to transpose temporal to spatial properties. Notwithstanding this proposition, the incongruous finding of a significant proportional difference of temporal ICs in the Spatio-temporal condition remains unsolved. In Section 7.4 we suggest further work to find a solution this phenomenon.

Furthermore, object and spatial features, having reached inferotemporal (for object features) and posterior parietal (for spatial information) cortex by means of feed forward and recurrent connectivity in the two visual processing streams are loosely bound through principles of hierarchical feature analysis. At this 'later' processing stage, information from the loosely bound objects continues to prefrontal cortex where memories (attractors) can be sustained even in the case of intervening stimuli presentation. However, we suggest that presentation of a subsequent intervening stimulus drives activity for that stimulus, and activity sustained for previous stimuli is suppressed and drifts from that basin of attraction to others. In recall, objects are located by their perceived spatial properties which, for sequential presentation, will have drifted into other basins of attraction. Some of which will be spurious attractors. In this case, the recalled memory will be close to the actual memory, but not exactly. This could explain why the proportion of observed temporal ICs was significantly greater

than expected in the Spatio-temporal condition and not in the Temporal condition where we surmise that observers were simply guessing.

In the behavioural experiment (Chapter 5) variables comprised three features: size, orientation and colour, plus spatial location (four in the Spatial and Spatio-temporal conditions, and one in the Temporal condition). In summary we predict that object feature binding takes place through cooperation (excitation) and competition (inhibition) in each sub-module in the multimodular system using local (Hebbian) learning rules. We further suggest that interference in inferotemporal cortex is likely responsible for temporal ICs when spatial in addition to temporal properties are available. When no spatial properties are available, observers guess which object features are bound.

6.2.6 Discussion of cognitive model in relation to behavioural data

The model predicts that where there are no unique spatial properties (as in the Temporal condition) observers are likely to use a guessing strategy to determine which features belong to which object. Support for this appears to come from results derived in the Spatio-temporal condition where both spatial and temporal responses were simultaneously collected. That is when both spatial and temporal properties are simultaneously available, the proportions of observed spatial single errors and ICs significantly exceeded expected errors and ICs; whereas the proportions of observed temporal single errors were at chance levels, but temporal ICs significantly exceeded expected ICs. What neural correlates could account for this phenomenon?

If object feature binding takes place as a result of external or internal feed forward activation driving competition which in turn produces top-down bias, spatial properties could have an advantage over temporal properties due to the dedicated spatial system operating in the dorsal pathway. One may argue however, that all information processing is essentially spatio-temporal as it happens within and across spatially separate systems over time. Nevertheless, results from the behavioural study suggest that spatial properties are primary and temporal properties secondary.

In each condition, following the presentation of the object(s), a checkerboard mask was presented at fixation for the same period of time as the SOA for the observer in that condition. Several models of backwards masking

exist (Section 3.3.4), however, our model predicts that the mask interferes with information stored as a decaying trace in early visual areas, rather than by replacing (Enns and DiLollo, 2001) or ‘catching up’ with the earlier stimulus (Breitmeyer, 1984). In this way, interference is likely due to lateral inhibitory mechanisms, as suggested by Bugmann and Taylor (1994; 2005). We predict therefore that activation in both modules VS and DS, will be disrupted, but activation in the PF module is sustained for previously presented stimuli. As STMs are stored as attractors in the PF module, the most recent memory is the most strongly activated. When the mask is presented, activation for the previous stimulus is sustained, and activity for the mask is fed forward to PF through the magnocellular pathway. This activity is sustained as an attractor in addition to the attractor for the previous stimulus. Thus there are two attractors representing two memories. In addition, when the second stimulus is presented, neurons responding to its features feed forward through VS, and neurons responding to its spatial properties feed forward through DS to PF where the integrated features and their location form another attractor and so on until the fourth mask is presented. Following the delay period corresponding to the SOA, the post-cue is presented. The external input activates neurons in VS, particularly in V4 which processes colour. This activation propagates towards IT and onwards to V-PF where the STMs for the loosely integrated features and their locations are sustained. This increased activation acts as top-down bias through both VS and DS modules.

The novel model presented here integrates the three most influential theories of object feature binding as described in Chapter 4, are i) coincidence and convergence (Hubel and Wiesel, 1962; Barlow, 1972; Rao et al., 1997); ii) attention to spatial location (Treisman and Gelade, 1980; Reynolds and Desimone, 1999); and iii) temporal synchrony (Von der Malsburg, 1981; Singer and Gray, 1995; Gray, 1999). The rationale for this is described below.

Barlow (1972) proposed that neurons at later processing stages code for features loosely grouped into whole objects through the convergence of their co-activated relevant feature detectors (Subsection 4.1.1), and stable convergent connections could be established by means of modified Hebbian learning (Hebb, 1949) through neurons which respond to low order combinations of features at each stage of visual processing (Fukushima, 1975) (Chapter 2). Furthermore, by

incorporating fixed-feature combination neurons, a hierarchical feature analysis system can implement spatial relations between features in hierarchical layers so that neurons in 'higher' layers (in our model, IT, PP and PF) respond to more complex combinations. Individually hierarchical convergence and coincidence could not be complete binding mechanisms, but they could operate together, in addition to additional mechanisms.

The spatial attention binding theory (Subsection 4.1.2), for example the Feature Integration Theory (Treisman and Gelade, 1980), suggests that distributed representations of object features are conjoined by attending a region of visual space. Without focused attention features from different objects may become erroneously combined into illusory conjunctions (ICs). Support for the primacy of spatial information in attention comes from event related potential (ERP) studies which show that location-based selection influences the feed forward information flow through the visual system. However, Hopf et al. (2005) emphasised that temporal attention also allows different sequences of location, feature and object-based selection mechanisms to meet the demands of the task in hand. We suggest this might explain the significant result for observed temporal ICs in the Spatio-temporal condition. That is, in the case of bound features, but not for single features, where there are unique spatial properties, temporal properties are translated to spatial properties recall.

Electrophysiological recordings (Chelazzi et al., 1993) support claims that attention is serial (Treisman and Gelade, 1980; Duncan and Humphreys, 1989; Wolfe, Cave et al., 1989; Corchs and Deco, 2001). Chelazzi et al.(1993) propose that after initial parallel processing, competition leads to target identification. In our model, the parallel processing takes place in the visual streams (VS and DS) and the competition takes place in PF driven by the post-cue stimulus. This also supports later work which suggests attention is an emergent property of bottom up and top down neural mechanisms (Desimone and Duncan, 1995; Corchs and Deco, 2001) and the dynamic behaviour of synchronous activity of pools of neurons corresponding to different properties of the same object. Furthermore, winning objects in one system become dominant in other systems (Duncan et al., 1997b). The model proposed in this chapter and illustrated in Figure 6.2.1 allows these diverse propositions to cooperate in the following manner.

Coincidence and convergence (as seen in hierarchical feature analysis) operate throughout the visual pathways as evidenced through the larger receptive fields of neurons at 'later' areas. Thus a representation for 'loosely connected' features of an object is processed in IT (Hommel, 1998), and the representation for its location processed in PP. Coincidence of activity is moderated by lateral inhibition (competition) between pools of neurons in VS, and results in synchronous firing throughout the hierarchy. Simultaneously, in DS, response to the location of the external input, results in activation in pools of neurons responding to that location. Lateral connections between VS and DS at IT and PP allow the location and the features to be combined. Feed forward connectivity to PF allows sustained activations in pools of neurons representing the composite object at a specific location. We propose that rapid oscillations facilitate dynamical cooperation between neuronal pools in different brain areas and enhance synchronous activity between submodules V-PF and D-PF. This in turn drives enhanced firing which acts as top-down bias to IT and PP. Thus cooperation would complement competition to allow the simultaneous activation of several competitive pools. Object and spatial attention, produced by dynamic interactions between the VS and DS modules is thus an emergent property arising from this top-down bias from PF to IT or PP respectively (Deco and Lee, 2001). This attentional bias activates neuronal pools simultaneously and creates simultaneous dynamical states to guide the dynamics to concentrate at a specific spatial location or specific object features and thus helps bind coherent objects.

Presentation times in the experiment (Chapter 5) were considerably shorter than the 500ms considered by Deco and Rolls (2002) to constitute a static stimulus. Neurons which respond to the same sensory object might fire in precise temporal synchrony, while no synchronisation should occur between cells activated by different objects (e.g. Engel et al., 1991b; Niebur et al., 1993). In this case, binding information is conveyed through neuronal firing patterns in which temporal properties may be characterised by some phase, frequency, or both (Sohal, 1995). However, synchronization alone cannot be used as a binding mechanism when multiple patterns are active, since a synchronously firing neuronal ensemble signifies a single event only (Ritz and Sejnowski, 2000). Furthermore,, synchrony can arise through the synchronous firing of individual neurons, or more plausibly through synchronous firing within a population of

neurons. Corchs and Deco (2001) proposed that synchronous oscillations are induced by, and are dependent on an attentional mechanism. We postulate that the attentional mechanism is an emergent property of top-down processing emanating in PF.

In memory retrieval, neurons that code for more than one representation become activated, and to avoid representational interference due to cross-talk (Schacter, 1989), binding may be involved in memory encoding through non-stationary synchronized activity (Raffone and van Leeuwen, 2001). Such dynamics may enable the integration of local features into global representations (van Leeuwen et al., 1997), and is more likely than static or phase-locked neural oscillators (Fujii et al., 1996) to be used for perceptual and working memory cortical coding.

The dorsal stream is dedicated to processing object features; the ventral stream is dedicated to processing spatial properties. No such channel exists for temporal properties. Evidence has shown that neural firing is stochastic and that little additional information is available from spike timing or interspike intervals. Despite synchrony being a useful binding agent, it is problematic when multiple objects are present. In agreement with Amit et al. (1994), we argue that attractor activity in IT cortex could explain the conversion of temporal into spatial correlations between the representation of sequential stimuli. Additionally we suggest that the attractor activity may ‘drift’ differentially temporally and it is this which is responsible for incorrect temporal bindings.

To test the assumptions made in the cognitive model presented in this chapter, the proposed computational model would require modes for learning, maintenance and recall. In the learning mode, V4 and IT would be trained with Hebbian learning during presentations of composite objects at one of the four spatial locations. In the maintenance mode, nodes in PF and IT would continue to fire (e.g. for object 1). However, when the mask is presented, activity would be disrupted in IT, but would continue in PF at a reduced rate. When object 2 is presented, nodes in PF and IT would fire and when the second mask is presented, activity would be disrupted in IT, but continues in PF at a reduced rate. Thus firing in PF is synchronised for both objects, but in different phases with neurons representing the most recent object firing most strongly. This process would continue until the four objects, each followed by a mask have been presented. At

this stage neurons would fire above threshold in PF for the four objects. On presentation of the colour post-cue, nodes in VS would be activated and the patterns of features which represented the object of that colour (the target) would begin to fire more strongly. This happens rapidly as the neurons would be already firing above threshold. Inhibition biases the response in both VS and DS through lateral connections. Hence the memory for that object is reactivated in PF.

The models present a novel modular systems approach to object feature binding in VSTM. Although implementation of the computational model will not be completed as part of this thesis, and is a task for further work, parameters for its implementation are suggested in the section that follows.

6.3 OUTLINE OF THE PROPOSED COMPUTATIONAL MODEL

A variety of theoretical approaches and levels of abstraction are available when modelling memory (Brunel and Nadal, 1997). The architecture of the computational model in this thesis is based on biological and physiological evidence. At the cellular level, interneurons in the cerebral cortex are inhibitory, whereas pyramidal neurons are excitatory. Interactions between interneurons and neurons facilitate orientation specificity in V1 neurons, and are likely involved in the formation of memory fields in prefrontal cortex (Silito and Murphy, 1986). Memory fields may be maintained by associatively modified recurrent collateral connections between neighbouring neurons which process attractor states in autoassociative networks (Amit, 1995; Rolls and Treves, 1998). The attractor dynamics demonstrated in low-level computational models of working memory can be scaled up to higher, systems levels of abstraction able to reproduce psychophysical data from memory experiments (Chapter 2).

Attractor neural network models form memory states during learning. When a stimulus is presented, it elicits a specific configuration of activity that is learned via Hebbian synaptic modifications. These modifications enable the network to sustain an active representation of the stimulus in its absence. Hopfield (1982) defined the first full mathematical formalisation of Hebb's (1949) ideas on the neural assembly, the learning rule, the role of the connectivity, and the neural dynamic, as an associative memory model based on formal neurons. In addition, attractor neural networks are able to store and

retrieve many different memories as patterns of neural activity which can be recalled from short term memory (STM) when the network is provided with a fragment of the original input. In order to operate as a STM, attractor networks require internal feedback (recurrent connections). However, feedback can produce serious interference between stored memories, and as feedback increases, so does interference. To decrease interference from internal feedback, biological systems demonstrate sparse connectivity (Section 2.2.6). In fact, neurophysiological experiments indicate that the fraction of neurons participating to a given 'memory state' in the observed area is very low (~ 0.01) (e.g. Miyashita, 1988), and computational models of memory have shown sparse connectivity decreases interference from feedback. To this end, we propose that the computational model is based on attractor dynamics, has sparse connectivity, and is trained by a modified Hebbian learning rule.

Attractor neural networks store and retrieve memories as patterns of neural activity via synaptic modifications (Hebb, 1949), and can form memory states (Hopfield, 1982) manifested as self-sustaining selective neural activity (e.g. Miyashita, 1988; Fuster, 1995) which enable the neural assembly to sustain an active representation of the stimulus in its absence, and recall a complete pattern on presentation of a similar or partial one (Kesner and Rolls, 2001). Thus, on presentation of the post-stimulus cue, neurons in IT would increase their activation because they are being stimulated both by bottom-up connectivity from V1 from the external visual cue, as well as top-down connectivity from PF where the memory of the integrated object features has been sustained.

The computational model would comprise the discrete independent, yet interrelated attractor network modules: VS, DS and PF as previously described in the cognitive model. The three modules would comprise submodules of independent recurrent nodes, which could maintain previously presented information over a short period of time through reverberating neural activity, connected by reciprocal recurrent connections. VS would be represented by an autoassociative recurrent attractor network; DS by a continuous attractor network; and PF by a competitive attractor network. In order for sustained, thus recurrent activity to take place, the recurrent pathway must be strong enough, and there must be a delay activity so that re-entry signal does not fall within the refractory period.

Recurrence would be modelled using a single node to represent a population (pool) of neurons with similar preferences. The recurrent connections would be trained using Hebbian learning such that simultaneous pre- and post-synaptic firing result in synaptic modification. In brief, the model should i) use leaky integrate and fire units (described in Section 2.2.6.1), because information is not carried by the precise form of the spike, but by its occurrence; ii) incorporate a balance of local recurrent inhibition (Compte et al., 2000), and long-range excitation (Shalden and Newsome, 1994); and iii) use sparse connectivity to reduce interference between stored memories which could be produced by internal feedback (Miyashita, 1988). These modules are more fully described in the following subsections.

6.3.1 The autoassociative recurrent attractor network: VS module

The VS module comprises submodules: V1, V2-V4 and IT. V1 processes form, colour and contour, and spatial information. A larger representation of the features is represented in V2-V4, and in IT, where receptive fields (RFs) are larger and neurons respond to more complex combinations of the features (this corresponds to hierarchical feature analysis). Thus the object becomes loosely bound. Spatial coordinates are not processed in this area, but in PP of the dorsal stream (DS) (Section 6.2.2) (again in accordance with hierarchical feature based analysis, see Sections 3.2.1 and 4.1.1). Thus V1 contains nodes which each represent pools of neurons with similar tuning preferences for the four dimensions of colour, size, and orientation (12 nodes in all). Mutual inhibition operates among these so that the feature dimension corresponding to the input is activated more strongly than the other three nodes for that feature. Feed forward and reciprocal feedback connections allow information from V1 to be represented in V2-V4, and information from V2-V4 to V1 allows a sharper image to emerge. From V2-V4, reciprocal connections send (receive) information from (to) IT through fully connected neurons trained by a modified Hebbian rule to include decay.

Feed forward connections send signals from IT to V-PF which in turn maintains activity representing the object even when an intervening stimulus is presented (in this case, the mask and subsequent composite objects). The activity in V-PF is maintained in the absence of the stimulus and during intervening

stimuli presentation, whereas in IT activity is disrupted by intervening stimuli. Dynamic interactions from lateral connections to DS lead to object and spatial attention which guides the dynamics to concentrate at a specific spatial location or object features (Section 4.3.9).

A proposed training mechanism for the VS network would be using patterns representing composite objects. In a recurrent network, cyclic activity enhances pattern completion until the learned output pattern is reached. In order to stop all the nodes becoming active when a specific subset is active, recurrent connections are tuned using Hebbian learning. After the external input pattern (composite object) is presented, each submodule in the network responds with a pattern more similar to a trained pattern. This response is fed back as input to the same network so that it responds with a pattern even closer to the desired pattern. Simultaneously, output from one submodule is fed forward to the next submodule in the hierarchy (V1 to V4, V4 to IT) so that in IT, a full representation of the object's features is represented. This is fed forward to PF (described in Section 6.3.3) and fed laterally to the PP submodule of the DS module (described in Section 6.3.2). In addition, lateral dynamic interactions between modules VS and DS lead to the emergence of object and spatial attention which guides the dynamics to concentrate at a specific spatial location or object features as suggested by Deco and Lee (2002) (Section 4.3.9). Cyclic activity stops once the learned output pattern is reached. This pattern is kept active in the network until the next pattern is applied. The activity of these 'loosely connected features' remains active in IT until the next pattern, the mask (in the case of the present study) is presented. This input results in firing of a different set of nodes in each of the submodules in VS, for the new stimulus, while the activity for the first stimulus is suppressed. However, the activity in V-PF for the object features of the first stimulus is sustained, and when input from IT representing the mask, is fed forward to V-PF, other nodes tuned to that stimulus also fire, but at a different phase. This process continues until all eight frames have been presented (4 objects and 4 masks) at which point there are sustained activations for each object and the mask in V-PF. The next stage is described in Section 6.3.3.

Associative learning depends on relating pre- and post-synaptic activity imposed by an unconditioned stimulus (see Section 2.2.7.1). If the dynamic from

the recurrent network dominates, activity is driven away from the desired activity pattern. To solve this, the network operates in two phases: training and retrieval. Switching between training and retrieval phases might be accomplished biologically through signals propagated along the large synapses of mossy fibres in the hippocampal complex (in MT and incorporated into the DS module). Alternatively, it might be achieved via neuromodulators which modulate learning and enable switching through facilitating synaptic plasticity and enhancing neuronal firing. In addition, neuromodulators suppress excitatory synaptic transmission which can suppress the effects of recurrent collaterals, in turn this renders the neurons more responsive to external input.

6.3.2 The continuous attractor network (CANN): DS module

The DS module would comprise submodules: V1, V2-V3, MT/MST (combined into V2/MT) and PP. As in VS, V1 performs as described in Section 6.2.1.1. A larger representation of the features is represented V2-V3, and MT (in accordance with hierarchical feature analysis). Information from this area is fed forward to PP. As before, mutual inhibition operates so that the spatial location corresponding to the spatial location of the input is activated more strongly than the other three nodes for that location. Again, feed forward and reciprocal feedback connections allow information from V1 to be represented in V2/MT and send or receive information from or to PP. From PP to D-PF, recurrent connections allow top-down biasing to spatial properties.

Because we are interested in spatial properties of visual processing, a CANN can be used to represent spatial information (see subsection 2.2.7.4). The weights are excitatory for short distances, and inhibitory for long distances between nodes which gives the model local cooperation and global competition. Local cooperation encourages ongoing firing of neighbouring nodes, and long-range competition reduces the firing of remote nodes. A collection of active nodes is termed ‘an activity packet’ (Trappenberg, 2002).

As in the recurrent autoassociator described previously, once the initial external input pattern is presented, each submodule in the network responds with a pattern more similar to a trained pattern. This response is fed back as input to the same network so that it responds with a patterns even closer to the desired pattern. Simultaneously, output from one submodule is fed forward to the next

submodule in the hierarchy (V1 to V4, V4 to PP) so that in PP, a full representation of the object location is represented. This is fed forward to D-PF (described in Section 6.3.3) and laterally to IT submodule of the VS module (described in Section 6.3.1). The activity representing the location of the object remains active in PP until the next pattern, the mask (in the case of the present study) is presented. This input results in firing of a different set of nodes in each of the submodules in DS, for the new stimulus, while the activity for the first stimulus is suppressed. However, the activity in D-PF for the object location of the first stimulus is sustained, and when input from PP representing the mask's location, is fed forward to D-PF, other nodes tuned to that location also fire, but at a different phase. This process continues until all eight frames have been presented (4 objects and 4 masks) at which point there are sustained activations for each object's and the mask's location in V-PF. The post-cue stage is described in the following subsection.

6.3.3 The competitive self-organising network: PF module

Drawing on the design of Renart and co-workers (1999; 2001), the PF module comprises two reciprocally connected attractor models (V-PF and D-PF) representing areas 46v and 46d, with dynamic excitatory-inhibition balance achieved through organising the adjacent excitatory and inhibitory leaky integrate and fire nodes, with the same statistical properties, into a single subpopulation, each characterised by the mean firing rate of its neurons. Excitatory and inhibitory connections are sparse, and stronger within, than between each submodule (V-PF and D-PF). The generation of persistent activity occurs through tuned inhibition (Rao et al., 1999), while the balance between total excitation and inhibition generates highly irregular spike trains as seen in prefrontal cortex (Shalden and Newsome, 1994). Furthermore, coincident patterns of activity in IT and PF are associated by reciprocally connected Hebbian-like synapses so that sustained firing in PF can reactivate firing for the target in IT which would have subsided in the event of new stimuli. Reciprocal long range signals between modules are correlated with the representations of the external stimuli in each module. Amit et al. (1994) showed that attractor activity in IT cortex could explain the conversion of temporal into spatial correlations

between the representation of sequential stimuli observed in Miyashita (1988), thus it is appropriate in this model.

Information about spatial and object properties is shared and maintained through lateral connections between V-PF and D-PF. This activity forms a memory field which exhibits sustained activity during the presentation of intervening stimuli. On presentation of the post-cue (which for this thesis, comprises colour at the central fixation point), for a correct response, pools of neurons tuned to that colour fire in V1 and activity propagates throughout the hierarchy. Meanwhile, sustained activity in PF for the composite objects includes activity for one object with the same colour as the post cue. Thus feed forward input from IT to V-PF heightens firing for that colour and through lateral inhibition, reduces activity for the other objects. Furthermore, lateral connections between V-PF and D-PF connect the most active representation with its location. Feed back connectivity from PF acts as top-down bias to IT and PP such that activity in those areas is heightened from the external input (the colour post-cue) as well as the top-down bias. Thus both the top-down bias and bottom-up input lead to completion.

In a competitive self-organizing network, the collateral interactions within the network are fixed, but retain the short-range excitation and long-range inhibition common to the other two modules. VF comprises, reciprocally connected submodules V-PF (Section 6.3.1) and D-PF (Section 6.3.2). In addition to their reciprocal connections to each other, these submodules are reciprocally connected to IT and PP respectively. Nodes in both V-PF and D-PF exhibit sustained activity representing the objects and their locations. On presentation of the cue which follows the fourth mask, after the fourth object, activity propagates from V1 through the hierarchy until activity from IT and PP is fed forward to PF. At this stage, activity for the four objects and the mask is being sustained through synchronous oscillations for each object. When the activity for the cue reaches V-PF it enhances the firing of the nodes representing the object in that colour (the target) and because of competition in the network, activity for other objects is suppressed. This activity is fed back downwards throughout VS and DS as top-down biasing so that the firing for this feature enhances the other features associated with that colour in the submodules, and suppresses firing for other features and locations.

6.3.4 Predictions

Because attractor activity in the IT module is likely required to convert temporal into spatial correlations between the representation of sequential stimuli (Miyashita (1988), we predict that the model will make more spatially than temporally adjacent single binding errors and ICs. We predict this for a number of reasons: first, activity for a memory is not sustained in IT when intervening stimuli are presented therefore temporal relations could become confused; second, typically, under most circumstances, humans do not make binding errors, however, manipulation of the SOA to achieve a 50% error rate shows that a disproportionate number of errors are adjacent. This finding suggests that features migrate to the spatially closest object. In addition, we predict that the model will perform feature binding such that when a single feature is subsequently presented, the corresponding features of the loosely bound object are recalled. Thus the submodule VS operates ‘biased competition’ (Miller and Erickson, 1996) (Section 3.3.4) implemented through connections from a given module to a common inhibitory pool. Thus the more pools of the module that are active, the more active is the inhibitory pool and the more feedback inhibition will affect the pools in the module such that only the most excited pool(s) survive the competition. Subsequently, feature-based attention biases intermodular competition between V4 and IT. We are also interested in the spatial and temporal proximity of the object features to the target object features. To this end, DS contains four nodes representing the four spatial locations. In this module, spatial attention biases intermodular competition between V1, V4 and MT (Rolls and Deco, 2002).

DS comprises V1-V4-PP submodules which feed forward to PF. PF comprises v46 and d46 which are reciprocally connected. PF provides short term memory. D-PF (d46) feeds back to PP and generates an attention bias for a particular spatial location; V-PF (v46) feeds back to IT and generates an attention bias for particular features to specify the target object.

6.4 EVALUATION OF MODEL

The human visual system comprises more than 50% of posterior cortex. Some of these are involved in memory, others are specialised for memory. Chapter 3 detailed specificity in the visual and visual memory systems. Yet, despite the known specificity of particular brain areas, humans can simultaneously and precisely perceive and bind fine detail and localize visual signals despite the fact that much of the visual cortex is involved in identifying individual object features independent of their location, and other regions are dedicated to localizing objects in space independent of identity. The mechanisms underlying such phenomena are not fully understood. Nevertheless, it is clear that several brain areas operate in a distributed, parallel manner. Such complexity is inherently problematic to understand and even to model. In order to achieve computational tractability, an appropriate level of abstraction needs to be adopted.

The model system presented in this thesis comprises representations of seven neural areas. Clearly, this naivety restricts the model in terms of plausibility and generalizability, but maintains tractability. The author has concentrated on those areas considered to be primarily implicated in object feature binding in VSTM. To that end, the model has maintained biological plausibility, as it is based on physiological and biological evidence. It is also theoretically plausible in that it combines and integrates the three leading object binding theories. It is novel its integrative approach. However, because the computational model has not been implemented, it is possible only to speculate as to how and where the neural correlates of object feature binding take place. Nevertheless, the cognitive model is detailed and appropriate to describe the phenomena apparent in the literature and the behavioural data presented in this thesis. This model has been used to design a computational model, also based on existing neural models of binding, and parameters for its future implementation are given here and discussed in Subsection 7.4.

6.5 DISCUSSION AND CONCLUSIONS

The cognitive model presented here has been derived from results from the behavioural experiment (Chapter 5), which showed that spatial properties are primary. We interpreted this by suggesting that spatial properties have a dedicated processing stream whereas temporal properties rely on memory systems which are subject to interference from several sources. For example, inferotemporal cortex, implicated in maintaining a stimulus memory in the short term until it is suppressed by another, is involved in transposing temporal to spatial properties, and there is a lack of parietal activation when stimuli are presented serially at the same location (Rees et al., 1997). Thus memories for sequential information would be more difficult to accurately recall. We further suggest that composite memories in prefrontal cortex are uniquely identified by their features and their spatial location, and that spatial properties are primary and temporal properties secondary. In addition, to maintain biological plausibility and computational tractability, the model incorporates evidence from biological, psychological and physiological and computational studies (Chapters 2, 3 and 4). To this end, we suggest that object feature binding takes place as a result of external or internal feed forward activation driving competition. This in turn produces top-down bias. In such a scenario, spatial properties would have an advantage over temporal properties due to the dedicated spatial system operating in the dorsal pathway. Our model thus comprises three main sub-systems: the dorsal stream (spatial information processing), the ventral stream (object feature information processing), and prefrontal cortex for maintaining memories when intervening stimuli are presented.

This multimodular systems model incorporates the three main object binding theories, thus the cognitive model proposes that binding takes place through attention (e.g. Treisman and Gelade, 1980; Moran and Desimone, 1985; Koch and Ullman (1985) Wolfe et al. (1989) (Chelazzi et al., 1993), temporal synchrony of neural firing (e.g. Singer and Gray, 1995; Eckhorn, 1999), and coincidence (Hubel and Wiesel, 1962) or convergence (Barlow, 1972). To the author's knowledge, no other cognitive model has presented this integrated and unified solution to the problem of object feature binding.

Many models of binding lack completeness or have demonstrated some lack of biological plausibility, or both. Nevertheless, such models can provide

some important insights into the binding problem (Chapter 4). Despite the naivety of the proposed computational model, in part due to the complexity of the neural systems involved, many existing models of binding have demonstrated some lack of biological plausibility, but yet have provided some important insights into the binding problem (such as Ungerleider, 1995; Raffone and Wolters, 2001; Domijan, 2003; van der Velde and de Kamps, 2003). The author intends that the proposed cognitive model be used to inform a biologically plausible computational model to be implemented at a later date. Once implemented, results from the computational model could be used to test assumptions made by the cognitive model, and would thus lead to a further understanding of exactly where and when object features become bound.

6.6 SUMMARY

To our knowledge, the model described in this chapter is novel in that it presents an integrated theory of object feature binding. It encompasses the three current leading contenders for binding solutions in a biologically plausible multimodular attractor network system. The rationale and design of the multimodular cognitive model, and the architecture and operation of the three sub-modules have been described and schematically illustrated. Using the same architecture and connectivity, the cognitive model is intended to be used to inform the design of a computational model, realised in a modular attractor network. The author proposes implementation of the cognitive model takes place as future work.

DISCUSSION AND CONCLUSIONS

INTRODUCTION

This thesis focuses on object feature binding. Thus its concern lies with information processing in both the visual system and in visual working memory. In order to gain insight into visual and visual working memory systems, this thesis has focused on biological, physiological and psychological aspects of information processing as well as on cognitive modelling. This chapter describes and synthesises the experimental and modelling methodologies, rationales and findings. Limitations of the work carried out and proposed further work are presented prior to the contributions to knowledge and conclusions proposed in this thesis.

Recall that the aim of this doctoral thesis is to further understand how and where object feature binding takes place by means of a behavioural experiment and cognitive modelling. In order to achieve this aim, a number of objectives were established and achieved. These objectives are listed below:

- To thoroughly review and critique literature relating to object feature binding and visual short-term memory
- To outline, review and critique artificial neural systems in general and models of object feature binding in particular
- To design and conduct a behavioural experiment to test hypotheses relating to spatio-temporal aspects of object feature binding
- To develop a cognitive model based on existing evidence to explain the resultant behavioural data
- To use the cognitive model to inform the design of a biologically plausible multimodular attractor network model to be implemented at a later date.

7.1 SUMMARY OF LITERATURE

In Chapter 2, several ANN models developed in an attempt to further understand brain function, were described and compared to their biological counterparts. Descriptions included biological and artificial systems from the composition and function of single neurons and artificial nodes, to networks of networks. A brief outline of the development of ANNs, transform functions, integrate-and-fire models and learning algorithms was given. Major differences between artificial nodes and biological neurons, and between networks of biological neurons and ANNs were outlined. Chapter 3 outlined the physical aspects of the visual system, neurological evidence supporting function specificity, masking and visual working memory. Despite the visual system being typically conceptualised as a hierarchy of information processing stages working in parallel to analyse different visual attributes, humans recognise and react rapidly to internally or externally generated stimuli which may be novel, embedded in complex scenes, or both. The mechanisms behind this phenomenon remain largely unknown.

Chapter 4 was concerned with literature relevant to object feature binding and visual short term memory. In doing so proposed solutions to the binding problem were described and critiqued. The outcome was that although each proposed solution goes some way to solve the problem, none alone is sufficient. Furthermore, the role of visual short term memory (VSTM) in object feature binding was discussed given the evidence that even when an object's features are integrated correctly in perception, they can be erroneously combined in memory. We thus were interested in spatial and temporal aspects of binding in VSTM, questioning whether the processes that mediate these phenomena are automatic or independent. In the final section Chapter 4, computational models of object feature binding in VSTM were described. Although several models have been proposed and implemented, to date, no single model adequately solves the object feature binding problem.

The behavioural experiment and threats to validity were described in Chapter 5. An evaluation of the experimental methodology is outlined in Section 7.2. Chapter 6 detailed the cognitive model and proposals for a computational model

to be implemented as future work. An evaluation of the model is given in Section 7.2.2.

7.2 EXPERIMENTAL AND MODELLING METHODOLOGIES

7.2.1 Evaluation of experimental methodology

Design: The experiment was designed and conducted using a within-subjects delayed response task comprising a battery of sixty-four composite objects each with three features and four dimensions in each of the three conditions. The within-subjects design was chosen to achieve greater power, less error variance and reduce Type 2 error rate.

Participants: Each volunteer participant carried out each of the 3 experimental conditions only once. This was considered to be the most efficient for minimizing learning effects and fatigue.

Environment: The experimental environment was controlled, and the stimuli were designed for their simplicity rather than their realism, therefore the findings should not be generalised to object feature binding in real-life situations without further refinement.

Mask: The checkerboard mask was used before and between each stimulus, but it did not appear to mask the objects. A mask comprising objects from the presentation (i.e. coloured rectangles in different orientations and sizes) might have masked more successfully.

Stimulus onset asynchrony (SOA): The staircase procedure (Cornsweet, 1962) used to establish the desired SOA that would achieve a ~50% error rate for each observer in each condition for both features was implemented following pilot studies in which the error rate was set at a 50% over all conditions of the experiment proper. Error rates from the pilot studies were too low (e.g. up to 100% correct in some cases) for meaningful analyses to be conducted. Despite setting the individual SOAs after trials, the 50% error rate was not achieved in

the study proper for all observers. The result was that a mean lower than 50% was achieved in the spatial condition and a mean greater than 50% was achieved in the temporal condition. This discrepancy could affect the analysis when expected and actual values are taken into account. A possible method of achieving 50% error might be to repeat several observations with each observer until 50% errors were reported in each condition. However, this method would introduce other problems such as learning effects and fatigue. Because of the large number of presentations in each of the conditions, the slight variation in error rate from the desired 50%, in two out of three conditions, did not compromise the results of the study overall.

Size: To maintain an equal number of dimensions for each of the three features, four sizes were determined. The size ratio was 4:3:2:1 so that size 1 was 4 times larger than size 4 and so on. Verbal reports suggest that the intermediate two sizes proved difficult to distinguish when either was shown without a different object (i.e. with no comparison available) as in the temporal and spatio-temporal conditions.

Orientation: The four orientations (horizontal, vertical, right and left) selected were the most appropriate and least uncorrelated possible.

Response recording: In earlier pilot studies, responses were automatically recorded when the observers responded with key clicks which meant that an immediate initial response could not be changed. This appeared to be a problem for some observers, so the recording method was changed to manual recording to a check list. In addition, verbal comments were recorded to provide a richer analysis of the experimental findings.

Software: Macromedia Director was used to design and run the experiments because it was appropriate, available and the author had experience using this program.

Statistical Analysis: Results were analysed using the Chi-squares statistic and *t*-tests.

7.2.2 *Evaluation of modelling methodology*

To develop a plausible model, biological, physiological and psychological evidence was considered and incorporated into the design of the cognitive model. Although a variety of techniques and approaches is available, the high complexity of the independent, yet integrated neuronal systems involved in object feature binding, demand a high level abstraction. Modelling such complex systems at low levels of detail would be beyond the scope of this thesis and would likely prove intractable in terms of existing computational power and analytical methods.

Because prefrontal (PF), inferotemporal (IT) and prefrontal (PF) cortex have been implicated in visual short-term memory (VSTM), the model was designed to represent and integrate these systems. A short-term memory has been described as a dynamical system which exhibits sustained activation during a delay period (that is when no external stimuli are present). A short term memory was thus modelled as an attractor.

The cognitive model proposes that binding takes place through attention (e.g. Treisman and Gelade, 1980; Moran and Desimone, 1985; Koch and Ullman (1985) Wolfe et al. (1989) (Chelazzi et al., 1993), temporal synchrony of neural firing (e.g. Singer and Gray, 1995; Eckhorn, 1999), and coincidence (Hubel and Wiesel, 1962) or convergence (Barlow, 1972). A hierarchical feature analysis system, starting with a low-level description of the object and building features based on earlier layers in the hierarchy, enables neurons in 'higher' layers to respond to more complex combinations.

The proposed cognitive model is biologically plausible, and could be implemented as a computational model using Hebbian synaptic modifications to enable the sustained active representation of the stimulus in its absence. A 'forgetting factor' suggested for the computational model would prevent the system growing without bound, and asynchronous updating would be realistic. Despite the naivety of the proposed computational model, mainly due to the complexity of the neural systems involved, many existing models of binding are incomplete (such as Ungerleider, 1995; Raffone and Wolters, 2001; Domijan, 2003; van der Velde and de Kamps, 2003), but have provided some important insights into the binding problem

7.3 CONTRIBUTIONS TO KNOWLEDGE

Findings from work undertaken in this thesis have led to the following contributions to knowledge:

- A finding that no single existing binding theory is adequate. This contribution is derived from a comprehensive review and critique of the literature relating to object feature binding.
- A finding that no single existing binding model is adequate. Derived from a comprehensive review and critique of the literature relating to computational modelling of the binding problem and visual short term memory.
- The finding that spatial properties are key in binding, and that temporal properties are used when no spatial properties are available. Derived from the design of a novel behavioural experiment incorporating and combining both spatial and temporal object properties as well as featural proximity; and analyses of the resultant data.
- A novel multimodular cognitive model describes object feature binding mechanisms in visual short term memory. Derived from findings stated above.
- A unified object binding theory which states that object feature binding takes place in visual short term memory by means of attention to object features or spatial location depending which is more strongly activated. Activation is sustained by means of recurrent connectivity.
- A proposed novel computational model based on the cognitive model for later implementation.
- A unified object feature binding theory derived from the points stated above. The model proposes that object feature binding takes place through a combination of hierarchical feature analysis, spatial attention and temporal synchrony.

Findings from the behavioural experiment support the hypotheses proposed in Chapter 1, and in doing so support work of several others in neuroscience and psychology.

In the Spatio-temporal condition:

- spatial properties aid object features binding
- when both spatial and temporal properties are available, temporal properties hinder binding (please see Section 5.4, p.170)
- the differences between observed and expected spatial errors and ICs were significantly greater than between observed and expected temporal errors and ICs
- there were significantly more spatially than temporally adjacent errors and ICs.

When comparing temporal errors and ICs in the Spatio-temporal and Temporal conditions there were:

- significantly greater differences in the Spatio-temporal condition between observed and expected errors and ICs
- significantly more adjacent errors and ICs in the Spatio-temporal condition.

Overall:

- Features from spatially adjacent objects are highly likely to migrate, whereas features from temporally adjacent objects are much less likely to migrate.
- We propose that composite memories in prefrontal cortex are uniquely identified by both their features and their spatial location. In this way, a scene comprising a moving object or objects (or in the case of the experiment, rapid sequential presentations) is represented by synchronous activity for the object at the different spatial, but not temporal positions.
- Our study supports findings of Treisman and Schmidt (1982), Prinzmetal et al. (1995) and Cohen and Ivry (1990) that spatial properties are key in object feature binding.
- Both spatial and temporal coherence, to differing degrees, may be involved in object feature binding (Maunsell and Gibson, 1992; Blake

and Yang, 1997), perhaps driven by distinct mechanisms (Holcombe and Cavanagh, 1999).

- Attention is an emergent property of recall (Desimone and Duncan, 1995).
- ICs occur mostly when no unique spatial location exists as a result an incorrect perception of feature location. This finding supports (Prinzmetal et al., 1995).
- ICs are more likely to occur between adjacent items (Cohen and Ivry, 1990).

The model:

The author's concern was with the mechanisms of object feature binding. That is, how does information, distributed in patterns of neural firing, result in coherent representations in the absence of the original stimulus, and what is the effect of proximity and spatial and temporal properties? How does object feature binding occur in visual short-term memory (VSTM)?

The derived model is a novel multimodular network comprising information processing in the visual streams (dorsal and ventral) and prefrontal cortex. The visual streams incorporate sub-modules representing visual areas involved in both spatial and object visual processing. The submodules are independent but integrated by means of recurrent connectivity. This is described in details in Chapter 6.

7.4 LIMITATIONS AND FURTHER WORK

Object feature binding is inherently interesting. However, in retrospect the original project was ambitious and likely beyond the scope of a single PhD thesis. The body of literature on the subject is vast. Furthermore, the author now considers that an attempt to encompass computational neuroscience and behavioural experimentation in a single project was naïve. Since project conception, and revision, the focus has changed from computational neuroscience to cognitive neuroscience which is not only less ambitious, but also more tractable. Nevertheless, the demand to understand the neural mechanisms underlying object feature binding remains. The author, in this thesis, has built on

existing evidence integrated with the behavioural findings from the experiment (Chapter 5), to produce cognitive model describing a unified theory of object feature binding. She considers this model can be used as a basis for a computational model that could be implemented at a later date.

The following section describes further work beyond the time scope of this thesis, but which the author considers worthwhile and potentially insightful.

7.4.1 *Further experimental work*

The behavioural experimental design was finalised after pilot studies had been conducted and modifications made to lead to a more appropriate, efficient and useful design. In retrospect, the author is aware of the complexity of the experimental design used in this study. Analysis and hypothesis testing of the resultant data were therefore problematic. A series of simpler experiments each dealing with a particular hypothesis from general to specific could be designed. Data analysis and hypothesis testing would then be more straightforward. This would allow replication, for theory testing to be carried out more easily.

In retrospect the author would make further modifications which are described below.

Coloured rectangles were used in the behavioural experiment. The colours used could be less correlated to reduce any possible confusion among them. The author used four different sized rectangles. Using three, rather than four might prove easier for observers to discriminate among as some observers pointed to the fact it was difficult to discriminate between intermediate sized objects when they were presented sequentially (and hence without a large or small) referent. Four different shapes, rather than sizes could be used to eliminate this problem while maintaining 4 dimensions for each feature.

The time lag between each object presentation in the temporal and spatio-temporal conditions could be varied so that rather than have equal time differences between presentation of each stimulus, the stimulus could be pseudo-randomly presented adjacent to, distant and more distant *in time* from the target. This proposed revision to the temporal presentation of each object would more closely match the spatial presentation of each object.

The experiments provided a large body of data. These could be analysed along several additional dimensions. For example, each feature could be analysed

individually to discover frequencies and interrelationships of each dimension of both features. It would be interesting to discover whether the temporal sequence of objects affected the response. For example, did a large stimulus dominate if presented in a particular temporal position? Would a purple stimulus be less confused with the other colour dimensions as it is less correlated? Furthermore, despite the checkerboard mask, the author suspects that the post-cue colour was often carried through to the next presentation so that it determined the target object in the following presentation.

7.4.2 Further modelling work

The cognitive model provides a starting point for a computational model. Because of the complexity of the biological involved in object feature binding in a delayed match to stimulus task, the author needed to adopt a high level of abstraction when designing the models. The following sections describe the enhancements that might be made in order test and evaluate the assumptions and predictions stated in this thesis, and lead to a clearer understanding of object feature binding in visual short-term memory. Integrate-and-fire model neurons were proposed to add to the sensitivity of the model. In order to be more biologically realistic, the model should implement a softer competition than winner-take-all (WTA) such as the Softmax function (Bridle, 1990). Softer competition should result in faster firing neurons firing relatively much faster than the slower ones after competition.

The current model used short-range excitation and long-range inhibition between the neurons so that the network could develop topological maps (e.g. von der Malsburg, 1973). In the brain, short-range excitatory connections between neurons, and long-range inhibitory connections mediated by interneurons are common. This characteristic not only groups ensembles of neurons with similar response profiles, it also reduces the length of connections between them to enable fast information exchange.

A more ambitious task would be to extend the model to encompass other related brain regions such as hippocampus complex which has also been implicated in visual short-term memory. Ultimately, the system could be further developed so that representations of the entire visual system from retinal input to behavioural response are modelled.

7.4.3 *Summary of further work*

As stated above, the experimental design used in this study is complex. A series of experiments could be carried out to test each hypothesis individually and build a theory of object feature binding.

Notwithstanding the suggestions above, neural modelling is by definition an abstraction. In computational neuroscience, the intention is not to recreate a brain, but to understand the processes therein. Were the suggestions for further work proposed in Section 7.3 carried out, the behavioural experiment could provide further insights into object feature binding, and the model would be both more plausible and therefore more able to explain how object feature binding takes place. However, the author believes that a simple model that represents complex systems is desirable. The model developed as part of this thesis is a starting point towards this and ultimately towards a further understanding of object feature binding.

7.5 AUTHOR'S PUBLICATIONS

To date, three publications have been accepted from this work:

Mair, C. (2006) Towards a Further Understanding of Object Feature Binding: a Cognitive Neuroscience Perspective. Poster presented at BPS Undergraduate/Graduate Student Conference, Bournemouth University, 22nd March 2006

Mair, C., Shepperd, M., Cartwright, M., Kirsopp, C., Premraj, R. and Heathcote, D. (2004) Understanding Object Feature Binding through Experimentation as a Precursor to Modelling. Proceedings of the 8th Neural Computation and Psychology Workshop, Connectionist Models of Cognition and Perception II. H. Bowman and C. Labiouse (Eds) Progress in Neural Processing 15. World Scientific, 295-305.

Mair, C. (2003) Towards a Further Understanding of Object Feature Binding. Proceedings of the 7th International Conference on Cognitive and Neural Systems (ICCNS), Boston, USA, May 2003.

The publications below are not connected with this thesis, but have been part of my full-time work as Research Assistant.

Journal publications

Song, Q. Shepperd, M.J., Cartwright, M. and **Mair, C. (2006)** Software Defect Association Mining and Defect Correction Effort Prediction, *IEEE Transactions on Software Engineering* 32(2), pp69-82.

Mair, C., Shepperd, M.J. and Jorgensen, M. (2005) An analysis of data sets used to train and validate cost prediction systems, *ACM SIGSOFT Software Engineering Notes*, 30(4), pp1-6.

Mair, C., Kadoda, G., Lefley, M., Phalp, K., Schofield, C., Shepperd, M.J., and Webster, S. (2000) An investigation of machine learning based prediction systems, *Journal of Systems Software* 53(1), pp23-29.

Conference Publications

Mair, C. and Shepperd, M.J. (2006) Looking at Comparisons of Regression and Analogy-based Software Project Cost Prediction, Proceedings of *International Conference on Software Engineering Research and Practice (SERP'06)*, Las Vegas, June 2006.

Mair, C., Shepperd, M.J. and Forselius, P. (2006) An Empirical Analysis of Software Productivity, Proceedings of *3rd Software Measurement European Forum (SMEF 2006)*, Rome, 2006.

Mair, C. and Shepperd, M.J. (2005) The Consistency of Empirical Comparisons of Regression and Analogy-based Software Project Cost Prediction, Proceedings of 4th *International Symposium on Empirical Software Engineering (ISESE)*, Noosa Heads, Australia, November 2005.

Song, Q., Shepperd, M.J. and **Mair, C.** (2005) Using Grey Relational Analysis to Predict Software Effort with Small Data Sets, Proceedings of the *11th IEEE International Software Metrics Symposium (Metrics05)*, Como, Italy, September 2005.

Mair, C. and Shepperd, M.J. (2004) Making Software Cost Data Available for Meta-Analysis, Proceedings of the *International Conference on Empirical Assessment of Software Engineering (EASE)*, Edinburgh, May 2004.

Premraj, R., Twala, B. and **Mair, C.** (2004) Productivity of Software Projects by Business Sector: An Empirical Analysis of Trends. Proceedings of the *10th International Software Metrics Symposium*, Chicago, USA, 14-18 Sept., 2004.

Mair, C. and Shepperd, M.J. (1999) An investigation of rule induction based prediction systems,' Proceedings of *IEEE ICSE Workshop, Empirical Studies of Software Development and Evolution*, Los Angeles, May 1999.

Bennett, K., Burd, E., Kemerer, C., Lehman, M., Lee, M., Madachy, R., **Mair, C.**, Sjoberg, D. and Slaughter S. (1999) Empirical Studies of Evolving Systems. *Empirical Software Engineering*, 4, 4.

REFERENCES

- Abbott, L.F., & Nelson, S.B. (2000). Synaptic plasticity: taming the beast. *Nature Neuroscience*, 3 (1178-1183.)
- Abeles, M. (1982). Local Cortical Circuits. Berlin: Springer Verlag.
- Abeles, M. (1991). Corticonics: neural circuits of the cerebral cortex. (Cambridge.: CUP.
- Abeles, M. (1994). Firing rates and well-timed events. In: E. Domany, Schulten, K., and van Hemmen, J. L., (Eds.) (Ed.) *Models of Neural Networks*, 2 (p. Chapter 3.). New York: Springer.
- Abeles, M., Bergman, H., Gat, I., Meilijson, I., Seidemann, E., Tishby, N., & Vaadia, E. (1995). Cortical activity flips among quasi stationary states. *Proceedings of the National Academy of Sciences, USA.*, 92, 8616-8620.
- Abeles, M., Bergman, H., Margalit, E., & Vaadia, E. (1993). Spatiotemporal firing patterns in the frontal cortex of behaving monkeys. *Journal of Neurophysiology*, 70, 1629-1638.
- Adrian, E.D. (1926). The impulses produced by sensory nerve endings. *Journal of Physiology (London)*, 61, 49-72.
- Amari, S. (1977). Dynamics of pattern formation in lateral inhibition type neural fields. *Biological Cybernetics*, 27, 77-87.
- Amit, D., J. (1989). Modeling Brain Function: the world of attractor neural networks. (Cambridge: CUP.
- Amit, D.J. (1995). The Hebbian paradigm reintegrated: local reverberations as internal representations. *Behavioural and Brain Sciences.*, (18), 617-657.
- Amit, D.J., & Brunel, N. (1997). Model of global spontaneous activity and local structured delay activity during delay periods in the cerebral cortex. *Cerebral Cortex*, 7, 237-252.
- Amit, D.J., Brunel, N., & Tsodyks, M.V. (1994). Correlations of cortical Hebbian reverberations: experiment versus theory. *Journal of Neuroscience*, 14, 6435-6445.
- Amitai, Y., Friedman, A., Connors, B., & Gutnick, M. (1993). Regenerative activity in apical dendrites of pyramidal cells in neocortex. *Cerebral Cortex.*, 3, 26-38.
- Anderson, D., & McNeill, G. (1992). Artificial neural networks technology. *DACS State of Art Report* (pp. 1-87). Griffiss, New York: Rome Laboratory.
- Arbib, M.A. (1987). A View of Brain. In: E. F.E. Yates (Ed.) *Theory in Self-Organizing Systems: The Emergence of Order* (pp. 279-311.): Plenum.
- Arbib, M.T.M.P., 1995 (1998). The Handbook of Brain Theory and Neural Networks. (Boston, Mass, USA.: MIT Press.
- Ashby, F.G., Prinzmetal, W., Ivry, R., & Maddox, W.T. (1996). A formal theory of feature binding in object perception. *Psychological Review.*, (103), 165-192.
- Baddeley, A.D. (1986). Working Memory. (NY: OUP.
- Balint, R. (1909). Die Seelenlahmung des schauens, optische Ataxie, raumliche Storung der Aufmerksamkeit. *Monatsschrift Fur Psychiatrie und Neurologie*, 25, 51-81.

- Barbas, H., & Pandya, D.N. (1989). Architecture and intrinsic connections of the prefrontal cortex in the rhesus monkey. *Journal of Computational Neurobiology*, (286), 353-375.
- Barlow, H. (1989). Unsupervised learning. *Neural Computation*, 1, 295-311.
- Barlow, H.B. (1953). Summation and inhibition in the frog's retina. *Journal of Physiology*, 119, 69-88.
- Barlow, H.B. (1972). Single units and cognition: a neuron doctrine for perceptual psychology. *Perception*, 1, 371-394.
- Beaulieu, C., Kisvarday, Z., Somogyi, P., Cynader, M., & Cowey, A. (1992). Quantitative distribution of GABA-immunopositive and immunonegative neurons and synapses in the monkey striate cortex (area 17). *Cerebral Cortex*, 2, 295-309.
- Bernander, O., Koch, C., & Douglas, R.J. (1994). Amplification and linearization of distal synaptic input to cortical pyramidal cells. *Journal of Neurophysiology*, 72 (2743-2752.)
- Bi, G., & Poo, M. (1998). Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength and post-synaptic cell type. *Journal of Neuroscience*, 18, 10464-10472.
- Bienenstock, E. (1995). A model of neocortex. *Computation in Neural Systems*, 6, 179-224.
- Blake, R., & Lee, S.-H. (2000). Temporal structure in the input to vision can promote spatial grouping. In: S.-H. Lee, H.H. Bulthoff, & T. Poggio (Eds.), *BMVC* (pp. 635-653). Berlin: Springer-Verlag.
- Blake, R., & Yang, Y. (1997). Spatial and temporal coherence in perceptual binding. *Proceedings of the National Academy of Science, USA*, 94 (June 1997), 7115-7119.
- Breitmeyer, B.G. (1984). Recent models and findings in backward visual masking: a comparison, review and update. *Perception and Psychophysics*, 62, 1572-1595.
- Bressler, S.L., & Kelso, J.A.S. (2001). Cortical coordination dynamics and cognition. *Trends in Cognitive Science*, 5 (26-36)
- Bridle, J.S. (1990a). Probabilistic interpretation of feedforward classification network outputs, with relationships to statistical pattern recognition. In: F. Fogelman Soulie, & J. Herault (Eds.), *Neurocomputing: Algorithms, Architectures and Applications* (N.Y.: Springer Verlag.
- Bridle, J.S. (1990b). Probabilistic interpretation of feedforward classification network outputs, with relationships to statistical pattern recognition. In: J. Herault (Ed.) *Neurocomputing: Algorithms, Architectures and Applications* (N.Y.: Springer Verlag.
- Broadbent, D.E., & Broadbent, M.H.P. (1987). From Detection to Identification - Response to Multiple Targets in Rapid Serial Visual Presentation. *Perception & Psychophysics*, 42 (2), 105-113.
- Brockmole, J.R., Irwin, D.E., & Wang, R.F. (2003). The locus of spatial attention during the temporal integration of visual memories and visual percepts. ,. *Psychonomic Bulletin & Review*, 10, 510-515.
- Brockmole, J.R., Wang, R.F., & Irwin, D.E. (2002). Temporal integration between visual images and visual percepts. *Journal of Experimental Psychology: Human Perception & Performance*, 28, 315-334.

- Brunel, N., & Nadal, J.-P. (1997). Modelling memory: what do we learn from attractor neural networks. *Symposium: Memory: from neuron to cognition*. (Paris: Académie des Sciences,).
- Bugmann, G., & Taylor, J.G. (1994). The role of short-term memory in visual information propagation. *Extended Abstract Book of the International Symposium on Dynamics of Neural Processing*. (pp. 132-136.). Washington.
- Bugmann, G., & Taylor, k.J.G. (2005). A model of visual backward masking. *BioSystems*, (79), 151-158.
- Buhl, E.H., Halasy, K., & Somogyi, P. (1994). Diverse sources of hippocampal unitary inhibitory postsynaptic potentials and the number of synaptic release sites. *Nature*, 368, 823-828.
- Buzsáki, G., & Chrobak, J.J. (1995). Temporal structure in spatially organized neuronal ensembles: a role for interneurons networks. *Current Opinion in Neurobiology*, 5, 504-510.
- Camperi, M., & Wang, X.-J. (1998). A model of visuospatial working memory in prefrontal cortex: recurrent network and cellular bistability. *Journal of Computational Neuroscience*, (5), 383-405.
- Cave, K.R., & Zimmerman, J.M. (1997). Flexibility in spatial attention before and after practice. *Psychological Science*, 8, 399-403.
- Chelazzi, L., Duncan, J., Miller, E.K., & Desimone, R. (1998). Responses of Neurons in Inferior Temporal Cortex During Memory-Guided Visual Search. *Journal of Neurophysiology*, 80 (6), 2918-2940.
- Chelazzi, L., Miller, E., Duncan, J., & Desimone, R. (1993). A neural basis for visual search in inferior temporal cortex. *Nature (London)*, (363), 345-347.
- Cohen, A., & Ivry, R. (1991). Density effects in conjunction search: evidence for a coarse location mechanism of feature integration. *Journal of Experimental Psychology: Human Perception and Performance*, 17, 891-901.
- Cohen, J.D., & Ivry, R. (1990). Illusory conjunctions as a function of distance between objects inside and outside the focus of attention. *Journal of Experimental Psychology: Human Perception and Performance*, (16), 317-331.
- Cohen, J.D., Perlstein, W.M., Braver, T.S., Nystrom, L.E., Noll, D.C., Jonides, J., & Smith, E.E. (1997). Temporal dynamics of brain activity during a working memory task. *Nature*, (386), 604-608.
- Compte, A., Brunel, N., Goldman-Rakic, P., & Wang, X.-J. (2000). Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cerebral Cortex*, 10 (9), 910-923.
- Corbetta, M., Shulman, G.L., Miezin, F.M., & Peterson, S.E. (1995). Superior parietal cortex activation during spatial attention shifts and visual feature conjunction. *Science*, 270, 802-805.
- Corchs, S., & Deco, G. (2001). A neurodynamical model for selective visual attention using oscillators. *Neural Networks*, (14), 981-990.
- Cornsweet, T.N. (1962). The staircase method in psychophysics. *American Journal of Psychology*, 75 (pp. 485-491).
- Courtney, S.M., Ungerleider, L.G., Keil, K., & Haxby, J.V. (1997). Transient and sustained activity in a distributed neural system for human working memory. *Nature*, (386), 608-611.

- Crick, F., & Koch, C. (1998). Constraints on cortical and thalamic projections: the no strong loops hypothesis. *Nature (London)*, (391), 245-250.
- Critchley, M. (1953). *The Parietal Lobe*. (New York: Hafner).
- Czerwinski, M., Lightfoot, N., & Shiffrin, R.M. (1992). Automatization and training in visual search. *American Journal of Psychology*, 105, 271-315.
- Damasio, A., Damasio, H., & Van Hoesen, G.W. (1982). Prosopagnosia: anatomic basic and behavioural mechanisms. *Neurology*, (32), 331-341.
- Damasio, A., Yamada, T., Damasio, H., Corbett, J., & McKee, J. (1990). Central achromatopsia: behavioural, anatomic, and physiological aspects. *Neurology*, (30), 1064-1071.
- Davey, N., & Adams, R. (2002). Stochastic dynamics and high capacity associative memories. *Ninth International Conference on Neural Information Processing* (pp. 1666-1671).
- Dayan, A., & Thomas, J.R. (1995). Development of automatic and effortful processes in memory for spatial location of movement. *Human Performance*, 8, 51-66.
- de Kamps, M., & van der Velde, F. (2001). Using a recurrent network to bind form, colour and position into a unified percept. *Neurocomputing*, 38-40, 523-528.
- DeAngelis, G.C., Ohzawa, I., & Freeman, R.D. (1995). Receptive field dynamics in the central visual pathways. *Trends in Neurosciences*, (18), 451-458.
- Deco, G., & Lee, T.S. (2002). A unified model of object and spatial attention based on intercortical biased competition. *Neurocomputing*, (44-46), 775-781.
- Delvenne, J.-F., & Bruyer, R. (2004). Does visual short term memory store bound features. *Visual Cognition*, 11 (1), 1-27.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective attention. *Annual Review of Neuroscience*, 18, 193-222.
- Desimone, R., & Schein, S.J. (1987). Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. *Journal of Neurophysiology*, (57), 835-868.
- Desimone, R., Schein, S.J., Moran, J., & Ungerleider, L.G. (1985). Contour, colour and shape analysis beyond the striate cortex. *Vision Research*, 25, 441-452.
- D'Esposito, M. (2001). Functional neuroimaging of working memory. In: R. Cabeza, & A. Kingstone (Eds.), *Handbook of Functional Neuroimaging of Cognition* (pp. 293-327). Cambridge, Mass.: MIT Press.
- Destexhe, A., Mainen, Z.F., & Sejnowski, T. (1994). An efficient method for computing synaptic conductances based on a kinetic model of receptor binding. *Neural Computation*, 6, 14-18.
- DeValois, R.L., & DeValois, K.K. (1988). *Spatial Vision*. (NY: OUP).
- Di Lollo, V. (1980). Temporal integration in visual memory. *Journal of Experimental Psychology: General*, 109, 75-97.
- Domijan, D. (2003). A mathematical model of persistent neural activity in human prefrontal cortex for visual feature binding. *Neuroscience Letters*, 350, 89-92.
- Donk, M. (1999). Illusory conjunctions are an illusion: the effects of target-nontarget similarity on conjunction and feature errors. *Journal of Experimental Psychology: Human Perception and Performance*, 25 (5), 1207-1233.

- Dowling, J.E. (1987). *The Retina: An Approachable Part of the Brain*. (Cambridge, MA.: Harvard University Press,.
- Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, 113, 501-517.
- Duncan, J., & Humphreys, G. (1989). Visual search and stimulus similarity. *Psychological Review*, (87), 272-300.
- Duncan, J., Humphreys, G., & Ward, R. (1997b). Competitive brain activity in visual attention. *Current Opinion in Neurobiology*, 7, 255-261.
- Duncan, J., Ward, R., & Shapiro, K. (1994). Direct measurement of attentional dwell time in human vision. 1994, *Nature* (369), 313-315.
- Eckhorn, R. (1994). Oscillatory and non-oscillatory synchronizations in the visual cortex of cat and monkey. In: B. Lutkenhoner (Ed.) *Oscillatory Event-Related Brain Dynamics* (p. 115-134.). New York: Plenum.
- Eckhorn, R. (1999). Neural mechanisms of visual feature binding investigated with microelectrodes and models. *Visual Cognition*, 6 (3-4), 231-265.
- Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M., & Reitboeck, H.J. (1988). Coherent Oscillations - a Mechanism of Feature Linking in the Visual-Cortex - Multiple Electrode and Correlation Analyses in the Cat. *Biological Cybernetics*, 60 (2), 121-130.
- Edelman, G., M., & Tononi, G. (2000). *Consciousness: How Matter Becomes Imagination*. (London: The Penguin Press.
- Edelman, G.M. (1978). Group selection and phasic reentrant signalling: a theory of higher brain function. In: G.M.E.a.V.B. Hardcastle (Ed.) *The Mindful Brain* (pp. 51-100). Cambridge, Mass.: MIT Press.
- Edelman, G.M. (1987). *Neural Darwinism: The Theory of Neuronal Group Selection*. (New York: Basic Books.
- Edelman, G.M. (1989). *The Remembered Present: A Biological Theory of Consciousness*. (New York: Basic Books.
- Engel, A.K., Fries, P., König, P., Brecht, M., & Singer, W. (1999). Temporal Binding, Binocular Rivalry, and Consciousness. *Consciousness and Cognition*, 8, 128-151.
- Engel, A.K., Fries, P., & Singer, W. (2001). Dynamic predictions: oscillations and synchrony in top-down processing. *Nature*, 2
- Engel, A.K., König, P., Kreiter, A.K., & Singer, W. (1991a). Interhemispheric synchronization of oscillatory neural responses in cat visual cortex. *Science*, 252, 1177-1179.
- Engel, A.K., König, P., & Singer, W. (1992). The Functional Nature of Neuronal Oscillations - Reply. *Trends in Neurosciences*, 15 (10), 387-388.
- Engel, A.K., König, P., & Singer, W. (1991b). Direct physiological evidence for scene segmentation by temporal coding. *National Academy of Science USA*, 88 (pp. 9136-9140).
- Engel, A.K., Roelfsema, P.R., Fries, P., Brecht, M., & Singer, W. (1997). Role of the temporal domain for response selection and perceptual binding. *Cerebral Cortex*, 7 (6), 571-582.
- Enns, J.T., & DiLollo, V. (2001). What's new in visual masking. *Trends in Cognitive Science*, 4, 345-352.
- Fahle, M. (1993). Figure ground discrimination from temporal information. *Proceedings Biological Sciences*, 254, 199-203.

- Fahle, M., & Koch, C. (1995). Spatial Displacement, But Not Temporal Asynchrony, Destroys Figural Binding. *Vision Research*, 35 (4), 491-494.
- Farah, M.J. (2000). *The Cognitive Neuroscience of Vision*. (Oxford: Blackwell Publishers.
- Feldman, J.A. (1985). Four frames suffice: a provisional model of vision and space. *Behavioural and Brain Sciences*, (8), 265-289.
- Felleman, D., & Van Essen, D.C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, 1, 1--47.
- Felleman, D.J., Burkhalter, A., & Van Essen, D.C. (1997). Cortical connections of area V3 and VP of macaque monkey extrastriate visual cortex. *Journal of Computational Neurology*, (379), 21-47.
- Ferster, D. (1986). Orientation selectivity of synaptic potentials in neurons of cat primary visual cortex. *Journal of Physiology (London)*. 6, 1284-1301.
- Frackowiak, R.S.J. (1997).
- Francis, G. (2000). Quantitative theories of metacontrast masking. *Psychological Review*, 107, 768-785.
- Freeman, Ohzawa, & Walker (2001). Beyond the classical receptive field in the visual cortex. *Progress in Brain Research*, (134), 157-170.
- Friedman, H.R., & Goldman-Rakic, P. (1988). Activation of the hippocampus and dentate gyrus by working-memory: a 2- deoxyglucose study of behaving rhesus monkeys. 8, (8), 4693-4706.
- Friedman-Hill, S., Robertson, L.C., & Treisman, A. (1995). Parietal contributions to visual feature binding: evidence from a patient with bilateral lesions. *Science*, (9), 277-283.
- Friedman-Hill, S.R., Robertson, L.C., Desimone, R., & Ungerleider, L.G. (2003). Posterior parietal cortex and the filtering of distractors. *Neuroscience*, 100 (7), 4263-4268.
- Frien, A., Eckhorn, R., Bauer, R., Woelbern, T., & Kehr, H. (1994). Stimulus-Specific Fast Oscillations at Zero Phase Between visual areas V1 and V2 of awake monkey. *Neuroreport*, 5, 2273-2277.
- Fujii, H., Ito, H., Aihara, K., Ichinose, N., & Tsukada, M. (1996). Dynamical cell assembly hypothesis-theoretical plausibility of spatio-temporal coding in the cortex. *Neural Networks*, 9, 1303-1350.
- Fukushima, K. (1975). Cognitron: a self-organising neural network. *Biological Cybernetics*, 20, 121-136.
- Fukushima, K., & Miyake, S. (1982). Neocognitron: a new algorithm for pattern recognition tolerant of deformations and shifts in position. *Pattern Recognition*, 15 (6), 455-469.
- Funahashi, S., Bruce, C.J., & Goldman-Rakic, P. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of Neurophysiology*, (61), 331-349.
- Funahashi, S., Bruce, C.J., & Goldman-Rakic, P.S. (1993). Dorsolateral prefrontal lesions and oculomotor delayed-response performance: evidence for mnemonic "scotomas". *Journal of Neuroscience*. 1993, 13 (4), 1479-1497.
- Fuster, J. (1988). *The Prefrontal Cortex*. N.Y.: Raven.
- Fuster, J. (1995). *Memory in the cerebral cortex*. (Boston: MIT Press.
- Fuster, J.M. (1997). *The Prefrontal Cortex: Anatomy, Physiology, and Neuropsychology of the Frontal Lobe*. New York: Lippincott-Raven.
- Fuster, J.M., & Alexander, G.E. (1971). *Science*, 173, 652-654.

- Geisler, W., S., & Super, B.J. (2000). Perceptual organization of two-dimensional patterns. *Psychological Review*, 107, 677-708.
- Gerstmann, J. (1924 (1971)). Fingeragnosie: Eine umschriebene Störung der Orientierung am eigenen Körper. *Wiener Klinische Wochenschrift*, 1924, 37: 1010-1012. In: *Translation in Archives of Neurology*, 24 (pp. 475-476.). Chicago.
- Gerstner, W., & Kistler, W.M. (2002). Mathematical formulations of Hebbian learning. *Biological Cybernetics*, 87, 404-414.
- Ghose, G.M., & Freeman, R.E. (1997). Intracortical connections are not required for oscillatory activity in the visual cortex. *Visual Neuroscience*, 14, 963-979.
- Giersch, A., Humphreys, G.W., Boucart, M., & Kovacs, I. (2000). The computation of occluded contours in visual agnosia: evidence for early computation prior to shape-binding and figure-ground coding. *Cognitive Neuropsychology*, (17), 731-759.
- Gilbert, C.D. (1993). Circuitry, architecture and functional dynamics of visual cortex. *Cerebral Cortex*, (3), 373-386.
- Gilbert, C.D., & Wiesel, T.N. (1983). Clustered intrinsic connections in cat visual cortex. *Journal of Neuroscience*, 3, 1116-1133.
- Gilbert, C.D., & Wiesel, T.N. (1989). Columnar specificity of intrinsic horizontal and cortico-cortical connections in cat visual cortex. *Journal of Neuroscience*, (9), 2432-2442.
- Gisiger, T., Dehaene, S., & Changeux, J.-P. (2000). Computational models of association cortex. *Current Opinion Neurobiology*, 10, 250-259.
- Goldman-Rakic, P. (1987). Circuitry of the prefrontal cortex and the regulation of behaviour by representational knowledge. In: F.P.a.V.M. (Eds.) (Ed.) *Handbook of Physiology* (pp. 373-417). Bethesda: American Physiological Society.
- Goldman-Rakic, P. (1995). The cellular basis of working memory. *Neuron*, (14), 477-485.
- Goldman-Rakic, P. (1996). Regional and cellular fractionation of working memory. *Memory: recording experience in cells and circuits*, 93 (pp. 13473-13480). Irvine, CA.: NAC.
- Goldman-Rakic, P.S., Lidow, M.S., & Gallager, D.W. (1990). *Journal of Neuroscience*, 10, 2125-2138.
- Goodale, M.A., & Milner, A.D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, (15), 20-25.
- Gray, C.M. (1994). Synchronous Oscillations in Neuronal Systems: Mechanisms and Functions. *Journal of Computational Neuroscience*, 1, 11-38.
- Gray, C.M. (1999). The temporal correlation hypothesis of visual feature integration: still alive and well. *Neuron*, 24, 31-47.
- Gray, C.M., Engel, A.K., König, P., & Singer, W. (1989a). Stimulus-Specific Neuronal Oscillations in Cat Visual-Cortex - Receptive-Field Properties and Feature Dependence. *Perception*, 18 (4), 510-511.
- Gray, C.M., König, P., Engel, A.K., & Singer, W. (1989b). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronisation which reflects global stimulus properties. *Nature*, (338), 334-337.
- Gray, C.M., & Singer, W. (1989). Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *National Academy of Science USA*, 86 (pp. 1698-1702).

- Grinvald, A., Lieke, E.E., Frostig, R.D., & Hildensheim, E.R. (1994). Cortical point spread function and long-range lateral interactions revealed by real-time optical imaging of macaque monkey primary visual cortex. *Journal of Neuroscience*, 14, 2545-2568.
- Gross, C.G. (1976). Inferotemporal cortex in vision. In: J.M.S.a.A.N. Epstein (Ed.) *Progress in Psychobiology and Physiological Psychology* (pp. 77-123). New York: Academic Press.
- Grossberg, S. (1980). How does the brain build a cognitive code? *Psychological Review*, (87), 1-51.
- Grossberg, S. (1987). Competitive learning: from interactive activation to adaptive resonance. *Cognitive Science*, 11, 23-63.
- Grossberg, S. (1988). Nonlinear neural networks: principles, mechanisms, and architectures. *Neural Networks*, 1, 17-61.
- Grossberg, S., & Grunewalde, A. (1997). Cortical synchronization and perceptual framing. *Journal of Cognitive Neuroscience*, 9, 117-132.
- Grossberg, S., Mingolla, E., & Ross, W. (1994). A neural theory of attentive visual search. *Psychological Review*, 101 (470-489)
- Gruber, O., & von Cramon, D.Y. (2003). The functional neuroanatomy of human working memory revisited: evidence from 3-T fMRI studies using classical domain-specific interference tasks. *Neuroimage*, 19, 797--809.
- Hansel, D., & Sompolinsky, H. (1998). Modeling feature selectivity in local cortical circuits. In: C.K.a.I. Segev (Ed.) *Methods in Neuronal Modeling* (pp. 499-567). Cambridge, Massachusetts: MIT Press.
- Hartline, H.K. (1940). The receptive fields of optic nerve fibers. *American Journal of Physiology*, 130, 690-699.
- Hasher, L., & Zacks, R.T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, 108, 356-388.
- Haxby, J., Grady, C., Horwitz, B., Ungerleider, L., Mishkin, M., Carson, R., Herscovitch, P., Schapiro, M., & Rapoport, S. (1991). Dissociation of the spatial and object Visual Processing Pathways in Human Extrastriate Cortex. *National Academy Science USA*, 88 (pp. 1621-1625).
- Haxby, J.V., Ungerleider, L.G., Horwitz, B., Rapport, S.I., & Grady, C.L. (1995). Hemispheric differences in neural systems for face working memory. *Human Brain mapping*, (3), 68-82.
- Hayes, S.M., Ryan, L., Schnyer, D.M., & Nadel, L. (2004). An fMRI study of episodic memory: Retrieval of object, spatial, and temporal information. *Behavioral Neuroscience*, 118, 885-896.
- Healy, A.F. (1975). Coding of temporal-spatial patterns in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 14, 481-495.
- Heathcote, D., Walker, P., & Hitch, G.J. (1994). Feature independence and the recovery of feature conjunctions. *Journal of General Psychology*, 121 (3), 253-266.
- Hebb, D.O. (1949). *The Organization of Behavior: A neuropsychological theory*. (New York: Wiley).
- Heidmann, T., & Changeux, J.-P. (1982). Un modèle moléculaire de régulation d'efficacité d'une synapse chimique au niveau postsynaptique. *C. R. Academy of Science Paris*, 295, 665-670.
- Heywood, C.A., & Cowey, A. (1987). On the role of cortical area V4 in the discrimination of hue and pattern in macaque monkeys. *Journal of Neuroscience*, (7), 2601-2617.

- Hilgetag, C., O'Neill, M.A., & Young, M.P. (2000). *Philosophical Transactions of the Royal Society London.*, B355, 71-89.
- Hill, O.W., & Maodab, M.H. (1995). Spatial information and temporal representation in memory. *Perceptual and Motor Skills*, 81, 1339-1343.
- Hill, O.W., & Stuckey, R.W. (1993). Spatial coding of information on temporal order in short term memory. *Perceptual and Motor Skills*, 76, 110-124.
- Hodgkin, A.L., & Huxley, A.F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology*, 117, 500-544.
- Holcombe, A.O., & Cavanagh, P. (1999). The temporal resolution of visual feature binding. *The 7th Annual Workshop on Object Perception and Memory* (
- Hommel, B. (1998). Event files: Evidence for automatic integration of stimulus response episodes. *Visual Cognition*, (5), 183-216.
- Hommel, B., & Colzato, L.S. (2004). Visual attention and the temporal dynamics of feature integration. *Visual Cognition*, (11), 483-521.
- Hopf, J.-M., Scheoenfeld, M., A., & Jeinze, H.-J. (2005). The temporal flexibility of attentional selection in the visual cortex. *Current Opinion in Neurobiology.*, 15, 183-187.
- Hopfield, J.J. (1982). Neural networks as physical systems with emergent computational abilities. *Proceedings National. Academy Science USA*, 79, 2554-2558.
- Houck, M.R., & Hoffman, J.E. (1986a). Conjunction of colour and form without attention. Evidence from an orientation-contingent colour after effect. *Journal of Experimental Psychology: Human Performance and Perception*, 12 (2), 186-199.
- Houck, M.R., & Hoffman, J.E. (1986b). Conjunction of colour and form without attention. Evidence from an orientation-contingent colour aftereffect. *Journal of Experimental Psychology: Human Performance and Perception*, 12 (2), 186-199.
- Hubel, D.H., & Wiesel, T.N. (1959). Receptive fields of single neurons in the cat's striate cortex. *Journal of Physiology*, 148, 574-591.
- Hubel, D.H., & Wiesel, T.N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, 160, 106-154.
- Hubel, D.H., & Wiesel, T.N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology (London)*, (195), 215-243.
- Hubel, D.H., & Wiesel, T.N. (1977). Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society of London (B)*, (198), 1-59.
- Humphreys, G.W., & Riddoch, M.J. (1992). Interactions between objects and space vision revealed through neuropsychology. In: D.E.M.a.S.K. (eds.). (Ed.) *Attention and Performance XIV* (pp. 143-162.). Hillsdale.: Lawrence Erlbaum.
- Ito, M. (1984). *The Cerebellum and Neural Control*. (New York, NY: Raven Press.
- Jeffreys, J.G.R., Traub, R.D., & Whittington, M.A. (1996). Neuronal networks for induced "40 Hz" rhythms. *Trends in Neuroscience*, 19, 202-207.
- Jiang, Y., Kumar, A., & Vickery, T.J. (2005). Integrating Sequential Arrays in Visual Short-Term Memory. *Experimental Psychology*, 52 (1), 39-46.

- Jiang, Y., Olson, I.R., & Chun, M.M. (2000). Organization of visual short term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26 (3), 683-702.
- Joliot, M., U., R., & R., L. (1994). Human oscillatory brain activity near 40 Hz coexists with cognitive temporal binding. *Proceedings of the National Academy of Science USA*, 91 (pp. 11748-11751).
- Jonides, J., Smith, E.E., Koeppe, R.A., Awh, E., Minoshima, S., & Mintun, M.A. (1993). Spatial working memory in humans as revealed by PET. *Nature*, (363), 623-625.
- Keele, S.W., Cohen, A., Ivry, R., Liotti, M., & Yee, P. (1988). Tests of a Temporal Theory of Attentional Binding. *Journal of Experimental Psychology-Human Perception and Performance*, 14 (3), 444-452.
- Kesner, R.P., & Rolls, E.T. (2001). Role of Long-Term Synaptic Modification in Short-Term Memory. *Hippocampus*, 11, 240-250.
- Kessels, R.P.C., Postma, A., Kappelle, L.J., & De Haan, E.H.F. (2002b). Selective impairments in object-location binding, metric encoding and their integration after ischaemic stroke. *Journal of Clinical and Experimental Neuropsychology*, 24, 115-129.
- Kessler, K., & Kiefer, M. (2005). Disturbing Visual Working Memory: Electrophysiological Evidence for a Role of the Prefrontal Cortex in Recovery from Interference. *Cerebral Cortex*, 15, 1075--1087.
- Kleinfeld, D. (1986). Sequential state generation by model neural networks. *Proceedings of the National Academy of Science, USA.*, 83, 9669-9473.
- Kobatake, E., & Tanaka, K. (1994). Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *Journal of Neurophysiology*, 71, 856-867.
- Koch, C., & Segev, I. (2000). The role of single neurons in information processing. *Nature Neuroscience*, 3, 1171-1177.
- Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology*, (4), 219-227.
- Kohler, S., Moscovitch, M., & Melo, B. (2001). Episodic memory for object location versus episodic memory for object identity: Do they rely on distinct encoding processes? *Memory & Cognition and Emotion*, 29, 849-959.
- Kohonen, T. (1977). *Associative Memory: A System Theoretic Approach*. (Berlin: Springer-Verlag).
- Kohonen, T. (1982a). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43 (pp. 59-69).
- Kohonen, T. (1982b). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43, 59-69.
- Konig, P., & Engel, A.K. (1995). Correlated Firing in Sensory-Motor Systems. *Current Opinion in Neurobiology*, 5 (4), 511-519.
- Kopecz, K., & Schöner, G. (1995). Saccadic motor planning by integrating visual information and expectation on neural dynamic fields. *Biological Cybernetics*, 73, 49-60.
- Kopelman, M.D., Stanhope, N., & Kingley, D. (1997a). Temporal and spatial context memory in patients with focal frontal, temporal lobe, and diencephalic lesions. *Neuropsychologia*, 12, 1533-1545.

- Kopelman, M.D., Stanhope, N., & Kingsley, D. (1997b). Temporal and spatial context memory in patients with focal frontal, temporal lobe, and diencephalic lesions. *Neuropsychologia*, 35 (12), 1533-1545.
- Kosslyn, S.M., Koenig, O., Barrett, A., Cave, C., Tang, J., & Gabrielig, J.D.E. (1989). Evidence for two types of spatial representations: hemispheric specialization for categorical and coordinate relations. *Journal of Experimental Psychology: Perception and Performance*, 15, 723-735.
- Kuffler, S.W. (1953). Discharge patterns and functional organization of mammalian retina. *Journal of Neurophysiology*, 16, 37-68.
- Lamme, V., Super, H., & Spekreijse, H. (1998). Feedforward, horizontal, and feedback processing the visual cortex. *Current Opinion in Neurobiology*, (8), 529-535.
- Lee, D., & Chun, M.M. (2001). What are the units of visual short term memory. Objects or spatial locations? *Perception and Psychophysics*.
- Lee, S.-H., & Blake, R. (1999). Visual form created solely from temporal structure. *Science*, 284, 1165-1168.
- LeVey, S., Hubel, D.H., & Wiesel, T.N. (1975). The pattern of ocular dominance columns in macaque visual cortex revealed by a reduced silver stain. *Journal of Comparative Neurology*, (159), 559-575.
- Levy, N., Horn, D., Melijson, I., & Ruppín, E. (2001). Distributed synchrony in a cell assembly of spiking neurons. *Neural Networks*, 14, 815-824.
- Lisman, J.E., Fellous, J.M., & Wang, X.-J. (1998). A role for NMDA-receptor channels in working memory. *Nature Neuroscience*, 1, 273-275.
- Livingstone, M., & Hubel, D. (1988). Segregation of Form, Color, Movement, and Depth - Anatomy, Physiology, and Perception. *Science*, 240 (4853), 740-749.
- Logothetis, N.K. (1998). Object Vision and Visual Awareness. In: M.C. Raff (Ed.) *Current Opinion in Neurobiology*, 8 (pp. 536-544). London: Current Biology Publications.
- Luck, S.J., & Vogel, E.K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, (390), 279-281.
- Lund, J.S., Lund, R.D., Hendrickson, A.E., Bunt, A.H., & Fuchs, A.F. (1975). The origin of efferent pathways from the primary visual cortex, area 17, of the macaque monkey as shown by the retrograde transport of horseradish peroxidase. *Journal of Comparative Neurology*, (164), 287-304.
- Lund, J.S., & Wu, C.Q. (1997). Local circuit neurons of macaque monkey striate cortex. *Journal of Computational Neurology*, (384), 109-126.
- Maass, W., & Bishop, C. (1998). Pulsed Neural Networks. (Cambridge, Mass.: MIT Press.
- Maass, W., & Zador, A.M. (1998). Dynamic Stochastic Synapses as Computational Units. In: e. W. Maass and C. Bishop (Ed.) *Pulsed Neural Networks*. (pp. 321-336). Cambridge, MA.: MIT Press.
- MacAvoy, M.G., Gottlieb, J.P., & Bruce, C.J. (1991). *Cerebral Cortex*, 1, 95-102.
- Macknik, S.L., & Livingstone, M.S. (1998). Neuronal correlates of visibility and invisibility in the primate visual system. *Nature Neuroscience*, 1 (2), 144-149.
- Mainen, Z.F., & Sejnowski, T.J. (1995). Reliability of spike timing in neocortical neurons. *Science*, 268, 1503-1506.

- Marder, E. (1998b). From biophysics to models of network function. *Annual Review of Neuroscience*, 21, 25-45.
- Markram, H., Lubke, J., Frotscher, M., & Sakmann, B. (1997). Regulation of synaptic efficacy by coincidence of postsynaptic APOs and EPSPs. *Science*, (275), 213-215.
- Marr, D. (1982). *Vision*. (San Francisco: Freeman Publishers.
- Maunsell, J.H., & Gibson, J.R. (1992). Visual response latencies in striate cortex of the macaque monkey. *Journal of Neurophysiology*, (68), 1332-1344.
- Maunsell, J.H., & Van Essen, D.C. (1983). The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *Journal of Neuroscience*, (3), 2563-2586.
- Maunsell, J.H.R., & Newsome, W.T. (1987). Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience*, (10), 363-401.
- McClelland, J.L., McNaughton, B.L., & O'Reilly, R.C. (1995). Why are there complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, 102, 419-457.
- McCulloch, W.S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5, 115-133.
- McLelland, J.L., & Rumelhart, D.E. (1988). *Explorations in Parallel Distributed Processing*. (Cambridge, Mass. USA.: MIT Press.
- Melcher, D., Papathomas, T.V., & Vidnyanszky, Z. (2005). Implicit attentional selection of bound visual features. *Neuron*, ***
- Merigan, W.H., & Pham, H.A. (1998). V4 lesions in macaques affect both single and multiple viewpoint shape discriminations. *Visual Neuroscience*, (15), 359-369.
- Miall, D.S. (1989). Beyond the schema given: Affective comprehension of literary narratives. *Cognition and Emotion*, 3 (1), 55-78.
- Miikkulainen, R., Bednar, J., Yoonsuck, C., & Sirosh, J. (2005). *Computational Maps in the Visual Cortex*. (N.Y.: Springer.
- Miller, E., & Desimone, R. (1994). Parallel neuronal mechanisms for short-term memory. *Science*, (263), 520-522.
- Miller, E., Erickson, C.A., & Desimone, R. (1996). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *Journal of Neuroscience*, 16 (16), 5154-5167.
- Miller, E.K., & Cohen, J.D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24, 167-202.
- Miller, E.K., Gochin, P.M., & Gross, C.G. (1991a). Habituation-like decrease in the responses of neurons in inferior temporal cortex of the macaque. *Visual Neuroscience*, (7), 357-362.
- Miller, G.A., Gallanter, E., & Pribram, K.H. (1960). *Plans and the Structure of Behaviour*. (New York: Henry Holt.
- Milner, A.D., & Goodale, M.A. (1995). *The Visual Brain in Action*. (Oxford: OUP.
- Milner, B., Johnsrude, I., & Crane, J. (1997). Right medial temporal-lobe contribution to object location memory. *Philosophical Transactions/Royal Society of London. Biological Sciences*, 352, 1469-1474.
- Milner, P.M. (1974). A model for visual shape recognition. *Psychology Review*, 81 (6), 521-535.

- Mishkin, M., Ungerleider, L.G., & Macko, K.A. (1983). Object vision and spatial vision: two cortical pathways. *Trends in Neuroscience*, (6), 414-417.
- Miyashita, Y. (1988a). Neuronal correlate of visual associative long-term memory in the primate temporal cortex. *Nature*, 335, 817.
- Miyashita, Y. (1988b). Neuronal correlate of visual long term memory in the primate temporal cortex. *Nature*, (335), 68-70.
- Moran, J., & Desimone, R. (1985). *Science*, (229), 782-784.
- Moscovitch, M., Kapur, S., Kohler, S., & Houle, S. (1995). Distinct neural correlates of visual long-term memory for spatial location and object identity: A position emission tomography study in humans. *Proceedings of the National Academy of Sciences USA*, 92, 3721-3725.
- Mueller, H.J., & Elliott, M.A. (2001). Neural binding of space and time: an introduction. *Visual Cognition*, 8 (3/4/5), 273-285.
- Mumford, D. (1992). On the computational architecture of the neocortex.II. The role of cortico-cortical loops. *Biological Cybernetics*, (66), 241-251.
- Munk, M.H.J., Roelfsema, P.R., Konig, P., Engel, A.K., & Singer, W. (1996). Role of reticular activation in the modulation of intracortical synchronisation. *Science*, 272, 271-273.
- Murata, A., Gallese, V., Luppino, G., Kaseda, M., & Sakata, H. (2000). Selectivity for the shape, size, and orientation of objects for grasping in neurons of monkey parietal area AIP. *Journal of Neurophysiology*, 83 (5), 2580-2601.
- Murthy, V.N., & Fetz, E.E. (1992). Coherent 25-35 Hz oscillations in the sensorimotor cortex of the awake behaving monkey. *National Academy of Sciences USA*, 89 (pp. 5670-5674).
- Nairne, J.S., & Dutta, A. (1992). Spatial and temporal uncertainty in long-term memory. *Journal of Memory and Language*, 31, 396-407.
- Naveh-Benjamin, M. (1987). Coding of spatial location information: An automatic process? *Journal of Experimental Psychology: Learning, Memory and Cognition*, 13, 595-605.
- Neuenschwander, S., & Singer, W. (1996). Long-range synchronization of oscillatory light responses in the cat retina and lateral geniculate nucleus. *Nature*, 379, 728-733.
- Niebur, E., Koch, C., & Rosin, C. (1993). An oscillation based model for the neuronal basis of attention. *Vision Research*, (33), 2789-2802.
- Nirenberg, S., & Latham, P.E. (1998). Population coding in the retina. *Current Opinion in Neurobiology*, 8 (pp. 488-493).
- Nissen, M.J. (1985). Accessing features and objects. In: O.S.M. Marin (Ed.) *Attention and performance XI* (Hillsdale, N.J.: Lawrence Erlbaum Associates).
- Nobre, A.C., Coull, J.T., Walsh, V., & Frith, C.D. (2003a). Brain activations during visual search: contributions of search efficiency and feature binding. *NeuroImage*, (18), 91-103.
- Nobre, A.C., Coull, J.T., Walsh, V., & Frith, C.D. (2003b). Brain Activations during Visual Search: Contributions of Search Efficiency versus Feature Binding. *NeuroImage* 18, 18 (p. 91-103).
- Nowak, L.G., & Bullier, J. (1997). The timing of information transfer in the visual system. In: *Extrastriate Cortex. Cerebral Cortex*, 12 (pp. 205-241). New York: Plenum press.

- Oja, E. (1982). A simplified neuron model as a principle component analyzer. *Journal of Mathematical Biology*, 15, 267-273.
- Optican, L.M., & Richmond, B.J. (1987). Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. *Journal of Neurophysiology*, (57), 162-178.
- Palanca, B.J.A., & DeAngelis, G.C. (2005). Does neuronal synchrony underlie visual feature grouping? *Neuron*, ***
- Pandya, D.N., & Yeterian, E.H. (1995). Architecture and connections of cortical association areas. *Cerebral Cortex*, (4), 3-61.
- Panzeri, S., & Schultz, S.R. (2001). A unified approach to the study of temporal, correlational and rate codings.
- Panzeri, S., Schultz, S.R., Treves, A., & Rolls, E.T. (1999a). Correlations and the encoding of information in the nervous system. *Proceedings of the Royal Society London B*, 266 (pp. 1001-1012.).
- Parkin, A.J., Walter, B.M., & Hunkin, N.M. (1995). Relationships between normal aging, frontal lobe function, and memory for temporal and spatial information. *Neuropsychology*, 9, 304-312.
- Patterson, D. (1996a). Artificial Neural Networks. (Singapore: Prentice Hall.
- Patterson, D.W. (1996b). Artificial neural Networks. (Singapore.: Prentice Hall.
- Payne, B.R., Lomber, S.G., Villa, A.E., & Bullier, J. (1996). Reversible deactivation of cerebral network components. *Trends in Neuroscience*, (19), 535-542.
- Perrett, D., Mistlin, A.J., & Chitty, A.J. (1987). Visual neurons responsive to faces. *Trends in Neurosciences*, 9 (10), 358--363.
- Phaf, R.H., & Wolters, G. (1997). A constructivist and a connectionist view on conscious and non-conscious processes. *Philosophical Psychology*, (10), 287-307.
- Pollen, D.A. (1999). On the neural correlates of visual perception. *Cerebral Cortex*, 9 (1), 4-19.
- Postma, A., Kessels, R.P.C., & Van Asselen, M. (2003). The neuropsychology of object-location memory. In: G.L. Allen (Ed.) *Human spatial memory: Remembering where* (pp. 143-160). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Potter, M.C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology*, 2, 509-522.
- Pouille, F., & Scanziani, M. (2004). Routing of spike series by dynamic circuits in the hippocampus. *Nature*, 429, 717-723.
- Prinzmetal, W. (1995). Visual Feature Integration in a World of Objects. *Current Directions in Psychological Science*, 4 (3), 90-94.
- Prinzmetal, W., Henderson, D., & Ivry, R. (1995). Loosening the Constraints On Illusory Conjunctions - Assessing the Roles of Exposure Duration and Attention. *Journal of Experimental Psychology-Human Perception and Performance*, 21 (6), 1362-1375.
- Raffone, A., & van Leeuwen, C. (2001). Chaos and neural coding: is the binding problem a pseudo-problem? *Behavioural and Brain Sciences*, 24 (4)
- Raffone, A., & van Leeuwen, C. (2003). Dynamic synchronization and chaos in an associative neural network with multiple active memories. *Chaos*, 13 (3), 1090-1104.
- Raffone, A., & Wolters, G. (2001). A cortical mechanism for binding in visual working memory. *Journal of Cognitive Neuroscience*, 13, 766-785.

- Rainer, G., Asaad, W.F., & Miller, E.K. (1998). Memory fields of neurons in the primate prefrontal cortex. *Proceedings of the National Academy of Science USA*, 95 (pp. 15008-15013.).
- Rao, R.P.N., & Ballard, D.H. (1997). Dynamic Model of Visual Recognition Predicts Neural Response Properties in the Visual Cortex. *Neural Computation*, 9, 721-763.
- Rao, S.C., Rainer, G., & Miller, E.K. (1997a). integration of what and where in the primate prefrontal cortex. *Science*, 276 (pp. 821-824).
- Rao, S.C., Rainer, G., & Miller, E.K. (1997b). Integration of what and where in the primate prefrontal cortex. *Science*, 276, 821-824.
- Rao, S.G., Williams, G.V., & Goldman-Rakic, P. (1999). *Journal of Neurophysiology*, (81), 1903.
- Rauch, A., La Camera, G., Luescher, H.-R., Senn, W., & Fusi, S. (2003). Neocortical pyramidal cells respond as integrate-and-fire neurons to in vivo-like input currents. *Journal of Neurophysiology*, 90, 1598-1612.
- Raymond, J.E., Shapiro, K.L., & Arnell, K.M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, 18, 849-860.
- Redish, A.D., & Touretzky, D.S. (1996). Modeling interactions of the rat's place and head direction systems. In: M.E. Hasselmo (Ed.) *Advances in neural information processing systems*, 8 (Cambridge, Massachusetts: MIT Press.
- Rees, G., Frackowiak, R.S.J., & Frith, C. (1997). Two modulatory effects of attention that mediate object categorisation in human cortex. *Science*, 275 (835-838.)
- Renart, A., Moreno, R., de la Rocha, J., Parga, N., & Rolls, E.T. (2001). A model of the IT-PF network in object working memory which includes balanced persistent activity and tuned inhibition. *Neurocomputing*, (38-40), 1525-1531.
- Renart, A., Parga, N., & Rolls, E.T. (1999). The relation between perception and short term memory modelled by interacting recurrent network modules.
- Renart, A., Parga, N., & Rolls, E.T. (2000). A recurrent model of the interaction between PF and IT cortex in delay memory tasks. *NIPS99* (pp. 171-177). Cambridge, MA: MIT Press.
- Reynolds, J.H., & Desimone, R. (1999). The role of neural mechanisms of attention in solving the binding problem. *Neuron*, 24 (1), 19-29.
- Richmond, B. (2001). Information Coding. *Science*, 294, 2493-2494.
- Riehle, A., Grun, S., Diesmann, M., & Aertsen, A. (1997). Spike synchronization and rate modulation differentially involved in motor cortical function. *Science*, 278, 1950-1953.
- Rieke, F., Warland, D., de Ruyter van Steveninck, R., & Bialek, W. (1997). Exploring the Neural Code. *Computational Neurosciences* (Cambridge, Mass.: MIT Press.
- Ritz, R., & Sejnowski, T.J. (2000). Synchronous oscillatory activity in sensory systems: new vistas on mechanisms. *Current Opinion in Neurobiology*, 7 (536-546.)
- Robertson, L., Treisman, A.M., Friedman-Hill, S., & Grabowecky, M. (1997). The Interaction of Spatial and Object Pathways: Evidence from Balint's Syndrome. *Journal of Cognitive Neuroscience*, 9 (3), 295 - 317.

- Robertson, L.C. (2003). Binding, spatial attention and perceptual awareness. *Nature Neuroscience*, 4 (Feb, 2003), 93-102.
- Robson, C. (1995). Real World Research. (Cambridge, Mass. USA: Blackwell Inc.
- Roelfsema, P.R., König, P., Engel, A.K., & Singer, W. (1996). The role of neuronal synchronization in response selection: A biologically plausible theory of structured representation in the visual cortex. *Journal of Cognitive Neuroscience*, 8, 603-625.
- Rolls, E.T. (1989b). Functions of neuronal networks in the hippocampus and neocortex in memory. In: J.B.a.W.B. (eds.) (Ed.) *Neural Models of Plasticity: Experimental and Theoretical Approaches*. (San Diego, USA.: Academic Press.
- Rolls, E.T. (2000). Memory systems in the brain. *Annual Review of Psychology*, 51, 599-630..
- Rolls, E.T. (2001a). Functions of the primate temporal lobe cortical visual areas in invariant visual object and face recognition: computational mechanisms. In: C.Musio (Ed.) *Vision: The Approach of Biophysics and Neurosciences* (pp. 396-408.). London: World Scientific.
- Rolls, E.T. (2001b). The rules of formation of the olfactory representations found in the orbitofrontal cortex olfactory areas in primates. *Chemical Senses*, 26 (595-604)
- Rolls, E.T., & Deco, G. (2002). Computational Neuroscience of Vision. (Oxford: OUP.
- Rolls, E.T., & Tovee, M.J. (1994). Processing speed in the cerebral cortex and the neurophysiology of visual masking. *Proceedings of the Royal Society London. B.*, 257, 9-15.
- Rolls, E.T., & Tovee, M.J. (1995). The Responses of Single Neurons in the Temporal Visual Cortical Areas of the Macaque When More Than One Stimulus Is Present in the Receptive-Field. *Experimental Brain Research*, 103 (3), 409-420.
- Rolls, E.T., & Treves, A. (1998). Neural networks and Brain Function. (Oxford: OUP.
- Rosenblatt, F. (1958). The Perceptron: a probabilistic model for information storage & organization in the brain. *Psychological Review*, 65, 386-408.
- Rosenblatt, F. (1961). Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms. (Washington, DC.: Spartan Books.
- Roskies, A.L. (1999). The binding problem - Introduction. *Neuron*, 24 (1), 7-9.
- Ross, W.D., Grossberg, S., & Mingolla, E. (2000). Visual cortical mechanisms of perceptual grouping: interacting layers, networks, columns and maps. *Neural Networks*, (13), 571-588.
- Saiki, J. (2003). Spatiotemporal characteristics of dynamic feature binding in visual working memory. *Vision Research*, 43, 2107-2123.
- Sakai, K., & Passingham, R.E. (2004). Prefrontal selection and medial temporal lobe reactivation in retrieval of short-term verbal information. *Cerebral Cortex*, 4, 914--921.
- Salin, P., & Bullier, J. (1995). Corticocortical connections in the visual system: structure and function. *Physiological Review*, (75), 107-154.
- Sato, T. (1989). Interactions of visual stimuli in the receptive fields of inferior temporal neurons in macaque. *Experimental Brain Research*, (77), 23-30.

- Schacter, D.L. (1989). Memory. In: M. Posner (Ed.) *Foundations of Cognitive Science*. (Cambridge, Mass.: Bradford Books.
- Schillen, T.B., & König, P. (1994). Binding by temporal structure in multiple feature domains of an oscillatory neuronal network. *Biological Cybernetics*, 70, 397-405.
- Schiller, P.H. (1993). The effects of V4 and MT lesions on visual performance in the rhesus monkey. *Visual Neuroscience*, (10), 717-746.
- Sereno, A.B., & Maunsell, J.H.R. (1998). Shape selectivities in primate lateral intraparietal cortex. *Nature*, 395, 500-503.
- Shadlen, M.N., & Movshon, a.J.A. (1999a). Synchrony unbound: a critical evaluation of the temporal binding hypothesis. *Neuron*, 24 (pp. 67-77).
- Shadlen, M.N., & Movshon, J.A. (1999b). Synchrony unbound: a critical evaluation of the temporal binding hypothesis. *Neuron*, 24, 67-77.
- Shadlen, M.N., & Newsome, W.T. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *Journal of Neuroscience*, 18 (10), 3870-3896.
- Shafritz, K.M., Gore, J.C., & Marois, R. (2002). The role of the parietal cortex in visual feature binding. *PNAS*, 99 (16), 10917-10922.
- Shalden, M.N., & Newsome, W.T. (1994). *Current Opinion in Neurobiology*, (4), 569.
- Shapley, R.M., & Perry, V.H. (1986). Cat and monkey retinal ganglion cells and their visual functional roles. *Trends in Neurosciences*, 9, 229-235.
- Shepherd, G.M. (1979). *The Synaptic Organization of the Brain*. (Oxford.: OUP.
- Silito, A.M., Jones, H.E., Gerstein, G.L., & West, D.C. (1994). Feature-linked synchronisation of thalamic relay cell firing induced by feedback from the visual cortex. *Nature*, (369), 479-482.
- Silito, A.M., & Murphy, P.C. (1986). *Neurotransmitters and Cortical Function*. (pp. 167-186.). New York: Plenum Press.
- Singer, W. (1994). Putative functions of temporal correlations in neocortical processing. In: Davis (Ed.) *Large-scale neuronal theory of the brain*. (Boston Mass: MIT Press.
- Singer, W. (1996). Neuronal synchronization: a solution to the binding problem. In: R.R.L.a.P.S.C. (Eds.) (Ed.) *The Mind-Brain Continuum; Sensory Processes*. (pp. 101-130). Cambridge, Mass.: MIT Press.
- Singer, W. (1999). Neuronal Synchrony: A Versatile Code for the Definition of Relations? *Neuron*, 24, 49-65.
- Singer, W., Engel, A., Kreiter, A., Munk, M., Neuenschwander, S., & Roelfsema, P. (1997). Neuronal assemblies: necessity, signature and detectability. *Trends in Cognitive Science*, 1 (7), 252-261.
- Singer, W., & Gray, C.M. (1995). Visual Feature Integration and the Temporal Correlation Hypothesis. *Annual Review of Neuroscience*, 18, 555-586.
- Skaggs, W.E., Knierim, J.J., Kudrimoti, H.S., & McNaughton, B.L. (1995). A model of the neural basis of the rat's sense of direction. In: T.K. Leen (Ed.) *Advances in neural information processing systems*, 7 (pp. 173-180.). Cambridge, Massachusetts: MIT Press.
- Smirnakis, S.M., Berry, M.J., Warland, D.K., Bialek, W., & Meister, M. (1997). Adaptation of retinal processing to image contrast and spatial scale. *Nature*, 386 (69-73.)
- Smith, E.E., & Jonides, J. (1997). Working memory: A view from neuroimaging. *Cognitive Psychology Review*, 33, 5-42.

- Smith, M.L., & Milner, B. (1981). The role of the right hippocampus in the recall of spatial location. *Neuropsychologia*, 19, 781-793.
- Smith, M.L., & Milner, B. (1989). Right hippocampal impairment in the recall of spatial location: encoding deficit or rapid forgetting? *Neuropsychologia*, 27, 71-81.
- Sohal, V. (1995). A Binding Game. *The Harvard BRAIN: Harvard's Undergraduate Neuroscience Magazine*, II (I)
- Sougne, J. (2000). Short term memory in a network of spiking neurons.
- Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, (74), 1-29.
- Sporns, O., Gally, J.A., Reeke, G.N., & Edelman, G.M. (1989). Reentrant signaling among simulated neuronal groups leads to coherency in their oscillatory activity. *Proceedings of the National Academy of Sciences*, 86, 7265-7269.
- Sporns, O., Tononi, G., & Edelman, G.M. (1991). Modeling perceptual grouping and figure-ground segregation by means of active reentrant connections. *Proceedings of the National Academy of Sciences U.S.A.*, 88 (pp. 129-133).
- Stefurak, D.I., & Boynton, R.M. (1986). Independence of memory for categorically different colours and shapes. *Perception and Psychophysics*, 39 (3), 164-174.
- Stepankova, K., Fenton, A.A., Pastalkova, M.K., & Bohbot, V.D. (2004). Object-location memory impairment in patients with thermal lesions to the right or left hippocampus. *Neuropsychologia*, 42, 1017-1028.
- Stopfer, M., Bhagavan, S., Smith, B.H., & Laurent, G. (1997). Impaired odour discrimination on desynchronization of odour-encoding neural assemblies. *Nature*, 390, 70-74.
- Stringer, S.M., Rolls, E.T., & Trappenberg, T.P. (2004). Self-organising continuous attractor networks with multiple activity packets, and the representation of space. *Neural Networks*, 17, 5-27.
- Tanaka, K. (1997). Mechanisms of visual object recognition: monkey and human studies. *Curr. Opin. in Neurobiol.*, 7, 523-529.
- Thorpe, S.J., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, 381, 520-522.
- Tomita, H., Ohbayashi, M., Nakahara, K., Hasegawa, I., & Miyashita, Y. (1999). Top-down signal from prefrontal cortex in executive control of memory retrieval. *Nature*, 401, 699--703.
- Tononi, G., Sporns, O., & Edelman, G.M. (1992b). Reentry and the problem of integrating multiple cortical areas: simulation of dynamic integration in the visual system. *Cerebral Cortex*, 2 (2), 310-335.
- Tootell, R.B.H., Silverman, M.S., Devalois, R.L., & Jacobs, G.H. (1983). Functional-Organization of the 2nd Cortical Visual Area in Primates. *Science*, 220 (4598), 737-739.
- Toth, L.J., Kim, D.S., Rao, S.C., & Sur, M. (1997). Integration of local inputs in visual cortex. *Cerebral Cortex*, (7), 703-710.
- Trappenberg, T.P. (2002a). *Fundamental of Computational Neuroscience*. (Oxford: OUP.
- Trappenberg, T.P. (2002b). *Fundamentals of Computational Neuroscience*. (Oxford: OUP.

- Trappenberg, T.P. (2003). Why is our working memory so large? *Neural Information Processing-Letters and Reviews*, 1 (3), 97-101.
- Treisman, A. (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 194-214.
- Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology*, 6 (2), 171-178.
- Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12 (1), 97-136.
- Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, (14), 107-141.
- Treisman, A., Sykes, M., & Gelade, G. (1977). Selective attention and stimulus integration. In: S.D. (Eds.) (Ed.) *Attention and Performance VI* (Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Treves, A., & Rolls, E.T. (1991a). What determines the capacity of autoassociator networks in the brain? *Network*, 2, 371-397.
- Treves, A., & Rolls, E.T. (1991b). What determines the capacity of autoassociator networks in the brain? *Network*, 2, 371-397.
- Treves, A., & Rolls, E.T. (1994). A computational analysis of the role of the hippocampus in memory. *Hippocampus*, 4, 374-391.
- Tsodyks, M.V., & Markram, H. (1996). Plasticity of Neocortical Synapses Enables Transitions Between Rate & Temporal Coding. *Lecture Notes in Comp. Sci*, 1112 (pp. 445-450). Berlin: Springer.
- Tuckwell, H. (1988). Introduction to theoretical neurobiology. (Cambridge: CUP.
- Ungerleider, L.G. (1995). Functional brain imaging studies of cortical mechanisms for memory. *Science*, 270, 769-775.
- Ungerleider, L.G., Courtney, S.M., & Haxby, J.V. (1998). A neural system for human visual working memory. *Proceedings of the National Academy of Science USA*, 1998, 95 (3), 883-890.
- Ungerleider, L.G., Gaffan, D., & Pelak, V.S. (1989). Projections from inferior temporal cortex to prefrontal cortex via the uncinate fascicle in rhesus monkeys. *Experimental Brain Research*, (76), 473-484.
- Ungerleider, L.G., & Mishkin, M. (1982). Two cortical visual systems. In: R.J.W. Mansfield (Ed.) *Analysis of Visual Behavior*. (pp. 549-586). Cambridge, Mass: The MIT Press.
- Usher, M., & Cohen, J.D. (1999). Short-term memory and selection processes in a frontal-lobe model. In: A.O. (Eds.) (Ed.) *Connectionist Models in Cognitive Neuroscience* (pp. 78-91): Springer Verlag.
- Usher, M., & Donnelly, N. (1998). Visual synchrony affects binding and segmentation in perception. *Nature*, 394, 179-182.
- Usher, M., & Niebur, E. (1996). Modelling the temporal dynamics of IT neurons in visual search: a mechanism for top-down selective visual attention. *Journal of Cognitive Neuroscience*, (8), 311-327.
- Van Asselen, M., Postma, A., & van der Lubbe, R.H.J. (2002). The influence of aging on spatial and temporal order memory. *6th Dutch Endo-Neuro Meeting* (Doorwerth, The Netherlands.
- van der Velde, F. (1997). On the use of computational modelling behaviour. *Network: Computational Neural Systems*, 8, 1-32.

- van der Velde, F., & de Kamps, M. (2003). A model of visual working memory in PFC. *Neurocomputing*, 52-54, 419-424.
- van der Voort van der Kleij, G.T., de Kamps, M., & van der Velde, F. (2004). Increasing number of objects impairs binding in visual working memory. *Neurocomputing*, 58-60, 599-605.
- van Essen, D. (1985). Functional architecture of primate visual cortex. In: E.G. Jones (Ed.) *Visual Cortex*, 3 (pp. 259-329): Plenum Press.
- van Leeuwen, C., Buffart, H., & van de Vegt, J. (1988). Sequence influence on the organization of meaningless serial stimuli: economy after all. *Journal of Experimental Psychology: Human Perception and Performance*, 14 (481-502.)
- van Leeuwen, C., & Raffone, A. (2001). Coupled non-linear maps as models of perceptual pattern and memory trace dynamics. *Cognitive Processing*, 2, 67-111.
- van Leeuwen, C., Steyvers, M., & Nooter, M. (1997). Stability and intermittency in large-scale coupled oscillator models for perceptual segmentation. *Journal of Mathematical Psychology*, 41, 1-26.
- van Leeuwen, C., Verver, S., & Brinkers, M. (2000). Visual illusions, solid/outline in variance, and non-stationary activity patterns. *Connection Science*, 12, 279-297.
- van Rullen, R., & Thorpe, S.J. (2001). Rate coding versus temporal order coding: what the retinal ganglion cells tell the visual cortex. *Neural Computation*, 13, 1255-1283.
- Vecera, S.P., & Farah, M.J. (1997). Is visual image segmentation a bottom-up or an interactive process? *Perception & Psychophysics*, 59, 1280-1296., 59, 1280-1296.
- Vogel, E.K., Woodman, G.F., & Luck, S.J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 27 (1), 92-114.
- Von der Malsburg, C. (1973). Self-organisation of orientation sensitive cells in the striate cortex. *Kybernetik*, 14, 85-100.
- Von der Malsburg, C. (1981). The Correlation Theory of Brain Function. *Biophysical Chemistry, MPI*, 81-82.
- von der Malsburg, C. (1983). How are nervous structures organized? In: A.J. Mandell (Ed.) *In Synergetics of the Brain, Proceeding of the International Symposium on Synergetics*. (pp. 238-259): Springer.
- von der Malsburg, C. (1995). Dynamic link architecture. In: M. Arbib (Ed.) *The Handbook of Brain Theory and Neural Networks*, (pp. 329-331.). Cambridge, Mass: MIT Press.
- Von der Malsburg, C., & Schneider, W. (1986). A Neural Cocktail-Party Processor. *Biological Cybernetics*, 54 (1), 29-40.
- Walker, P., & Cuthbert, L. (1998). Remembering visual feature conjunctions: visual memory for shape colour associations is object-based. *Visual Cognition*, 5 (4), 409-455.
- Walsh, V. (1999). How does the cortex construct color? *Proceedings of the National Academy of Science USA*, 96(24) (pp. 13594-13596).
- Wang, X. (2001). Synaptic reverberation underlying mnemonic persistent activity. *Trends in Neuroscience*, 24, 455-463.

- Werner, G., & Mountcastle, V.B. (1965). Neural activity in mechanoreceptive cutaneous afferents: Stimulus-response relations, Weber functions, and information transmission. *Journal of Neurophysiology*, 28, 359-397.
- Wersing, H., & Ritter, H. (1999). Feature binding and relaxation labeling with the competitive layer model. *ESANN'1999* (pp. 295-300). Bruges: D-Facto.
- Wheeler, M.E., & Treisman, A. (2002). Binding in short term visual memory. *Journal of Experimental Psychology: General.*, (131), 48-64.
- Whittington, M.A., Traub, R.D., & Jefferys, J.G.R. (1995). Synchronized oscillations in interneuron networks driven by metabotropic glutamate receptor activation. *Nature*, 373, 612-615.
- Williams, G.V., & Goldman-Rakic, P.S. (1995). *Nature (London)*, 376, 572-575.
- Wilson, F.A., O'Sclaidhe, S.P., & Goldman-Rakic, P.S. (1993). Dissociation of spatial and object processing domains in primate prefrontal cortex. *Science*, (260), 1955-1958.
- Wilson, H., & Cowan, J. (1972). Excitatory and inhibitory interaction in localised populations of model neurons. *Biological Cybernetics*, (12), 1-24.
- Wilson, H., & Cowan, J. (1973). A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik*, (13), 55-80.
- Wolfe, J.M., Cave, K.R., & Franzel, S.L. (1989). Guided Search - an Alternative to the Feature Integration Model For Visual-Search. *Journal of Experimental Psychology-Human Perception and Performance*, 15 (3), 419-433.
- Woodman, G.F., & Vogel, E.K. (2005). Fractionating Working Memory: Consolidation and Maintenance Are Independent Processes. *Psychological Science*, 16 (2), 106-113.
- Wu, Y. (2002a). Encoding colour and shape from different parts of an object in visual short term memory. *Perception and Psychophysics*, (64), 1260-1280.
- Yoshioka, T., Blasdel, G.G., Levitt, J.B., & Lund, J.S. (1996). relation between patterns of intrinsic lateral connectivity, ocular dominance, and cytochrome oxidase-reactive regions in macaque monkey striate cortex. *Cerebral Cortex*, (6), 297-310.
- Zeki, E. (1997). The colour and motion systems as guides to conscious visual perception. *Cerebral Cortex*, (12), 777-809.
- Zeki, S. (1973). Colour coding in rhesus monkey prestriate cortex. *Brain Res.*, 53, 422-427.
- Zeki, S. (1978). Functional specialisation in the visual cortex of the rhesus monkey. *Nature*, 274, 423-428.
- Zeki, S. (1993a). The visual association cortex. *Curr. Opinion Neurobiol.*, 3, 155-159.
- Zeki, S. (1993b). The visual association cortex. *Current Opinion in Neurobiology*, 3, 155-159.
- Zhang, K.-C. (1995). Dynamical representation of spatial orientation: a model of head direction cells. , pp . . *Proceedings of the 2nd Joint Symposium on Neural Computation* (pp. 24-25.). University of California, San Diego, and California Institute of Technology.
- Zihl, J., Von Cramon, D., Mai, N., & Schmidt, C.H. (1991). Disturbance of movement vision after bilateral posterior brain damage. *Brain*, (114), 2235-2252.
- Zipser, K., Lamme, V., & Schiller, P.H. (1996). Contextual modulation in primary visual cortex. *Journal of Neuroscience*, (16), 7376-7389.