

**CORTICAL SELF-ORGANIZATION DURING OBJECT DISCRIMINATION**

by

Edward Rzempoluck

BA., Simon Fraser University 1991

MA, Simon Fraser University 1993

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF  
THE REQUIREMENTS FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY  
in the Department  
of  
PSYCHOLOGY

© Edward Rzempoluck 1995

SIMON FRASER UNIVERSITY

October 1995

All rights reserved. This work may not be reproduced in whole or in part, by photocopy or other means, without permission of the author.

## Approval

NAME: Edward Rzempoluck  
DEGREE: Doctor of Philosophy (Psychology)  
TITLE OF THESIS: Cortical Self-Organization During Object Discrimination

### EXAMINING COMMITTEE:

Chair: Jean Koepke  
Associate Professor of Psychology

---

Harold Weinberg  
Senior Supervisor  
Professor of Kinesiology

---

Barry L. Beyerstein  
Associate Professor of Psychology

---

Thomas L. Richardson  
Associate Professor of Kinesiology

---

Christopher Davis  
Internal External Examiner  
Associate Professor of Psychology

---

J. A. Scott Kelso  
External Examiner  
Professor of Psychology  
Professor of Biology  
Florida Atlantic University

DATE APPROVED: 13 Dec 1995

## PARTIAL COPYRIGHT LICENSE

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

### **Title of Thesis/Project/Extended Essay**

Cortical Self-Organization During Object Discrimination

---

---

---

---

**Author:** \_\_\_\_\_

(signature)

Edward Joseph Rzemoluck

(name)

Dec. 15 1995

(date)

## **Abstract**

Camouflaged object discrimination was studied with two objectives: to develop a model of object discrimination, and to develop a neuroelectric index of camouflaged object discrimination. Scalp potentials were recorded for 8 second intervals in two conditions. In the experimental condition subjects, 3 female and 3 male right-handed university students with no known neurological disorders, ages 23 to 47, viewed images depicting target objects embedded in a camouflaging background that delayed target discrimination. Subjects signaled discrimination by blinking. In a control condition subjects were instructed to blink at will. Linear inter-channel association within theta band EEG was estimated using cross-correlation and coherence analyses; general association, using mutual information analysis. Correlation and coherence increased over the 1 second interval preceding discrimination, between multiple regions with a larger increase between more widely separated areas, and an inverse relationship between association and separation. Associations initially included occipital and left temporal regions, developing into a bilateral pattern involving left and right frontal and temporal areas, and evolving into an organization, immediately prior to discrimination, that included bilateral occipital, temporal, central, frontal and prefrontal regions. Mutual information showed a similar pattern, indicating a strong linear component to interregional association. The Discrimination Index was defined as the ratio of cross-correlation mean to variance. The index increased by 78% immediately prior to discrimination in the picture condition, and decreased by 11% over the same interval in the control condition. A model is presented, according to which a unitary percept is the emergent result of a process of self-organization within a network of interregional signaling in which information interchange between multiple and wide-spread cortical regions in successive iterations accomplishes a recursive series of transformations of the original retinal representation, through which elementary features are bound into a population of successively more complex ensembles, which are in turn selected according to the goodness of match with memory templates. A successful match terminates the iterations, accomplishing the target-background discrimination. The model suggests a dynamical system within which the pattern of interareal signaling, driven by energy relaxation, self-organizes in order to coordinate the processing resources within multiple and widespread cortical regions.

**Key Words:** visual perception; camouflage; scalp potentials; self-organization model

**The Secret Sits**

**We dance round in a ring and suppose,  
But the Secret sits in the middle and knows.**

**Robert Frost, 1942**

## **Acknowledgments**

I thank my senior supervisor, Dr. Harold Weinberg for his initial suggestion that this question might be an interesting one to study, for his ongoing support over the course of this work, and for providing an environment in which the exchange of ideas, amongst all members of the Brain-Behaviour Laboratory was a natural and fruitful process. I thank all members on my committee for the wealth of ideas and points of view that were presented to me, as well as, importantly, for their patience. I thank also the members of the Brain-Behaviour Laboratory for many helpful suggestions, and for their participation as subjects in this study. Finally I thank my sweetheart who put up with seeing the back of my head for weeks and months as I sat at the computer.

This work was supported in large part by fellowships from the Natural Sciences and Engineering Research Council, and from the Department of Psychology, as well as by a Doctoral Stipend from Simon Fraser University. I am grateful for these sources of support, without which this study would not have been possible.

Simulnet is a trademark of E. J. Rzempoluck.

## Table of Contents

Approval .....	ii
Abstract.....	iii
Acknowledgments .....	v
List of Tables .....	vii
List of Figures .....	viii
I Introduction .....	1
1 Overview.....	2
2 Psychoneural Identity: Can Mental Processes be Studied? .....	6
3 Neurophysiological Models of Cortical Integration.....	11
4 Studies of Visual Perception.....	20
5 A Model of Neural Processes in Object Discrimination.....	34
5.1 Predictions.....	42
II Linear Analyses.....	47
6 Cross-correlation analysis.....	48
7 Topographical Distribution of Correlations.....	79
8 Topographical Distribution of Net Correlations .....	95
9 Mutual Information Analysis.....	101
10 Coherence Analysis .....	110
11 Topographic Distribution of Coherence Differences .....	118
12 The Discrimination Index .....	122
III Nonlinear Analyses .....	126
13 Neural Network Analysis .....	127
14 Correlation Dimension Analysis .....	149
IV Discussion .....	171
15 Summary of Findings.....	172
16 A Neural Basis for Object Discrimination.....	174
17 Cortical Self-Organization.....	180
18 Relating the Present Findings to Alternative Models .....	192
19 Perception and Awareness.....	195
20 Extensions.....	197
21 Applications .....	198
Appendices .....	200
Appendix 1 Effect of Noise on Correlation .....	201
Appendix 2 - Stimulus Pictures .....	205
References .....	211

## List of Tables

Table 6.1 Results of analysis of variance of intercorrelation .....	60
Table 6.2 Mean intercorrelations .....	60
Table 6.3 Short vs. long intercorrelations .....	61
Table 6.4 Results of analysis of variance of lag.....	65
Table 6.5 Mean lags.....	65
Table 7.1 Correlation change .....	88
Table 7.1 Correlation change .....	89
Table 7.1 Correlation change .....	90
Table 8.1 Average correlation difference.....	99
Table 9.1 Results of analysis of variance of mutual information .....	105
Table 9.2 Mean mutual information.....	105
Table 9.3 Short vs. long mutual information.....	106
Table 10.1 Results of analysis of variance of coherence .....	114
Table 10.2 Mean coherence .....	114
Table 10.3 Results of analysis of variance of phase .....	115
Table 12.1 Discrimination Index.....	124
Table 13.1 Generalized Regression Neural Network Scores.....	143
Table 13.1 Generalized Regression Neural Network Scores.....	144
Table 13.2 T-test Results for Network Scores.....	145
Table 14.1 Correlation Dimension Estimates.....	164
Table 14.1 Correlation Dimension Estimates.....	165
Table 14.2 T-test Results for Correlation Dimensions.....	166
Table A1.1 Signal-To-Noise Ratios with Theta Band Filtering.....	202

## List of Figures

Figure 6.1 Typical stimulus picture .....	50
Figure 6.2 Electrode topography .....	52
Figure 6.3 Typical wave-form ensemble.....	52
Figure 6.4 Typical eye-blink wave-form.....	53
Figure 6.5 Correlation vs. time .....	61
Figure 6.6 Correlation vs. electrode spacing.....	62
Figure 6.7 Short and long distance correlations.....	63
Figure 6.8 Lag vs. time .....	66
Figure 6.9 Lag as a function of electrode spacing.....	67
Figure 6.10 Correlations vs. time, no-blink condition .....	68
Figure 6.11 Correlations vs. distance, no-blink condition .....	69
Figure 6.12 Association Fibers .....	73
Figure 6.13 Commisural Fibers .....	73
Figure 6.14 Projection Fibers .....	74
Figure 7.1 Topography of correlation differences, subject 4.....	82
Figure 7.2 Topography of correlation differences, subject 5.....	83
Figure 7.3 Topography of correlation differences, subject 7.....	84
Figure 7.4 Topography of correlation differences, subject 8.....	85
Figure 7.5 Topography of correlation differences, subject 9.....	86
Figure 7.6 Topography of correlation differences, average across subjects .....	87
Figure 8.1 Topography of net difference correlations, subject 5.....	97
Figure 8.2 Topography of net difference correlations, average across subjects .....	98
Figure 9.1 Mutual information vs. time .....	107
Figure 9.2 Mutual information vs. inter-electrode distance.....	108
Figure 10.1 Coherence vs. inter-electrode distance.....	116
Figure 11.1 Topography of coherence differences, average across subjects .....	120
Figure 12.1 Intercorrelation variance.....	123
Figure 12.2 Discrimination Index.....	124
Figure 13.1 Generalized Regression Neural Network classification results .....	146
Figure 14.1 Correlation dimension estimates for the before-blink epoch.....	166
Figure 14.2 Correlation dimension estimates for the after-blink epoch.....	167
Figure 14.3 Mean correlation dimensions.....	168
Figure 16.1 The Cortical Self-Organization Model .....	178
Figure 17.1 Prototypical inter-regional signaling topologies .....	183

Figure 17.2 Reciprocal interconnections in the visual system ..... 183  
Figure A1.1 Pearson product-moment correlation as a function of signal-to-noise ratio. .... 203  
Figure A3.1 Stimulus pictures..... 206

# I Introduction

## 1 Overview

The subject of this work was the connection between the process of mental model creation, and the underlying neuronal events. The phenomenon that was studied was the discrimination of a camouflaged object from its background. Such discrimination, it is suggested, involves to the formation of a mental model concerning the relationship of the target object and its background.

One approach to the study of model creation was taken by the Gestalt school of Psychology, which had its beginnings in Germany in the late nineteenth century. In 1890, Christian Ehrenfels discussed the "Gestalt qualities" of a perception. Ehrenfels questioned how it was that certain perceptual experiences could maintain their form, or gestalt, in the face of changes in sensory qualities (Leahey, 1987). An example is the perception of a melody as being essentially unaltered through transpositions of key. Gestalt psychology challenged the reductionist model of Wilhelm Wundt's structuralism, according to which complexes could usefully be analyzed into component parts. According to Gestalt psychological teaching, meaning emerges from the organization of the component parts of a perceptual experience. Consequently, such meaning is lost when the experience is analyzed into its components.

Gestalt psychology posited a number of rules of organization. One such rule is closure. As an example of closure, the image of an incomplete circle may nevertheless be perceived to represent a circle, rather than an arc segment. The arc and the background, and the relationship between them, organize to form a gestalt, a whole, out of which emerges the perception of the circle. This process, by which the interrelationships among the elements of the image give rise to a unitary percept can be viewed as an example of model creation. The model that is formed, on the basis of the immediate sensory data together with prior experience, is that of a complete but occluded, circle.

Ivan Pavlov in 1935 criticized Gestalt psychology for its idealistic basis, which he suggested was mentalistic and lacking in a physiological foundation (Brennan, 1991). Gestalt psychology addressed this criticism with the concept of isomorphism, a correspondance between the structural relationships in the perceptual field of an individual, and the underlying electrical brain field. While the perceptual field is evoked by sensory activity, the brain field is the result of electrochemical processes. The exact nature of this isomorphism has however remained obscure (Leahey, 1987). Karl Lashley expressed the question as "how [do] the specialized areas of the cerebral cortex interact to produce the integration evident in thought and behavior". The present study explored one facet of this question, the connection between the perceptual phenomena and underlying neuronal events during the process of model creation.

Such model creation occurs chronically and automatically throughout our interactions with our environment. At a perceptually simple level this process is involved in the circle completion phenomenon. At a more complex level model creation is involved in the perceptual-cognitive process of visually discovering an object that is initially indistinguishable from a camouflaging background. In the one case the model is that of an occluded circle. In the other case the model is that of the forms and corresponding identities of the initially camouflaged target and its background.

Traditionally, studies of brain function have involved the recording of EEG responses to stimuli whose physical characteristics change between conditions. In contrast, the present study used an unchanging stimulus, which however was designed to evoke a changing perception. It was to this changing perception, of the constant stimulus, that brain electrical activity was recorded. EEG signals were recorded while a subject performed the task of discriminating a camouflaged target. This task will be referred to as object discrimination. By *object discrimination* is meant the phenomenon that occurs when, after a period of visual examination, an individual is able to eventually distinguish a target object from within the context of a visual embedding matrix intended to have the effect of camouflaging the target, and to then identify the target as belonging to a particular semantic category. As an example, an initially indecipherable image might eventually be recognized as representing a bird sitting in grass (Figure 2.1). The image, which is first interpreted in one way, later comes to be interpreted in a different way. This study aims to identify the neurophysiological events that are associated with such shifts in interpretation to an unchanging stimulus. An experiment will be designed that will provide the conditions under which such discrimination can take place. The results of this experiment will be used to develop a description of the neural activity underlying object discrimination.

The direction taken in this work was determined in part by the findings of previous studies of both an empirical and theoretical nature. Empirical studies include investigations of brain function from two general methodological approaches. The first approach emphasizes the study of signal properties, and in particular the correlational structure, of single cell activity, during tasks involving visual perception. The second, complementary approach, through the study of single-cell and field potentials, and through the study of the effects of lesions, emphasizes the delineation of the anatomical and functional systems involved in the object discrimination process. Theoretical work includes a number of models which have been proposed to account for visual perception in terms of cortical integration. Models that will be discussed in Section 3 of the present Unit include the reentrant cortical integration (RCI) model developed by Finkel and Edelman (1989), and the convergence zone framework proposed by Damasio and Damasio (1993). Elements of these models, which view perceptual integration from the perspective of interregional organization, will be shown to be subsumed by a more

general description. According to this proposed description, the integration of perceptual experience is the outcome of a process of cortical self-organization driven by a mechanism of energy relaxation. This mechanism provides a natural direction for the evolution of the configuration of interregional signaling, without requiring the invocation of higher level coordinating mechanisms to direct the process of sensory integration.

Unit I contains discussions of a number of fundamental issues. Section 1 presents an overview of the present work. Section 2 discusses psychoneural identity, and more particularly the question as to whether mental events can be studied by physical means. Section 3 discusses a number of models of visual perception through cortical integration. Section 4 reviews a selection of neurophysiological studies of visual perception. Section 5 outlines a proposal for a model of the neuronal processes underlying visual object discrimination, and discusses a number of predictions that can be derived from this model.

Unit II details a number of experiments and analyses designed to test these predictions, using methods that can be classed as linear analytic procedures. Section 6 will describe an experiment in which interregional associations during object perception are examined, by looking at changes in the pattern of cross-correlations as a function of time, and as a function of the relative distance between pairs of electrode sites. Chapters 7 and 8 elaborate on the data obtained from the experiment described in Section 6. These sections examine the topographical distribution of the intercorrelations. Section 7 deals with changes in cross-correlation between all pairs of electrode sites. Section 8 looks at the topographical distribution of the mean correlation, that is, the extent to which the signals from each electrode site are correlated with the signals from all other electrode sites. Chapters 9 through 11 describe analyses of the data using alternative measures of association. Section 9 looks at the analysis of interregional coupling using mutual information, a measure of general association. Section 10 repeats this analysis using measures of coherence and phase, while Section 11 looks at the topographical distribution of coherence. Section 12 addresses an issue related to possible applications of the results of this study, the issue of the changes with time of two statistics computed for the cross-correlations as a function of time, the mean and variance. The Discrimination Index is then defined, as a factor which uses information about both the mean and variance of the correlations, and it is shown how this index changes with time before and after the moment of discrimination.

Unit III details analyses of the results of this study using nonlinear analytic techniques. Section 13 will describe the application of a neural network classifier that is trained to detect features in the recorded signals associated with the visual discrimination event. A number of previous studies applying neural networks to the problem of neuroelectric signal analysis are discussed. Section 14 deals with correlation dimension, a measure of dynamical system

complexity. The correlation dimension and its application to neuroelectric signal analysis is discussed.

Unit IV presents a discussion of the implications of the results of this study for a number of issues. Section 15 summarizes the findings of the present study. Section 16 elaborates on the model of neuronal processes associated with object discrimination. This elaboration continues in Section 17 with a discussion of the concept of cortical self-organization in perception. In Section 18 the present findings are related to alternative models of cortical integration. Section 19 speculates on the relationship between perception and conscious awareness. Section 20 lists a number of extensions of the present work, while Section 21 discusses potential practical applications of the results of this study.

Appendix 1 addresses a methodological issue, the effect of signal-to-noise ratio on strength of correlation. Appendix 2 shows the stimulus images used in the experiment.

## 2 Psychoneural Identity: Can Mental Processes be Studied?

The fundamental proposition underlying the association between physical evidence such as neuroelectric measurements and mental activity is the principle of psychoneural identity. Psychoneural identity is the assumption that every mental state or event can be identified with a corresponding distinctive brain state, and that this brain state in turn is an in-principle specifiable physiological occurrence (Reber, 1985). Regan (1989, p. 167) suggests that there is no general agreement among neuroscientists on this issue. He cites John Eccles as an instance of a neuroscientist who finds this position 'inconceivable' on the basis that neural machinery cannot be rich enough to act as the ground of consciousness and memory (e.g., Eccles, 1981). Although such an extreme position of rejection of psychoneural identity may be held by only a small minority of workers in this field, perhaps less infrequent may be an underlying feeling, expressed or unexpressed, that the scope of human potential, seemingly limitless in its capacity for innovation, is too wide and rich to be the product of less than one and a half kilograms of brain tissue.

Attempts have been made to try and reconcile psychoneural identity with the intuition that mental events transcend deterministic physics. Penrose (1989) for example has taken the position that mental phenomena such as consciousness are noncomputable, that these phenomena cannot be reduced to the level of algorithms carrying out deterministic computations. In support of this position, he has suggested that the small physical scale of microtubules within neurons is of the right order to allow quantum effects to exert a significant influence on the macroscopic level, and in this way bring an element of indeterminacy to the workings of the brain, freeing the results of neural activity from the apparent constraints imposed by determinism.

On the other hand even an admittedly superficial analysis of the information storage capacity of the brain reveals a perhaps unexpected depth. First of all, assuming that mental phenomena are not uncomputable, that is, mental phenomena could be expressed in algorithmic terms, then the question may be asked, what are the limits of such computation? Alan Turing (1937) proved that any procedure that could be expressed as a finite algorithm could be carried out by a universal Turing machine. A universal Turing machine is a device that, suitably programmed, can carry out any computation that can be expressed as an algorithm of finite length. An essential requirement of this Turing machine is an unlimited storage capacity. Taking  $10^{12}$  as the order of magnitude of the number of neurons in the brain (Kandel and Schwartz, 1985), with on the order of  $10^4$  synapses per neuron (Rosenweig and Leiman, 1989) gives a figure of  $10^{16}$  as the order of magnitude of the number of interconnections in the human brain. In general, information can be stored within a system such as the brain, in terms of the pattern of

synaptic connections, using at one extreme local representation of each bit of information, and at the other extreme a totally distributed representation of each bit. With local representation of the information, that is, with each bit of information coded at one single location in terms of the synaptic strength at one particular synapse, the number of storage units needed to store  $n$  bits of information would be  $n$ . The storage capacity  $C$  of such a system with  $n$  synaptic nodes would thus be simply

$$C \propto n$$

In this case, the brain, with on the order of  $10^{16}$  storage locations, would in principle be capable of storing on the order of  $10^{16}$  bits of information. At the opposite extreme from local representation is distributed representation of information. With distributed representation, a collective or network of interconnected synapses, rather than a single synapse, codes each bit of information. Importantly, any one synapse is assumed to be able to take part in many such networks and thus to participate in the encoding of multiple bits of information. With distributed representation, the information is encoded, not in terms of the strength of a single synapse of a single neuron, but rather in terms of a pattern of interconnections amongst the multiple neurons comprising a network. With such distributed information representation, the number of units of storage needed to code  $n$  bits of information is on the order of  $\log n$  (Gallant, 1993), assuming total connectivity and the ability for individual cells to participate in an unlimited number of separate networks. The storage capacity  $C$  of a system with  $n$  synaptic nodes would now be

$$C \propto 10^n$$

With the distributed representation method of encoding information, the brain would have a maximum storage capacity of on the order of 10 raised to the power of  $10^{16}$ . Again, it must be emphasized that the extent to which this storage capability could be realized would depend on the degree of interconnectivity of the network, and on the extent to which individual cells could take part in multiple networks. Anatomical studies reveal that a significant proportion of the mass of the brain is devoted to such interconnectivity, in the form of the diverse association tracts, commissures and projection fibers that make up the medullary centers (Barr and Kieman, 1988, p. 244). Admittedly superficial, this observation nonetheless suggests that interconnectivity must play a significant role in the functioning of the brain.

These information storage capacity estimates for the brain using distributed representation are only very broad order of magnitude values. Furthermore, it may be more plausible to consider that the brain makes use of some combination of local and distributed representation. Nevertheless, capacities even remotely near magnitudes such as 10 raised to the power  $10^{16}$  would seem to be large enough to account for the richness and diversity of

mental life, and the seemingly endless labyrinths of creative thought. If it can be accepted that mental events are not uncomputable, and that such events could be expressed as finite algorithms running on some equivalent of a universal Turing machine within the brain, then at least on the basis of the vast storage capacity available to such a machine, the limits of the computations that could be carried out by such a machine would appear to be correspondingly vast. Arguments, such as Eccles (1981) proposes, for the existence of mind separate from brain would seem to be an attempt to support a philosophical position on the basis of an incomplete reading of neurophysiological data.

The tendency to attribute a quality in the data to the object under scrutiny rather than to our interpretation of the data is not new. Consciousness, as the background against which the events of mental life are played out, is on the one hand compellingly real, and on the other hand seemingly irreducible to a physical basis. That consciousness appears to possess such a noumenal character, argues Churchland (1995), may be only the result of the lack of an appropriate conceptual framework. He cites as examples Ptolemy who in the first century dismissed the possibility of gaining knowledge about stars and planets because of their great distance from us, and August Comte who in the nineteenth century similarly argued that the constituents of stars could never be known because of their remoteness. Churchland's (1995) argument is that if an entity is apparently unknowable to us, then this may tell something about us, rather than about the entity. In the present case, the apparent lack of a conceptual framework from within which the phenomena of consciousness might be understood should be used to suggest that such a framework can not exist.

The view that mental events are in some way fundamentally distinct from material phenomena can be traced at least to the ideas of Gottfried Leibniz. Leibniz proposed a thought experiment (Leibniz, 1965) in which, if one were reduced to a minute size and were able to enter the brain and examine the machinery and processes of the brain, one would nevertheless be unable to find evidence of features of mental life such as thoughts, sensations and desires. On the basis of this argument this Leibniz concluded that such features must be separate from the physical machinery of the brain. Churchland (1995) points out that, in analogy with the earlier examples, Leibniz made the assumption that absence of evidence of the physical nature of mental entities was taken as evidence of their absence. Again, the alternative exists that we do not have the conceptual framework that would allow us to perceive such entities, whether or not in fact these entities constituted a separate level of reality or whether they were identical with some arrangement of neuronal elements. In making this argument, Leibniz takes a position that appears to be similar to the perspective adopted by more recent thinkers (e.g., Eccles, 1981) who assume that our inability to understand or to find a location within the brain for consciousness implies that consciousness must transcend the physical bounds imposed by

neurophysiology, rather than allowing for the possibility that we may be bringing insufficient analytical power to bear on the question. In taking this position, Leibniz appears to be assuming what he is trying to prove.

This argument, extended in contemporary philosophy by Searle's (1987) analogy of the Chinese room, thus suggests that entities such as understanding and knowing are distinct from and somehow transcend the neurophysiological machinery upon which they are based. It remains possible however that terms such as consciousness, knowing and feeling are labels that have been applied to phenomena in a way that is analogous to the way in which constellations were once constructed in the night sky. This fallacy of reification would seem to be an ever-present danger that might be particularly likely when consciousness is desperately striving to understand consciousness itself: there is no frame of reference for this understanding that is independent of the entity that is attempting to understand.

A further argument for the transcendent, non-physical nature of feelings has been put forward by Nagel (1974) who proposed the question of what it might be like to be a bat. While we might learn in complete detail about the neural functions of a bat's brain, we still would not, argues Nagel, know what it is actually like to be a bat. Since we can know only indirectly what it is like to be a bat, the experience of being a bat cannot be captured by the study of physical systems. This experience therefore must somehow transcend such physical systems. Churchland (1995) points out that there is a conflation in this argument between ways of knowing, and the nature of the object of the knowing itself. The bat knows directly what it is like to be a bat because of causal connections between its sensory and neural systems. A neuroscientist may infer, and thus know indirectly, what it might be like to be a bat, through studying this neurophysiology. These are statements about two ways of knowing, direct and indirect, with a common referent, in this case the experience of being a bat. This difference in ways of knowing is made more evident using an example with a more concrete referent. While I have direct knowledge of what it is like for me to be thirsty as a result of causal connections between various systems of my body, an observer can note from my behaviour signs that indicate this thirst, and know in this indirect way that I am thirsty. In both the case of knowing directly that I am thirsty, and in the case of the observer knowing this fact indirectly, the object of the knowing is the same: a constellation of physiological responses to a particular state of my body. This purely physical referent is in the one case known directly and in the other case known only indirectly. The point is that being known to an outside observer only indirectly does not necessarily render the referent insubstantial or transcendental. In this way, that the qualia or 'raw feels' that constitute experience can be known only indirectly by an outside observer does not imply that such qualia are necessarily nonphysical.

In the following section the discussion will proceed to particular issues. Several neurophysiological models of visual perception will be discussed. Overall, these models emphasize the importance of interconnectivity within a distributed processing network, the functional significance of the signaling that takes place through the medium of these interconnections, and the relationship between the pattern of signaling and the architecture of the network itself.

### **3 Neurophysiological Models of Cortical Integration**

The models that will be discussed in this section have been proposed to account for aspects of visual perception in terms of cortical organization. The first of these models is the reentrant cortical integration model proposed as a test of the theory of neuronal group selection put forward by Edelman (1989). This model emphasizes the primacy, in perception, of correlational activity resulting from reentrant signaling between neuronal groups. A second model which will be discussed is the convergence zone framework, proposed by Damasio and Damasio (1993). A third model which will be reviewed briefly is Mishkin's (1993) model of object recognition.

#### **3.1 The Theory of Neuronal Group Selection**

Gilbert (1995) classifies theorizing about visual mechanisms into two extreme positions. At one end, relatively passive filter models identify processing stages, with the visual system at each stage organizing and filtering information to be passed on the next level. There neuronal systems act on various basis sets, that is orthogonal combinations of features which, when assembled, can represent any arbitrary object. At the other extreme of theorizing, a position that may be referred to as a generalist view endows neural connections with the capacity to form and reform through interactions with the environment, so that any neuron or group of neurons can emulate any filter characteristic.

A model, suggested by Gilbert (1995) as being exemplary of the generalist position, is Gerald Edelman's Theory of Neuronal Group Selection (TNGS) (e.g., Edelman, 1987, 1992). A part of the TNGS that forms a prototypical example of a neuronal model of perception from a theoretical perspective is Edelman's view of perceptual categorization, the function that he views as "fundamental in any attempt to relate physiology to psychology." (Edelman, 1992, p. 89).

The TNGS is motivated by several observations. One observation is that the world does not appear to come prepackaged into perceptual categories. Second, there is observed to be a high degree of variation in brain structure both between and within individuals. According to Edelman's (1987) theory of neural Darwinism, of which the TNGS is an extension, in order for an organism to adapt to such an a priori unlabelled world, the organism's neural system acts according to a process of selection on this observed neural variability. Edelman proposes that the appropriate level at which to consider neural systems is a neuronal group, a neural structure involving multiple interconnected neurons. Neuronal groups exhibit a high degree of variability in terms of functional anatomy, or internal wiring. According to Edelman's theory, competitive activity between neuronal groups enforces a process of selection which operates on these highly variable neuronal groups. This process of selection works to promote the activity and survival of

some neuronal groups over others. Selection leads, in Edelman's terms, to differential amplification of particular variants within neural populations. The result of this selection process is that those neuronal groups are selected for whose activity is reinforced as a result of an individual's interactions with the environment.

The TNGS is thus based in part on the observation that there does not exist a precise point-to-point wiring scheme in any sufficiently rich neural system, such as that of invertebrates. This observation, along with the continuous rather than discrete nature of neural signals, Edelman appears to feel is support for an argument against conceptualizing the brain as a computer, performing computation. Neural wiring in such a system exhibits what Edelman terms degeneracy (1989, p. 50). Degeneracy refers to the observation that there does not appear to exist a one-to-one mapping between the activity of any particular system of neuronal groups, and any particular output or sets of outputs. Thus, multiple neuronal groups may be associated with the same output, while a single group may be associated with multiple outputs. This observation would appear to be fundamentally a statement of the distributed nature of neural systems. In distributed systems, all activity such as information processing and storage, is distributed among a system of processing elements, and is carried out in an asymbolic, or subsymbolic (Smolensky, 1988) form. This non-computational view of neural operation, Edelman (1989) contrasts with the information processing view of neural activity, in which functions are pre-programmed into systems which then carry out their actions in a correspondingly defined way, using computational procedures based on symbol manipulation.

Thus, building on the concept of neural Darwinism, Edelman (1987) proposed the theory of neuronal group selection, a feature of which deals with the manifest ability of neural systems to perceptually categorize an a priori unlabelled world. This ability is the result of two processes. First, in embryonic development, selection acts to establish the general configuration of neuronal networks. Second, during development and as a result of behavioural interaction with the environment, selection of certain neuronal groups over others acts to establish functional circuits and maps. Maps are functions that transform sensory inputs into corresponding cortical representations. In this way, perceptual categorization can take place, as a result of interactions with the environment acting in concert with neuronal selection. The result is that within the brain, perceptual categories about the world are created.

An essential feature of the TNGS relevant to the process of perceptual categorization and, thus, to the present work is the process of complex reciprocal signaling between neuronal groups which Edelman refers to as reentry, and which he describes as "temporally ongoing parallel signaling between separate maps along ordered anatomical connections" (Edelman, 1989, p. 65). An analogy may be made with computer code modules that are designed to take in, process, and output information, with the possibility that one such module of code, referred to

as reentrant code, can originate data which then directly or indirectly reenters the originating module in a recursive loop. The analogy between such recursive code and neural functions should not however be pressed beyond noting that brain regions, on a general systems level, might be viewed as processing modules similarly capable of exchanging complex information, along the feedforward and feedback connections within the numerous association tracts, commissures and projection fibers. Edelman (1992, p. 95) suggests that this signaling should show up in the form of correlations between the activity of brain structures, and hence as correlations in the EEG.

Reentry is to be distinguished from simple feedback. This distinction might be supported by the following argument (although Edelman himself does not appear to argue in these terms). While the effects of feedback can in principle be predicted using linear analytic methods, the effects of reentry can not. This effect is the result in part of a topological difference between feedback and reentry: Reentry is distinguished from feedback in that reentry is characterized by a multiplicity of pathways along which the information transfer takes place. In alternative terms, signal transfer in reentry occurs in parallel, while signal transfer in simple feedback occurs in serial. This parallelism associated with reentry has the significant outcome that the individual signals have the potential for interacting at their destination. The effect of such interactions is a nonlinearity in the nature of the information exchange, with the result that the neuronal groups, associated through such reentrant signaling, comprise a nonlinear system. The behaviour of such nonlinear systems, although fully deterministic, has been shown to be in general unpredictable, and potentially chaotic (e.g., May, 1976). These mutual interactions inherent in reentrant signaling, among the multiple signals carried over multiple and reciprocal parallel paths among neuronal groups, result in a high degree of functional complexity.

As a result of such reentrant signaling, according to the TNGS, neuronal groups become coupled and coordinate their individual activities. In particular, reentrant signaling among multiple neuronal groups has the effect of synchronizing their operation, resulting in coherent activity among these groups.

The concept of reentrant signaling between neuronal groups is relevant to the present study of visual perception in two ways. First, according to the TNGS, reentry is the underlying neurophysiological factor that forms a substrate for the coupling between multiple neuronal systems, and that results in synchronization between the activity of such systems. Second, this reentry-based association is responsible for brain functions ranging in scale from perceptual categorization to consciousness.

Directly relevant to the present work, perceptual categorization, according to the TNGS, occurs as the result of the activity of multiple neuronal groups, coordinated through the process of reentrant signaling. In particular, according to the TNGS, perceptual categorization is the

result of reentrant signaling between neural systems, each dealing with to some extent orthogonal dimensions of the sensory features of a stimulus. Such systems implement functions which map a particular facet of the sensory information, onto particular cortical regions. According to the TNGS, for nontrivial perceptual categorization to take place, at least two such maps, each carrying information about a disjunctive aspect of the stimulus, must be connected through reentrant signaling. Through this signaling, the various disjunctive aspects of the stimulus are mutually bound, resulting in the formation of a higher-order response to the initial stimulus. The reentrant signaling among these several maps thus leads to the creation of a higher-order concept, making use of a subset of the orthogonal features of the stimulus.

Importantly, the reentrant signaling which synchronizes the activity of the connected maps acts to promote synaptic changes within these maps. Straightforwardly, this can occur when cells within such maps, interconnected through reentry and firing synchronously, are likely to receive multiple, simultaneous inputs. The more neurons, comprising an interconnected set of maps, that are firing synchronously, the more likely it is in turn for any one of the participating neurons to receive simultaneous multiple excitatory inputs, and consequently the more likely it is for these neurons to themselves fire. Repeated exposure to the same stimulus environment then acts to reinforce such neural connections through repeated activation of same synapses. These statements are of course a description of Hebbian learning (Hebb, 1949).

Edelman's (1989) view of perceptual categorization involves the notion of what he refers to as primary consciousness (p. 104), a fundamental level of awareness of the internal and external world. Primary consciousness is comprised of experiences such as mental images, and is constrained to a time history defined by the span of short-term memory (p. 24). Primary consciousness is the result of the following sequence of processes. Perceptual categorization of exteroceptive signals, of the world external to the individual, is carried out by reentrant signaling within and between cortical regions as well as between the cortex and subcortical regions such as the thalamus. In parallel with this process, perceptual categorization of internally generated, interoceptive, signals is mediated by reentrant interactions within and between limbic and brain-stem systems as well as by biochemical signaling systems. These two categorization systems may each be considered to represent maps. In the one case exteroceptive signals, and in the other case interoceptive signals are mapped onto cortical regions. In turn, and at the next level of abstraction of the information generated within the brain, the thalamo-cortical and the limbic-brain stem systems interact. Their information products are compared, presumably also through the process of reentrant signaling, in the hippocampus, septum and the cingulate gyri. The results of this comparison are in turn re-categorized in the cingulate gyri, temporal lobes and the parietal and frontal cortex. Edelman (1989, p. 156) refers to the results of this process as conceptual categorization. Although not a statement made by the TNGS, this conceptual

categorization might be considered to form a component of an organism's internalized model of the world. Significantly, this conceptual categorization itself then is reentrantly connected back to the neuronal systems involved in the perceptual categorization of new exteroceptive signals. The results of the conceptual categorization thus influence subsequent perceptual categorizations. Essentially, a current set of perceptual categories formed about the world are in this way modulated by the results of the previous conceptual categorization process. This effect would appear to be a statement about the influence of the internal world model, on on-going perception. How the world is interpreted at any one instant is influenced by how the world was interpreted at previous times. The TNGS proposes that this effect of conceptual categorization, or in alternative terms, of an internalized world model, on subsequent perceptual categorization is the basis for primary consciousness.

Primary consciousness, founded on recurrent signaling between neuronal populations, is thus itself the outcome of a recursive process in which notions of self and nonself are interpreted in terms of each other, resulting in a construct which on the one hand forms a context in which subsequent notions about the world are interpreted, and which on the other hand is itself reevaluated in terms of such notions about the world.

These ideas are relevant to the present study in the following way. The reentrant cortico-cortical and thalamo-cortical signaling that Edelman (1989) associates with perceptual categorization may be expected to show an effect of the experimental manipulation that will be carried out in the present work. In particular, the effects of reentrant signaling associated with object categorization and discrimination should be observable as correlated activity between the relevant cortical areas, when subjects are engaged in a task involving such categorization.

Reentry is the basis for a computational model of cortical integration, the reentrant cortical integration (RCI) model (Finkel and Edelman, 1989). This model demonstrates the efficacy of recursive information interchange, reentry among multiple local processing networks along reciprocal feedforward and feedback interconnections. By means of this reentrant signaling such networks are able to exchange information regarding stimulus feature discriminations performed locally by each of the networks, allowing all networks to mutually make use of the results of such discriminations. Reentrant signaling also allows responses to complex and illusory stimuli to be synthesized since information generated by each network eventually returns to that network along reciprocal pathways. These functions were facilitated by the elimination of conflicting responses generated by the individual networks, also as a result of the inter-regional reentrant signaling. A feature of the RCI is that no central coordinating structure is required to direct the process of feature integration and synthesis. Finkel and Edelman (1989) found that during this integration process, changes in network coupling strengths were associated with episodes of correlated activity among the local networks.

### **3.2 Neuronal Ensembles**

A second model of visual perception, and an example of a neuronal model of perception from a neurophysiological perspective, is Mishkin's (1993) model of object recognition, a model which in some general ways resembles Edelman's TNGS. According to Mishkin (1993), object recognition involves the re-activation, by a visual stimulus, of a previously formed cell assembly. A retinal stimulus sequentially activates large neuronal groups in the occipito-temporal circuit. Mishkin refers to these groups as neuronal ensembles, and suggests that they form the neuronal representation of the visual stimulus. In the latter stages of this occipito-temporal activation, sequentially-connected temporal, thalamic and frontal areas of the limbic system become involved. These latter activations then lead to the strengthening of feedback limbo-cortical and cortico-cortical synapses. Through this process, cell assemblies are created (Hebb, 1949). These cell assemblies, which form the stored representation of the stimulus, are subsets of the neuronal ensemble. Visual recognition, Mishkin (1993) suggests, occurs when a neuronal ensemble that has been activated by a visual stimulus in turn re-activates a such an existing cell assembly. While the neuronal ensembles are activated by feedforward signals from the retina, cell assemblies can be activated both, by feedback signals from the limbic system, as well as signals more directly from the neuronal ensemble. Mishkin's notion of memory activation with the involvement of the limbic system resembles Edelman's model of primary consciousness which involves, not so much sequential activation of, but the correlated activity between, the limbic system and primary and secondary sensory cortices.

### **3.3 The Convergence Zone Framework**

A third approach to the problem of perceptual integration, and also a model based on neurophysiological evidence, is the convergence zone framework proposed by Damasio and Damasio (1993). While Mishkin's (1993) model sees perception in terms of the creation and later reactivation of neuronal ensembles by a sensory stimulus, Damasio and Damasio's (1993) view of the cortical processes underlying perception and memory involves the storage and subsequent reactivation of a code that contains instead, only a key. This key encodes the pattern of interconnections among the various participating neuronal groups involved in the original perception, a pattern that is later reinstated during recognition and recall. Damasio and Damasio (1993) base their model of large-scale distributed processing in perception and memory on the findings of lesion studies in humans. As an example, perception is not disrupted by bilateral lesions of anterior temporal or prefrontal cortices, while lesions in many sensory regions do impair perceptual integration processes. Thus, within the visual system, lesions to posterior regions, including the inferior occipital cortex, impair the retrieval of stimulus features such as

color (Damasio, Yamada, Damasio, Corbett and McKee, 1980). Lesions to intermediate regions including the inferotemporal cortex, spare feature retrieval but can impair category level recognitions for items learned through vision alone (e.g., Warrington and Shallice, 1984). Lesions to anterior regions including bilateral medial temporal regions, nonmedial anterotemporal cortices, and parts of the inferotemporal cortex, impair retrieval of item level components, while leaving intact retrieval of features and category level components (Damasio, Tranel and Damasio, 1989). The authors interpret this pattern of findings as suggesting the existence of a hierarchy of knowledge retrieval or access. This hierarchy is associated with a corresponding hierarchy of cortico-cortical connections within the feedforward chain from primary visual cortex to entorhinal cortex. Retrieval of more complex knowledge is associated with pathways located towards the end of the feedforward chain, closest to the entorhinal cortex. Retrieval of more elementary or lower level knowledge is associated with connections located towards the start of this chain, closest to the primary visual cortex.

The findings of such lesion studies, the authors conclude, indicate that the integration of perceptual phenomena is not determined by any single cortical area. Instead, they suggest, such perceptual integration must involve multiple cortical areas involved in distributed processing of information, rather than involving processes within localized cortical regions. In turn, the authors propose that these large-scale networks are controlled by ensembles of neurons, of which they suggest there may be thousands. Damasio and Damasio (1993) term these ensembles convergence zones.

Functionally, convergence zones are groups of control neurons whose function, Damasio and Damasio (1993) propose, is to organize networks of interconnections over a wide range of scales, within both cortical and subcortical regions. Such organization, they suggest, involves the selective strengthening of a subset of the feedforward and feedback loops, connecting cortical and subcortical regions, that pass through a convergence zone. These groups of control neurons thus coordinate the activity of multiple cortical regions, and exert an influence during both initial perception and subsequent recall. During perception, convergence zones encode the pattern of interrelations between the associated sensory regions. During recall, convergence zones reinstate these associations. The effect of this mechanism, the authors suggest, is to coordinate processes that are distributed over the large-scale networks that are involved in perception and memory.

Neuroanatomically, convergence zones are suggested to be neuronal collectives that are focal points for multiple feedforward and feedback loops. Anatomical evidence for such zones includes the finding of large-scale divergent projections (Bressler, 1995). Examples include sites in the medial pulvinar nucleus of the thalamus that are found to project to multiple widespread cortical regions, with these cortical regions themselves being interconnected by cortico-cortical

association tracts (Asanuma, Andersen and Cowan, 1985). Bressler (1995) suggests that the function of the pulvinar nucleus may be to prime particular sets of cortical regions, and in this way to facilitate communication and interaction between these regions.

Convergence zones are thus proposed to be a mechanism for binding knowledge at various stages of complexity (Damasio and Damasio, 1994). In accord with the findings of lesion studies, low level convergence zones located in relatively posterior cortical regions function to bind elementary stimulus features, while higher level zones in more anterior cortical locations would bind correspondingly more complex feature transformations.

The convergence zone framework can be compared with Finkel and Edelman's (1989) RCI model based on Edelman's (1989) TNGS. In terms of architecture, Damasio and Damasio's (1993) model emphasizes a convergence-divergence topology. In this topology, convergence zones are centers onto which multiple pathways converge and from which multiple pathways diverge. In contrast, the RCI model of perception posits inter-cortical mappings that are relatively completely and reciprocally interconnected, both hierarchically and heterarchically. A second point of distinction between these two models is that the RCI model directly addresses the issue of the dynamical behaviour of inter-cortical communication during perception. The convergence zone framework emphasizes instead the changes in functional topology of cortical regions, and how these changes are mediated by key neuronal ensembles during perceptual and memory operations. These distinctions between the two approaches, which to some extent represent orthogonal and even complementary views on perceptual and memory processes, are reflected in the general ways in which tests of the models could be carried out. The convergence zone framework, developed in part on the basis of the results of lesion studies, would appear to be most directly testable by means of such studies. Preexisting or experimental lesions could be used to verify the existence and extent of the proposed convergence zones. While the convergence zone framework does imply coordinated activity among multiple cortical regions along the feedforward and feedback paths directed by a convergence zone, the RCI model speaks more directly to the dynamics of intercortical communication by means of reentry. The RCI model would correspondingly be more directly verified by studies of inter-cortical coherence. From a network topological perspective however, the convergence zone framework is consistent with the view that synchronous intercortical activity is mediated by one or more pacemakers, that is, centers which commonly drives the oscillatory activity of multiple neuronal regions resulting in synchronized activity among these regions. In contrast, the RCI model would appear to be relatively neutral on this issue, since architectural concerns are not the primary focus of this model. The RCI model rather focuses on a theoretical description of the systems level events, and the nature of the associated signaling, that take place during perception. Admitting the point that the RCI model and the convergence zone frameworks may be contrasted to only a limited

extent, the RCI model nevertheless does not entail the concept of synchronization by means of a pacemaker, and in this sense might be seen to be the more parsimonious view. A contrasting position to the pacemaker view of intercortical association is the notion that multiple cortical regions can self-organize through the agency of interregional signaling, to create a larger scale network with the individual cortical regions engaged in mutually synchronous activity. Finkel and Edelman's (1989) RCI model would appear to be more clearly consistent with such a view, that perception involves a self-organization of multiple cortical regions by means of reentry. The principle of cortical self-organization will form a central component of the model of perceptual integration that will be presented in the present study.

## **4 Studies of Visual Perception**

### **4.1 Studies of Correlated Activity in the Visual System**

One approach to the study of visual processes is the analysis of intercorrelations between the activity of individual cells in avian (Gray and Singer, 1989; Gray, Konig, Engel and Singer, 1989) and cat (Gray and Singer, 1989; Engel, Konig, Gray and Singer, 1990) visual systems.

In a prototypical study of correlational activity in the visual system, autocorrelation analysis of single-cell recording in area 17 of cat visual cortex has revealed that the activity at a high proportion of recording sites has a prominent periodic component, with a frequency in the gamma band, between 40 and 60 Hz (Gray, Konig, Engel and Singer, 1989). Cross-correlation analysis applied to pairs of recording sites showed that approximately half of the site pairs that were tested showed significant levels of linear association. The study next examined the patterns of auto and cross-correlation as a function of visual stimulus characteristics. A single long moving light bar stimulus was found to elicit oscillatory activity that was synchronous across individual cells within a group of cells, with particular groups of cells becoming synchronized depending on the orientation of the bar. This synchronization was marked by a zero relative phase angle, and persisted for recording sites with the same orientation preference, that were separated by distances of up to 7 mm in the case of two subjects. The wide separation between these recording sites suggested that the sites had non-overlapping receptive fields. Consequently, the authors hypothesized, these sites should be activated by both a single long moving bar, as well as by two shorter moving bars. When two shorter bars were moved in the same direction, correlated activity was found at the corresponding recording sites. When the two bars were moved in opposite directions however, the correlation disappeared. The authors suggest that possible neuroanatomical substrates for these correlations might be tangential connections within the visual cortex, or back-projections from other cortical regions. Furthermore, they propose that such interareal correlations may function as a mechanism to accomplish transient binding of stimulus features such as orientation, continuity and organization of motion. Such a binding process, they point out, is essential to elementary stages of visual analysis involved in, for example, figure-ground discrimination.

While horizontal connections within the striate cortex have been suggested to mediate gamma band oscillations (e.g., Gray and Singer, 1989), Gilbert (1995) suggests that it is not clear first, whether such oscillations result in binding or segmentation, and second, whether these oscillations are merely an epiphenomenon. That is, gamma band oscillations and dynamic

changes in receptive field properties may both be the result of synaptic changes within the horizontal connections.

Evidence of more widely distributed coherences, among visual regions of the macaque monkey, were found in a study by Bressler, Coppola and Nakamura (1993). In this study the subjects were trained to either press or ignore a bar in response to particular visual stimuli. Coherences in several frequency bands ranging from 12.5 Hz to 87.5 Hz, were computed between signals recorded from electrodes placed in a number of regions, including striate, prestriate, parietal and motor cortices. In the press condition, significant increases in coherence were found between striate and motor areas immediately preceding and following the actual response, and between striate and parietal areas in the interval between stimulus onset and response onset. In the ignore condition the striate-motor coherence was absent, while the striate-parietal coherence was similar to that in the press condition. Thus, the increase in coherence, an indication of synchronized activity, between striate and motor regions was a function of whether or not a motor response was required, while the increase in striate-parietal coherence occurred in both conditions. The magnitude of the coherence increases was roughly inversely proportional to frequency: The largest increases occurred in the lowest frequency band centered on 12.5 Hz. This finding is consistent with the results of previous studies with humans that showed multi-regional coupling below 10 Hz in scalp-recordings of averaged event-related potentials (Gevins, Morgan and Bressler, 1987). While coherence increases in the Bressler et al. (1993) study extended to the gamma band, there was no evidence of a relative increase in the gamma frequency band as had been suggested previously (e.g., Bressler, 1990). The authors concluded that the coherence increases were not the result of the appearance of the stimulus itself, but rather, appearing after stimulus onset, were the result of subjects' discrimination between the two different stimuli, and subsequent preparation to respond. They suggest that these coherence increases, appearing at multiple frequencies, reflect inter-cortical synchronization on multiple time-scales, affording cortical processing with a flexibility that would not be available were the coherence increases restricted to narrow frequency bands. They suggest further that the wide spatial distribution of the coherences indicates that such synchronization can take place between any cortical regions, and lends support to models of high-level functions such as perception and action that involve binding between multiple and wide-spread cortical regions.

## **4.2 Studies of Visual Processes in Non-human Primates**

Much work has been carried out in the investigation of the neuronal processes that are involved in object recognition and discrimination. Many of these studies have involved non-

human subjects, notably the macaque monkey, upon which procedures such as single-cell recordings are more conveniently carried out. For this reason, some caution would appear to be in order in applying these results to theorizing about human visual processes. With this caveat in mind, a number of studies will now be reviewed which bear on the questions of what neuronal systems are involved in object recognition, and how these systems interact to produce the phenomenon of integrated perception. The purpose of reviewing such studies in the present context is twofold. First, it is to sketch out in broad terms a picture of which systems act in what way during visual perception, in order to suggest features that might be expected to be present in the results of the present study. Second, it is as to provide a point of reference for later discussions of the implications of the finding of the present study for the cortical dynamics associated with object discrimination in humans. These studies also collectively demonstrate the complexity of the visual system, and therefore the limited value of attempting any analysis of the visual system based on a reductionist approach.

One perspective on the visual system suggests an analysis into a component involved with spatial properties of the stimulus and a component involved with stimulus features (Ungerleider and Mishkin, 1982; Mishkin, Ungerleider and Macko, 1983). According to this analysis, information flow within the visual system occurs within a dorsal parietal and a ventral temporal pathway, both originating in the primary visual cortex. The parietal pathway, directed towards parietal lobe components associated with spatial and motor activity, is involved in tasks related to the spatial location of objects in the visual field and spatial guidance of motor responses. The temporal pathway, directed towards the inferotemporal cortex, is associated with the analysis of visual form and pattern. Thus, the parietal pathway may be generally characterized as dealing with 'where' information regarding the stimulus, whereas the temporal pathway can be thought of as dealing with 'what' information.

More recent appraisals have suggested that such an analysis may be insufficient to account for anatomical and functional evidence linking these pathways (Maunsell and Ferrera, 1995). Anatomical evidence includes the finding of interconnections between the parietal and temporal pathways (Maunsell and Van Essen, 1983). Functional evidence includes the finding of extra-retinal signals in the visual cortex, related to functions such as eye position and movement (Anderson and Montcastle, 1983) and memory (Miyashita, 1988), suggesting that the temporal and parietal visual pathways may be specialized for extra-retinal signals similar to the way in which these pathways are specialized for signals of retinal origin. One study of memory effects in the visual system (Maunsell, Sclar, Nealey and DePriest, 1991) used a match to sample task. A macaque monkey was presented with a sample stimulus, a visual grating with one of 4 possible orientations, followed after a 600 to 800 ms interval by a similar test stimulus with one of the 4 orientations. The subject was trained to press a lever when the test and sample

orientations matched. A primary interest in this study was to determine how the cell responses were affected by the grating orientation for which the subject was looking. During the presentation of the test stimulus recordings were made from cells in V4 with receptive fields covered by the stimulus. While a majority of neurons responded to a particular test stimulus orientation irrespective of what stimulus had been presented as the sample, approximately 25% responded to a test stimulus orientation only when this orientation matched that of the sample stimulus. The authors suggest that these neurons are associated with a memory function in which the subject is remembering the orientation of the sample stimulus. Such cells are sensitive therefore, the authors propose, to the memory of a stimulus orientation, analogous to how the other tested cells are responding to stimulus orientation itself. In contrast neurons in area V1 showed relatively little effect of the orientation that the subject was seeking. The study next looked for evidence of spatial memory in cells in the parietal pathway, using moving dot patterns with 4 possible directions of motion. Cells were found that were sensitive to direction of motion of the sample, but the effects were relatively weak in the middle temporal area (MT), and only somewhat stronger in later stages that included the medial superior temporal area (MST) and area 7a. Interestingly the authors found that the response of cells in area V4 of the temporal pathway, to direction of motion, were as strong or stronger as the response of the cells in the parietal pathway. Maunsell et al. (1991) concluded that the association between performance on behavioural tasks and particular visual areas may not be well predicted by the response properties of the neurons within the respective areas. In place of the 'what/where' functional distinction proposed by Ungerleider and Mishkin (1982), Maunsell and Ferrera (1995) suggest that a more accurate description may be one proposed by Goodale and Milner (1992): The temporal pathway may deal with object identification ('what'), while the parietal pathway is concerned with visual guidance tasks involved with the guidance of visual behaviours ('how').

The complexity of the processing associated with visual perception is demonstrated also by studies showing that feedback projections within the visual system are associated with higher level influences such as attention. In the macaque, for example, the receptive field of neurons in area V4 has been found to be sensitive to attention, with the size of the receptive field becoming restricted when the animal is not attending to a stimulus (Moran and Desimone, 1985). In this study, it was found that the feature specificity of cells in area V4 can also be modified by attention. Neurons that were sensitive to stimulus orientation when the animal was attending to the stimulus became insensitive to orientation when the animal was not attending. Mishkin (1993) points out that the dorsal and ventral visual pathways contain both feedforward and feedback projections, and that the feedback projections may have a function in stimulus attention and memory. Gilbert (1995) similarly suggests that feedback projections within the visual system provide an anatomical substrate for such influences, influences that may extend to the

earliest levels of visual processing. In addition to such higher level influences on the visual system, Gilbert (1995) suggests that within the primary visual areas horizontal connections mediate higher-level effects such as context dependency. Cells within the striate cortex for example are interconnected by horizontal connections. Such connections, he suggests, enable individual cells to integrate information from extended cortical areas, and therefore from an extended portion of the visual field that, because of the visual topography of the striate cortex, is larger than the receptive field of individual neurons. Cells in the visual cortex are known to alter their functional specificity with changes in the sensory context within which a stimulus is presented (e.g., Kanisza, 1979). Such context dependency, Gilbert (1995) proposes, may be mediated by the horizontal connections within the primary visual areas. These horizontal connections have been found to target cells of similar functional properties such as, for example, orientation preference (Gilbert and Wiesel, 1989). Such findings, Gilbert (1995) suggests, call into question the notion of a neuronal receptive field.

The influence on the visual system of higher level effects such as attention and memory, suggest Desimone, Miller, Chelazzi and Lueschow (1995), is the result of feedback from the prefrontal cortex. Such feedback can have a priming effect on inferotemporal neurons, constituting a higher level influence on object recognition. The typically large number of objects in a visual scene requires selection processes that must function at all levels of the visual system. In early stages such processes involve pre-attentive functions involved with, for example, figure-ground separation, that serve to increase the saliency of anomalous shapes within the visual field. Complementing such relatively low-level processes are attentive and memory processes that are similarly engaged in the task of selecting subsets of the objects within the visual field. These memory processes consist of both automatic processes based on stimulus repetition that could bias a response towards novel stimuli, as well as cognitive processes involving working memory that take part in the analysis of temporal sequences of stimuli. Miller, Li and Desimone (1993) investigated such memory processes in the visual system of a macaque monkey with a match to sample task similar to that used by Maunsell et al. (1991). Recordings were made from cells in the anteroventral portion of the inferotemporal cortex (IT), an area which has been associated with visual memory (e.g., Meunier, Bachevalier, Mishkin and Murray, 1993). Subjects were presented repeatedly with initially novel sample stimuli with a different set of stimuli for each cell. Intervening stimuli were presented between each sample presentation. Between-trial analyses of recordings made during stimulus presentations revealed a number of findings. First, there was an expected response to the different stimuli. This finding was expected on the basis that cells in the IT cortex have been shown to be responsive to stimulus features such as shape, texture and color (e.g., Tanaka, Saito, Fukada and Moryia, 1991). Second, in approximately one-third of the IT cells, a

decrement in the response of the cells was found, with each successive sample presentation, demonstrating a familiarity effect. This response suppression effect lasted through 150 intervening stimuli. The pattern was repeated when a second, new, sample stimulus was used. The authors refer to this process as an adaptive mnemonic filter, and suggest that it may activate orienting and attentional systems to favor novel stimuli. Desimone et al. (1995) propose that the memory functions within the IT may be associated with cells whose response does not change over presentations, while the IT neurons with declining responses may be those that have coded non-critical features of the stimulus.

Miller et al. (1993) next examined short term memory effects in the anteroventral area of the IT cortex. In a variation of the original match to sample task, subjects were presented with a familiar sample stimulus, followed by from 1 to 6 test stimuli. Cell responses during the presentation of the final test stimulus showed a clear distinction between the case when the test stimulus matched or did not match the sample stimulus. Responses to matching test stimuli was suppressed relative to the responses for non-matching stimuli, even with 5 intervening stimuli corresponding to an interval of several seconds. This response suppression was found for approximately half of the tested cells. All tested cells showed a response preference related to stimulus features. The rapidity of these suppression effects, which occurred within 80 ms of stimulus onset, the authors claim argues against the cause being feedback from other areas, but suggests rather that a sensitivity to repetition may be intrinsic to the visual system.

The findings of a related study (Miller and Desimone, 1994) suggest that these memory effects are not due to low-level comparison processes, since similar effects were found when test items are presented in a different size or location on the retina (Lueschow, Miller and Desimone, 1993). Miller and Desimone (1994) addressed the question of whether the match suppression might be due to a voluntary, working memory process, or an automatic repetition detector. The paradigm in this study included trials presented as in the Miller et al. (1993) study, along with trials in which, as well as a possible match between sample and test stimuli, there might occur a match between test stimuli within a trial. If the response suppression mechanism involved a voluntary memory mechanism, then response suppression should be observed only for test stimuli that matched the sample and not to test stimuli that matched other test stimuli. If on the other hand response suppression is the result of a repetition sensitivity within the IT cortex then suppression should occur in both cases. As before subjects were trained to respond to matches between the sample and the final test stimulus. Responses were found to be equally well suppressed for both a match to sample and a match to an irrelevant test stimulus, supporting the view that the response suppression mechanism involves a sensitivity to repetition. However, a portion of the cells that showed significant memory effects were sensitive to matches to sample stimuli, but not to matches to test stimuli.

Together, these findings indicate the existence of two short-term memory systems in the anteroventral regions of the IT cortex, suggest Miller and Desimone (1994), an automatic repetition-sensitive mechanism and a voluntary working memory system. The authors propose that this working memory might depend on priming of IT neurons by ventral pre-frontal cortex, with which the IT cortex is substantially interconnected. Evidence in support of this suggestion includes the findings of lesion studies showing consequent performance decrements on matching to sample tasks (e.g., Mishkin and Manning, 1978).

In summary, the results of these behavioural studies with macaque monkeys indicate the existence of two short-term memory processes in the visual system, a mechanism sensitive to stimulus repetition effects and involved in orienting to novel stimuli, and a process involved with maintaining in short term memory a representation of an attended-to object in the visual field.

### **4.3 Generalizing to Humans**

There appear to be a number of reasons why the results of studies such as many of those cited above, that base their findings on work with non-human subjects, may only with limitations be generalized to humans.

First, studies of object discrimination using non-human subjects would in large part seem to make use of a single modality, involving primarily an analysis of the visual features of the object. In contrast, visual discrimination in humans can be relatively more complex. Both visual and auditory modalities can be involved, since humans have the option of making use of language related processes. Object discrimination by humans can involve a naming process, the retrieval of a label, along with an analysis of visual features. Furthermore, these modalities may interact. Auditory contextual cues for example can modify the processing of visual cues. The ubiquitous presence of interconnections between cortical areas in general and between primary sensory areas in particular at least provides a substrate for such interactions to occur.

Second, investigations of object discrimination processes can not reasonably be carried out in an intentional vacuum. For completeness sake, an analysis of the neuronal processes associated with object discrimination should entail consideration of the purpose for which the discrimination is occurring. In studies with human subjects, the manifold purposes associated with object discrimination in the real world devolve primarily, although not exclusively, into the relatively simple goals associated with complying with experimental instructions. These goals, whatever their specific nature may be, can in no way be compared with those of non-human subjects, and in particular in view of the paradigms necessitated by single cell recordings. These different higher level influences such as intention on visual perception should be reflected in corresponding differences in neural activity. In both human and non-human subjects the effects

of higher-level influences such as attention and memory on discrimination processes have been demonstrated. There would not appear to be any principled reason why these influences should not include the effects of factors such as intention, the goal or purpose for which the discrimination is occurring. It seems reasonable to conclude, on the basis of these factors, that caution should be exercised in generalizing from non-human to human subjects.

#### **4.4 Studies of Human Visual Processes**

One finding from studies of visual perception in non-human subjects that does appear to be applicable to human visual processes is the functional specialization of the visual system into featural and spatial processing systems. Evidence for a dorsal-ventral dichotomy in human visual systems has been found in both lesion studies, as well as in a series of functional imaging studies involving measurement of cerebral blood flow using tracer compounds labeled with radioactive isotopes.

An example of a lesion study involved two groups of patients, undergoing either left or right anterior temporal lobectomy (Hermann and Seidenberg, 1993). Patients were administered object recognition task involving face recognition, and spatial localization tasks involving line orientation. The patients performed the tasks immediately after the operations, and after an interval of 6 months. On both sets of tests a dissociation was found in patients' performance on the two types of tasks. Patients demonstrated a significant loss in facial recognition ability, while gaining a concomitant improvement in line orientation performance. This pattern of performance was found to be similar for both groups of patients, those who had undergone the left and the right anterior temporal lobectomies. The results show clearly that the anterior temporal lobectomy had a specific effect on the object recognition system within the occipitotemporal region, while leaving the spatial perception system in the occipitoparietal region relatively unaffected.

A lesion study by Newcombe and Ratcliffe (1987) made use of both case and postmortem data on 2 male subjects who had suffered brain trauma as young adults. Case 1 was injured at age 31 and died at age 70. Damage had occurred in the midsection of the right rolandic region, and the patient had scotomata in the left visual field. Case 1 was unable to perform cube counting and maze learning tasks but was able to recognize familiar faces and to perceive shadowed faces. Case 2 was injured at age 29 and died at age 50. Damage had occurred in the right temporal lobe, and the patient had visual impairment of the left upper quadrant of the visual field. Postmortem examination revealed a cyst in the right temporal lobe. In contrast with case 1, case 2 was able to perform cube-counting and maze-learning tasks but was unable to recognize familiar faces or to perceive shadowed faces, a reversal of the

symptoms displayed by Case 1. These results are generally consistent with an analysis of the visual system into dorsal and ventral components. Case 1 suffered damage to the rolandic region involving the parietal lobe, and thus areas associated with the dorsal visual pathway, and nevertheless was able to perform effectively on object recognition tasks, but not on spatial perception tasks such as maze learning. Case 2 on the other hand, suffered damage to the right temporal lobe, a region associated with the ventral visual pathway, and was able to perform the spatial perception tasks and unable to perform the object recognition tasks.

A study using measurements of regional cerebral blood flow (rCBF) found that during a spatial task involving dot-location matching an increase in blood flow occurred in the lateral occipital and superior parietal cortical regions. Correspondingly, during a face matching task, a task involving featural rather than spatial properties of a stimulus, an increase in blood flow occurred within a zone including the lateral occipital and posterior temporal cortices (Haxby, Grady, Horwitz, Ungerleider, Mishkin, Carson, Herscovitch, Schapiro and Rapoport, 1991).

Haxby and Horwitz (1994) carried out measurements of changes in rCBF by positron emission tomography, while subjects performed dot location and face matching tasks. The spatial task was associated with selective rCBF increases in dorsal occipital, superior parietal, and intraparietal sulcus cortex bilaterally and in dorsal right premotor cortex. In contrast the shape analysis task was associated with selective rCBF increases in the fusiform gyrus in occipital and occipitotemporal cortex bilaterally and in a right prefrontal area in the inferior frontal gyrus. Concurrently, decreases in rCBF were seen during both tasks, in auditory, auditory association, somatosensory, and midcingulate cortices.

Correlations between the values of normalized regional cerebral blood flow (rCBF) within several cortical regions were found in a study by Horwitz, Grady, Haxby, Schapiro, Rapoport, Ungerleider and Mishkin (1992) using a spatial dot location task and a shape analysis face matching task. During both types of tasks significant correlations were found between the changes in rCBF in the right hemisphere in an extrastriate occipital region, and in an inferior occipitotemporal area. The rCBF value in the extrastriate occipital region was similar in both types of tasks. The value of rCBF in the inferior occipitotemporal region however was higher during the face-matching task than during the dot-location matching task, supporting the view that the functional specialization of the inferotemporal region includes a sensitivity to object featural properties.

A study exploring age-related changes of the dorsal-ventral visual system dichotomy was carried out by Grady and Haxby (1992). In order to measure rCBF, positron emission tomographic scans were performed on subjects in two age groups, a young group, with a mean age 27 years, and an old group, with a mean age of 72 years. Subjects were asked to perform an object perception task that involved face matching, and a spatial perception task that involved

dot-location matching. Both age groups showed increases in rCBF values in occipitotemporal cortex during the object perception task, and in the superior parietal cortex during dot-location matching task. Interestingly, the old subject group showed higher levels of rCBF in both regions, during both types of tasks, than did the young subject group.

A cognitive behavioural study similarly found evidence in support of the object-spatial system dichotomy (Tresch and Sinnamon, 1993), by demonstrating selective interference between tasks associated with the two visual systems. A spatial perception task involved having subjects remember the location of a dot, while an object recognition task involved remembering the form of an object. Performance on the spatial memory task was found to be impaired when subjects were engaged in a second spatial task involving movement perception, but not when subjects were asked to perform an object recognition task. Correspondingly, performance on the object recognition task was impaired by a second task in this same category that involved discrimination, but not by a task involving spatial perception.

The findings of lesion, rCBF and cognitive studies all support the notion that human visual systems, like those of non-human primates, can be usefully factored into two components: a dorsal occipito-parietal system supporting spatial perception and a ventral occipito-inferotemporal system supporting feature perception.

The reviewed studies have generally made use of relatively elementary visual perception tasks. A study using a more complex visual task looked at whether a difference in scalp potential could be found corresponding to the difference between a self-generated mental image, and a mental image generated from a prior perception (Petsche, Lacroix, Lindner, Rappelsberger and Schmidt-Henrich, 1992). EEG amplitude and coherence were measured in two conditions, visualization of an abstract concept and visualization of a painting. In one task, subjects were asked to generate a mental image corresponding to an abstract concept, a task expected to involve thinking with images. In a second task subjects were asked to interpret a painting which they had previewed before the EEG recording session. This second task was expected to engage thinking with language. EEG recordings were analyzed in terms of theta (4 to 7.5 Hz), alpha (8 to 12.5 Hz), beta (13 to 18 Hz), beta 2 (18.5 to 24 Hz) and beta 3 (24.5 to 31.5 Hz) frequency bands, using measures of amplitude and coherence. A complex pattern of amplitude and coherence changes was found. In the abstract visualization task, amplitude decreases were found in all bands, at almost all electrode locations. Coherence increases included the left frontal and central regions in the beta 2 band, and right frontal, central, and temporal areas in the beta 3 band. Coherence decreases included the right frontal and temporal areas in the theta and alpha bands. In the painting interpretation task, amplitude decreases were found in the left hemisphere in the beta band, and in posterior regions in the beta 3 band. Coherence increases included the left frontal, central, temporal and parietal areas in the theta

band, the left central area in the alpha band, and left frontal areas in the beta bands. Coherence decreases included the right anterior region in the alpha band, and the right posterior area in the beta ranges. These results indicated that the differences in mental processes associated with self-generated and perceptually-inspired mental images, were reflected most consistently in electrical changes over the frontal regions. The authors concluded that mental imagery involves connections between multiple brain regions, and conclude generally that creative, mental activity appears to be reflected in amplitude and coherence changes of the EEG between multiple cortical regions.

Rappelsberger and Petsche (1988) similarly found that interregional coherence was affected by a visualization task, mental rotation of a cube. Subjects, 13 male and 18 female right-handers, were shown a cube which they were then asked to visualize rotating. A pattern of coherences was found that involved multiple cortical areas. Theta band coherence increases were found in right parietal and right temporo-occipital areas in males, and in the left hemisphere in females. Alpha band coherence decreases were found in the left occipital region in males and in bilateral occipital areas in females. Beta band coherence increases occurred in right parietal and left temporo-occipital areas in males, and in left parietal areas and right temporo-occipital areas in females. Coherence increases were found in all bands between left and right parietal areas in both males and females. The authors propose that degree of coherence between different brain areas may be related to functional couplings between these areas. In support of this view, Bust and Galbraith (1975) found that inter-regional coherence was directly related to the density of connections between the regions.

The results of these studies by Petsche et al. (1992) and Rappelsberger and Petsche (1988) have demonstrated that the dynamics of the neural processes underlying perceptual and cognitive tasks are reflected in changes in the pattern of correlations in the activity between multiple and wide-spread cortical areas.

#### **4.5 Structures and Sensitivities in the Visual System**

The association between behaviour involving object recognition and the occipital and temporal areas is supported by two lines of evidence, ablation studies and single cell recordings. On the one hand, profound deficits on visual discrimination and recognition tasks has been shown to result from ablation of the bilateral anterior IT cortex (e.g., Ungerleider and Mishkin, 1982), and of area V4 (e.g., Schiller and Lee, 1991). On the other hand, populations of cells in the IT cortex have been found to be sensitive to shape discrimination. Cells in the anterior IT cortex have been shown to more sensitive to moderately complex shapes rather than to simple features. Such cells, for example, have been shown to respond to hand-like shapes (Gross,

Rocha-Miranda and Bender, 1972), faces (Young and Yamane, 1992), and complex shapes not corresponding to familiar objects (Fujita, Tanaka, Ito and Cheng, 1992), as well as to be widely invariant to stimulus features such as position, size and color (e.g., Desimone, Albright, Gross and Bruce, 1984).

Young (1995) suggests that an adequate analysis of visual pathways should involve consideration of the high degree of interconnectivity between the parietal and inferotemporal pathways. An analysis by Felleman and Van Essen (1991) for example shows that the visual system contains at least 32 systems interconnected by over 300 pathways. This analysis concentrates on connections and areas in which the connections terminate: the connections are considered to be ascending or descending depending on whether the connections terminate in cell-rich or cell-poor areas respectively. The result is necessarily a unidimensional hierarchical arrangement for the visual system, suggests Young (1992), that does not take into account the possibility of non-hierarchical configurations such as connections between the parietal and occipito-temporal streams. A topological analysis sensitive to such multidimensional configurations conducted by Young and Scannell (1993) found that the occipitotemporal regions comprise a hierarchical organization distinct from the dorsal pathway, a structural distinction consistent with the functional dichotomy proposed by Ungerleider and Mishkin (1982).

At the same time Young (1995) suggests that a more complete analysis would need to consider connections between these two visual pathways, as well as between these paths and other cortical areas. Evidence supporting such connections includes the presence, within the occipito-temporal stream, of a discontinuity between anterior and posterior IT cortex (Tanaka et al., 1991). The posterior IT is characterized by small receptive fields and a sensitivity to simple features, while the anterior IT is marked by relatively large receptive fields, and cells that are preferentially sensitive to more complex features. This discontinuity, Young (1995) proposes, suggests the presence of elaborate callosal connections at the anterior IT. Further, the existence of the many interconnections between the IT and parietal pathways suggests that the dorsal stream may have functions more extensive than simply foveation. On the other hand signaling into the IT regions involves inputs from many areas, such as the limbic system, in which lesions have been shown to result in discrimination and recognition deficits (Mishkin and Appenzeller, 1987).

While cells of the anterior IT cortex have been shown to respond to complex shapes such as hands and faces, a question remains as to what shape elements such cells are sensitive. Fujita et al. (1992) found that cells in the IT cortex are organized into modules, each of which is sensitive to a particular pattern element. These elements are suggested to form a set of basis functions into which more complex visual shapes can be analyzed. Fujita et al. (1992) estimate the number of such modules to be on the order of 1000, with a resulting very large number of

possible combinations available to represent real-world objects. The response of such modules to real-world objects would be a population response, by a collection of cells within these modules. Interestingly, Tanaka et al. (1991) showed that such cells may respond on an exclusive-or basis. Cells which for example respond to a T-shape may not respond when the T is presented in combination with other elements, such as for example within a cross shape. Thus, such cells may respond to the simultaneous occurrence of one feature and the absence of a second feature. It may also be possible, in this example, that a cross-sensitive cell exists that inhibits the response of the T-sensitive cell. Such findings serve to underscore the complexity of recognition processes, and the limits of present knowledge about recognition mechanisms in the visual system.

Such evidence illustrating the complexity of the visual system suggests that a reductionist approach to vision may not be appropriate. Van Essen and DeYoe (1995) propose that the visual system consists of parallel processing streams, with a diverging-converging architecture. In their view, the visual system consists of multiple concurrently-operating streams. These streams, extending the length of the visual system from the retina to the anterior areas of the inferotemporal cortex, may in some locations condense into local networks of relatively fewer paths. At other locations these streams may expand into networks of relatively greater numbers of paths, with multiple feedback paths throughout the system. This topography is suggestive of the convergence zone framework proposed by Damasio and Damasio (1993). The overall perspective then, is that elementary visual cues and their resulting transformed intermediate products interact within the visual system in creating relatively high level results such as motion, form and depth. In their view, cues such as velocity, binocular disparity and orientation all contribute to streams that carry out analyses of motion, form and depth. Van Essen and Deyoe (1995) base their view on anatomical data, such as the finding of approximately 30 distinct areas within the macaque visual system, with on the order of 10 inputs and 10 outputs to each area (Felleman and Van Essen, 1991). Such interconnections, they point out, are generally arranged in reciprocal pairs, arguing against simple hierarchical models of visual processing. On the basis of the finding that the connections within such pairs terminate in different cortical layers, such reciprocal pairs have been identified as corresponding to forward and backward projections. Together, such findings allow the visual areas to be arranged into a network with a high degree of interconnectedness and complexity.

#### **4.6 Summary and Conclusions**

The findings of the studies reviewed in the previous section have underscored the complexity of the visual system. This complexity is demonstrated, for example, by the profuse

feedback projections within the visual system that have been suggested to be involved in higher-level influences on visual perception. These influences include attentional effects (Mishkin, 1993; Gilbert, 1995), a notion that is supported by the findings of attentional influences on the receptive field of neurons in V4 (Moran and Desimone, 1985). In turn, attentional effects have been suggested to be indistinguishable from memory effects within the visual system (Desimone et al., 1995; Miller et al., 1993). The effects of still more wide-ranging feedback influences on the visual system have been found in studies demonstrating the effects of prefrontal cortex on the ventral visual pathway, providing a mechanism for multiple memory systems within the inferotemporal cortex (Miller and Desimone, 1994). Along with feedback projections, horizontal interconnections within the primary visual cortex have been suggested to be involved in other higher level effects in the primary visual areas, such as context dependency and feature integration (Gilbert and Wiesel, 1989; Gilbert, 1995).

The clarity of the analysis of the visual system into a what-where dichotomy itself has been questioned (Young, 1995) on the basis several lines of evidence. These includes the complex neuroanatomy of, and the profuse interconnections between the two pathways (Maunsell and Van Essen, 1983; Felleman and Van Essen, 1991), the finding of extrastriate signals within these areas related to higher level effects such as attention and memory (Maunsell et al., 1991), the complex cell responses found in the inferotemporal cortex (Fujita et al., 1992; Tanaka et al., 1991; Young, 1992), and functional and anatomical discontinuities within the inferotemporal cortex (Tanaka et al., 1991). Such considerations led Maunsell et al. (1991) to suggest that the what/where functional description for the ventral-dorsal dichotomy be replaced with a higher level description involving an object identification-spatial guidance distinction.

Together, such findings suggest that a reductionist approach to the problem of object recognition may not be appropriate in view of the complexity of the interactions within and between the different areas of the visual system. Thus, it may not be possible, when discussing a relatively higher level and more complex visual processes in organisms such as humans, to associate these processes with limited areas of the brain. Rather it appears more appropriate to suggest that in humans visual perception, while involving the inferotemporal cortex along with the primary sensory areas of the occipital cortex, can be expected to critically engage as well other areas, such as central and frontal cortices.

## 5 A Model of Neural Processes in Object Discrimination

The description presented here of the neural events which underlie object discrimination emphasizes the importance of the structure and time-evolution of interregional associations in the process of perception. It is proposed that the complex of neural events associated with visual discrimination can be usefully parsed into a model involving 3 interacting, and concurrently operating, functions:

- (1) feature discrimination, a stimulus driven process in which elementary features of the central representation of a visual image are identified on the basis of characteristics such as lines, forms, edges and colors.
- (2) feature binding, a process in which the elementary visual features that have been identified then undergo binding and transformation resulting in feature ensembles of increased complexity and dimensionality. This process of feature binding is guided by past learning, in the form of memory templates, based for example on the temporal or spatial co-occurrence of visual elements.
- (3) matching or association, a model driven process in which features or feature ensembles over a range of scales of complexity are compared with existing memory templates. At the lowest levels of complexity such matching would occur between the central representation of elementary visual features and hard-wired representations in the visual system, and can therefore be identified with the feature discrimination described in (1). At higher levels, feature ensembles would be matched with correspondingly more complex memory templates that have been created through interactions between the individual and the environment that begin in early stages of ontogenesis and continue throughout the life of the organism. At these higher levels, this matching process can be identified with the feature binding described in (2).

Functions (1) and (2) thus represent points on a continuum rather than essentially unique operations, points that differ essentially in the complexity of the information packet being matched with existing memory templates. Furthermore, the carrying out of these functions might in turn involve multiple simultaneous operations that occur in parallel, a notion that is supported by the parallelism inherent in intra-cortical and inter-cortical signaling pathways. In the lower-level analyses of visual features for example, the retinotopic mapping of visual features within the primary visual areas allows analysis of elementary features in parallel. Demonstrations of

elementary feature analysis within the primary visual areas date from the work of Hubel and Wiesel (1962, 1968) who found that within the primary visual areas information corresponding to visual stimuli is analyzed in terms of relatively local visual elements such as edges and oriented lines. At higher levels, the feature binding process might analogously involve the creation of multiple simultaneous, and to some extent orthogonal transformations, each of which would then be available for comparison with existing memory templates. At all levels of complexity therefore, the comparison or association process might involve the simultaneous, parallel examination of a large population of associations.

According to the model, there is not expected to be a clear distinction between the particular cortical systems associated with these three functions, in terms of the type of processing that is carried out. Specifically, there are not expected to be cortical systems dedicated exclusively to performing on the one hand the functions of elementary feature analysis and feature transformation, and on the other hand memory matching. Rather it is suggested that the memory matching function is an operation that is inherent in the neuronal structures carrying out feature identification and feature or feature ensemble transformation. Thus, the cells and neuronal groups in regions extending from the primary visual areas to the anterior inferotemporal cortex that have shown sensitivities to features ranging in complexity from relatively simple to relatively complex respectively, can be considered, in this sense, to have encoded memories corresponding to such features and feature constellations, over time and as a result of interactions with the environment, or in the case of elementary visual elements, as result of genetic inheritance.

It is proposed then, that image feature analysis and transformation and memory matching occur within the same neuronal region, for any given range of feature complexity, and that the relative extent of the cortical areas that are involved in these operations is dependent on the complexity of the information being processed. The more complex the information being processed the larger the extent of the associated cortical regions. It is proposed that successful discrimination of visually complex depictions of real-world target objects embedded in a visually camouflaging matrix will eventually involve most cortical regions, including occipital, temporal, frontal and central. The three functions that have been proposed as components of the process of object discrimination are therefore suggested to represent functional rather than structural distinctions. Thus, for example, the operation of feature analysis might closely depend on finding a match in memory for a particular feature or group of features. In this broad concept of memory, the particular cells that have been demonstrated to have particular sensitivities, such as to elementary visual features in the primary visual cortex (Hubel and Wiesel, 1962), and to hands, faces, and other complex shapes in the anterior inferotemporal cortex (e.g., Fujita et al., 1992), can be considered to owe this sensitivity to the network of interconnections involving

these cells with some associated cell population. Again, such a network can be thought of as having encoded information about such stimulus features either through developmental experience, or as a result of genetically-guided wiring. All three of the processes proposed to be involved in object discrimination would thus occur within the primary visual areas for relatively simple features, within the primary visual areas together with inferotemporal regions for more complex feature ensembles, and within these areas together with frontal cortical regions for still more complex and higher-dimensional transformations of the image elements.

Importantly, it is proposed that during visual recognition this set of processes does not in general occur as a single-pass sequence, but rather, that visual discrimination of complex real-world objects typically involves multiple iterations of the three functions. Furthermore, when the eventual outcome is successful discrimination, these multiple iterations will involve over time successively larger cortical areas, as increasingly more complex image feature transformations are involved. Thus, it is presently proposed that visual discrimination can involve repeated iterations of the processes of elementary feature identification, feature binding or transformation, and ensemble matching or association.

To illustrate, discrimination of a relatively simple target object from a visual background might first involve identification of elements of the image such as lines, edges and simple shapes in a process that would essentially involve matching these visual elements with existing memory templates. These templates would be encoded as sensitivities to elementary features within early visual cortical regions. Once such relatively low-level matches have occurred, and correspondingly the elementary visual features have been identified, the features would then be available to be bound together by being transformed to form more complex feature ensembles in a following iteration. As a part of this transformation, such feature ensembles would be matched with prior learning, that is, with pre-existing memory templates encoded as sensitivities to more complex feature constellations. The existence of such relatively complex feature sensitivities has been hypothesized by Fujita et al. (1992), who suggest that the inferotemporal cortex may contain regions sensitive to pattern partials that could act as basis functions, combinations of which could then be assembled to create sensitivities to arbitrarily complex visual stimuli (Perrett and Oram, 1993). Successful visual discrimination will occur when this sequence of events results in a match in an iteration corresponding to a level of complexity of the feature ensembles that would be determined by a higher-level influence, such as the goal of the discrimination task. Thus, if the goal is to detect simple line segments the iterations would terminate at a relatively early iteration, corresponding to a relatively dimensionally simple feature ensemble. If however the goal is to detect a more complex shape then the iterations would proceed until a correspondingly more complex feature ensembles were created. If, within a given iteration, a sufficiently accurate match does not occur, then a subsequent iteration should take place. The

results of the first iteration might in some way be able to modify the component processes in the second iteration in such a way that the probability of success on the next iteration is increased. This next iteration could involve a re-analysis of the image for a new set of features that could then be used to create new feature ensembles. Alternatively, the next iteration could involve a re-transformation of the existing features to form new ensembles. Once again, the resulting features and feature ensembles would be available for matching with memory templates. Successful discrimination would result, if, over some number of iterations of this sequence, a match occurs between the feature ensembles and a pre-existing representation in memory.

For the purpose of a more concrete illustration, imagine that the task is to view a monochromatic image composed of short, nonlinear line segments, within which is embedded a disjointed circle formed of similar short and nonlinear line segments. Such an image was used by Ullman and Shashua (1988) in their computational model of object discrimination. When initially viewed, this image would trigger the discrimination process through which the individual line segments are detected and separated from the background. The discriminated line segments would then be available to the transformation and binding operation. Through this operation, the various informational dimensions of the image elements, including elementary dimensions such as shape and length, as well as higher order dimensions such as distributions of shape and length over the image, would be combined through some transformational function to form a feature ensemble. The feature ensemble would then be matched against pre-existing memory templates. This ensemble would represent a particular topographic organization of the elementary image features, such as for example, a complex curve. If the feature ensemble representing this complex curve did not find a matching memory template, a subsequent iteration would take place. This next iteration might involve further transformations of the existing feature ensembles to create more complex feature bundles. If at some stage of these iterations, determined by influences from higher level regions, a successful match with a memory template did not occur, then a following iteration might involve a resetting of this process, with again a re-transformation of the elementary image features into one or more novel feature ensembles. A new round of iterations of the transformational binding function would occur, again creating a sequence of successively more complex feature ensembles for matching with existing memory templates. Target discrimination would occur when the result of these iterations was eventually a successful match between a feature ensemble and an existing memory template.

More extensive versions of this sequence of events will occur when the discrimination task is sufficiently challenging. In the scenario described above, the image features are transformed and bound into a single feature ensemble. Instead, in each iteration a population of ensembles might be created, each of which contains a to some degree orthogonal transformation

of the original elementary features. This population would then be matched in parallel against a corresponding population of memory templates. The transformation products, the feature ensembles, of any one such iteration would then be selected on the basis of the goodness of the match between the ensembles and corresponding memory templates. Those feature ensembles that are able to make a sufficiently good match would then survive to the next iteration, to the next round of transformation and matching. Computational models of such processes exist, generally subsumed under the rubric of genetic algorithms. Approaches based on genetic algorithms, in which a population of transformations is evolved over some number of generations, have been shown to be capable of searching complex problem spaces with an efficiency that can surpass that of more traditional search methods used in the application of distributed network models, such as for example gradient descent methods (Goldberg, 1989; Holland, 1975).

In order for this evolutionary component of the proposed description of neuronal processes underlying object discrimination to make such a description a more economical one, two general conditions must be met. First, the task to be accomplished, in this case the generation of a feature ensemble that is able to find a match with an existing memory template, should be one which presents multiple competing, but only partially correct, solutions. Such tasks are referred to as having a solution space containing multiple local minima. Object discrimination would appear to be a clear example of such a task, in view of the large number of possible topographical configurations that are possible with even a modest set of elementary visual features. That is, even a relatively small number of features such as simple lines, orientation, and colors can interact to form a relatively large number of more complex shapes. Each of these shapes would represent a possible outcome of the transformational process, while only a small subset would succeed in finding a match in memory. Second, the substrate available to carry out the evolutionary program should be capable of massively parallel operations. This second condition would also appear to be well satisfied, by the high degree of interconnectedness both within and between neuronal regions.

What then are some possible mechanisms that could direct the process of transformation through which feature ensembles are created. Three possibilities are presented.

The first possibility is that feature ensemble formation might be guided by internal heuristics that have evolved through environmental interactions. Such a heuristic might for example be based on the temporal or spatial co-occurrence of features.

The second possibility is that the process of creating new transformations might be guided by the results of the previous iteration. Thus the previous iteration would result in some set of indices that would encode the characteristics of the mismatch or mismatches between feature ensembles and prior learning. These indices would then be available to serve as a

correction or feedback signal that would carry information required to correct the previous transformation step, and possibly even the initial image segmentation step.

This correction process is reminiscent of gradient-descent methods, such as the back-propagation algorithm, that have been developed to allow training to occur in neural networks, the computational models of biological neuronal networks. In contrast with the original back-propagation algorithm which required non-local information in order to iteratively update the values of the analogs of synaptic strengths, versions have been developed that not only use only strictly local information in computing a synaptic strength update, but also are more flexible in that the node transfer function, the input-output characteristic of a neuron analog, is made an attribute of each node and thus can vary from node to node (e.g., Fausett, 1990). The question of how well or badly such computational models represent neural processes is not addressed here. Pertinent to the present discussion is that such models nevertheless do succeed in demonstrating flexible learning characteristics on the basis of only local information at each step of the learning process.

What is required of such schemes however is that a correction or feedback signal be able to propagate back from a comparison stage, in which the actual and target outcomes are compared and a measure of the difference between the two is computed, towards intermediate and initial stages of analysis. This requirement for a feedback path in neural systems would appear to be well met by the numerous and ubiquitous back projections between and within all cortical regions. More specifically, in the present context of visual object perception, the occipitotemporal pathway proposed by Ungerleider and Mishkin (1982) as a neural system for object discrimination, contains extensive feedforward and feedback connections between the primary visual areas of the occipital cortex, and the inferotemporal (IT) cortex. Further, as Desimone et al. (1995) point out, based on memory priming studies in monkeys (Miller and Desimone, 1994), extensive connections exist between the prefrontal and IT cortices. Such connections, they suggest, may serve to prime the inferotemporal cortex, and thus function as a higher-level influence on the visual analytic functions carried out in the IT cortices.

A third possible mechanism by which the process of ensemble creation could be directed is energy relaxation. The direction taken by the transformations that create feature ensembles would be one which, over the course of successive iterations, would tend to minimize the level of energy within the associated neuronal system. The mechanism of energy relaxation is discussed in the last Unit of this work. In brief, it is proposed that the iterative process of feature transformation and binding is subserved by a self-organization of the pattern of inter-cortical signaling, a self-organization which has the effect of minimizing both the information content within the neuronal system involved in the discrimination process, that is the information required to describe the corresponding state of the system, and the level of energy within the neuronal

system. It is proposed that the more organized the configuration of inter-cortical signaling, the less energy is required to maintain this configuration, on the straightforward basis that an element of signaling should require some increment of energy to carry out.

The three possible mechanisms proposed as directors of the process of ensemble formation, learned heuristics, corrections by one iteration of the subsequent iteration, and energy relaxation, are not presented as mutually exclusive possibilities. Thus, the nature of the correction to the direction taken by one iteration by the results of a previous iteration can be thought of as decreasing the level of system energy. Both of these mechanisms operating in concert in this way may in turn, at least to some extent, involve a learning process in which components of these mechanisms would develop over the course of ontogenesis as a result of interactions between the individual and the environment. Furthermore, one way in which guidance for the direction taken by the feature transformations might be transferred from one iteration to the next is by means of a genetic algorithmic mechanism as described above, in which a process of selection operates on a population of transformation products, based on the degree of match between a feature ensemble and a corresponding memory template. In turn, the degree of match between a feature ensemble and a memory template might be related to the energy level of the associated neuronal systems. Thus, it is suggested, the level of energy required to maintain an interregional signaling configuration corresponding to some degree of match between an ensemble and a memory template would be inversely proportional to the degree of match. The better the match, the less energy that would be required to maintain the corresponding interregional signaling configuration.

To summarize this description of the neuronal processes underlying visual discrimination, it is proposed that progressively, over a short interval of time prior to discrimination, an increasing proportion of the brain engages in cooperative activity. This activity can be characterized as a cortical self-organization in which the interchange of increasingly complex information occurs in successive iterations that continue until discrimination has been achieved. These iterations involve operations that can be parsed, it is proposed, into a set of functions consisting of (1) feature analysis, a stimulus-driven process, in which elements of the image are identified; (2) feature binding, a process through which visual elements or lower-level feature ensembles are transformed into higher level ensembles; and (3) memory matching, as existing memory templates are successively approximated in terms of the feature ensembles. It is suggested that these 3 functions represent only functional divisions of the neuronal processes that underlie object discrimination, and are not mutually distinct in terms of the underlying neuronal systems or the schedule on which these functions are carried out. This set of functions are carried, it is suggested, out within the same neuronal populations for any given level of complexity of the feature transformation.

Thus, discrimination of a target object from an embedding visual matrix involves a process that may be viewed as successively approximating the results of prior learning represented by existing memory templates, in terms of a sequence of increasingly complex transformations of the elements of the central representation of retinal signals within the primary visual cortex, until the results of a transformation sufficiently well approximate information stored within existing memory templates. The corresponding cortical dynamics can be conceptualized as a process of self-organization of the topological structure of the signaling between multiple cortical regions. The ultimate result of this self-organization is an emergent unitary percept. This description will be referred to as the Cortical Self-Organization (CSO) model of the neuronal processes underlying visual discrimination. The CSO model is schematically diagrammed in Figure 16.1

As a corollary intended to engage practical applications of the results of the present study, it is suggested that the changes in correlation that occur during the visual discrimination process can be summarized using information contained within the two indices, correlation mean and correlation variance. It is proposed that these two measures can be combined to yield an overall summary index which will be referred to as a Discrimination Index. This index is suggested to be a measure of the degree to which a target object has been discriminated from an embedding visual context. Practical application of the Discrimination Index might include its use as a real-time index of a subject's state of attention and level of performance on a task requiring visual perception.

In this section the CSO model has been outlined, as a description of the neuronal events associated with visual discrimination. Particular elements of the CSO model will be tested in the present work, by attempting to find, associated with successful discrimination of a target object, particular characteristics of the EEG. These characteristics are outlined in the following section.

## 5.1 Predictions

A number of statements will be derived from the CSO model that are related to the experimental paradigm to be used in this study, and in particular to the kinds of analyses that will be performed on the resulting data. Such statements will involve characteristics of the scalp potentials that should be observable during visual discrimination. These statements will in turn be used to generate a number of specific predictions.

The successive iterations of the operations of image element identification, transformation and matching, when building towards eventual discrimination, are suggested to involve successively larger extents of cortex. These cortical regions are expected to initially involve the primary visual cortex, along with inferotemporal areas, and to eventually involve larger cortical areas including frontal and central regions. The processes of feature binding and association, dealing with increasingly complex and multi-dimensional transformations of the features of the visual image, need to access memory templates located within correspondingly greater extents of cortex. In doing so, therefore, not only are more and more widely separated cortical areas involved, but also the exchange of information, the signaling, between these areas should become increasingly coordinated as this increasing number of cortical areas participate in common process. This common process, consisting of the analysis and binding of image elements and association of these elements with existing memory templates carried out in an iterated sequence, thus involves an increasing number of cortical areas exchanging increasingly complex information. The earlier events in this process, identification and transformational binding of elementary image elements may, it is suggested, be carried out by distinct cortical areas operating to some extent independently, and each associated with a particular set of dimensions of the data. The process of binding occurring within these separate cortical regions would then be reflected in multiple centers of oscillatory activity whose frequency and phase characteristics would be unique to each region. In terms of scalp potentials the observable effect would be an aperiodic signal at each electrode site with relatively low levels of association between sites. In the later stages of the object discrimination process there would be an increase in the amount of inter-regional signaling which would occur as the separate cortical regions engage in the attempt both, to mutually associate their feature ensembles, and to associate these ensembles with previously learned visual memories. As a consequence, the oscillatory activity associated with the separate cortical areas would become increasingly synchronized. The result, it is suggested, should be an increasing level of association between the scalp potentials measured over these cortical areas. Thus, it is suggested, the effect of the inter-regional signaling is to mutually synchronize the oscillatory activity occurring within separate cortical regions.

To summarize, the following neural events are proposed to occur over a short interval of time preceding the moment of discrimination. When a visual image is initially attended to, cortical areas within the primary visual cortex such as V1, that have been shown to be involved in elementary feature identification and discrimination, are expected to be active. Since a clear structural division between systems performing feature discrimination and feature transformation is not expected to exist, the primary visual areas engaged in feature analysis are also expected to be involved in relatively local, low-level transformations of the data, into relatively low-dimensional ensembles. At the same time such regions might be expected to be engaged in a relatively limited amount of mutual signaling or information interchange as the competition for cortical communications resources favors intra-regional rather than inter-regional signaling (Thatcher, Krause, and Hrybyk, 1986). The relatively independent nature of the activity within these cortical regions should result in mutually asynchronous field potentials across the different cortical regions, and a correspondingly low level of association between the scalp electrical activity over these regions. Over time, and as the moment of discrimination approaches, the level of inter-regional signaling is proposed to increase. Multiple cortical regions engage in an increasing level of information exchange, corresponding to the increasingly complex, high-dimensional feature transformations that are being created, and that involving increasingly greater neuronal populations within the respective regions. The result is an escalating level of mutual synchronization of the oscillatory activity occurring across these regions. Correspondingly, it is expected that potentials measured at multiple points on the scalp will be characterized by an increasing level of mutual association.

Thus, imminently successful visual discrimination is distinguished from eventually unsuccessful image analysis by the involvement of increasingly larger cortical areas as the ultimately successful transformation or feature ensemble is able to find a match in terms of an increasing number of information dimensions, and hence in terms of the information stored in increasingly large and more numerous cortical areas associated with these higher-dimensional transformations of the data. Through the mechanism of inter-regional signaling, the synchronous involvement of these ever larger and more numerous cortical areas then results in an ever increasing degree of correlation between the electrical field potentials associated with these regions, in some interval of time preceding discrimination. A very general analogy might be a network of coupled oscillators, with each oscillator representing the activity of one cortical area carrying out the process of feature binding through transformation. The coupling strength between the oscillators, representing the level of inter-regional signaling, would thus increase as the moment of discrimination approaches.

These elaborations of the CSO model allow a number of predictions to be made in the context of the paradigm used in the present study. It is suggested that, when the perceptual task

is to discover a camouflaged object embedded within a complex image, the associated EEG will show a number of characteristics that should be observable during a short interval of time preceding the moment of discrimination.

1. The magnitude of intercorrelations between all cortical areas is expected to increase with time during an interval preceding the moment of discrimination. Initially, these intercorrelations are expected to increase between areas associated with relatively low and moderate level visual feature analysis and transformation, the occipital and inferotemporal areas. Subsequently, these intercorrelations are expected to involve larger cortical extents, as the feature analysis and transformation processes result in increasingly complex feature ensembles that then require matching with prior learning within correspondingly larger areas of cortex.
2. Averaged over time, the magnitude of intercorrelations is expected to vary approximately in inverse proportion to the physical distance between the corresponding electrodes.
3. Intercorrelations between cortical regions that are widely spatially separated are expected to increase more than correlations between cortical areas that are closer together. A consequence is that the variability in the magnitude of these intercorrelations is expected to decrease.

These three characteristics are motivated by features of the CSO model, as well as by the results of previous studies. The first of these characteristics, the increase in the magnitude of intercorrelations, has been suggested to be a consequence of the phenomenon that multiple regions of the brain engage in an increasing degree of reciprocal signaling as the moment of discrimination approaches. In alternative terms, during the interval preceding successful discrimination there should be an increasing rate of information interchange between cortical regions. The suggestion of the initial involvement of primary visual and inferotemporal regions is motivated in part by the results of lesion studies, primarily using macaque monkeys, that have defined the functional specializations of these areas in the process of object discrimination (e.g., Damasio, Damasio and Tranel, 1990; DeYoe and Van Essen 1988; Gilbert, 1992; Ungerleider and Mishkin, 1982). The suggestion that frontal areas are involved in the later stages of the object perception process is motivated in part by the results of studies demonstrating attention-related effects within the visual system as a result of communication between anterior inferotemporal and prefrontal areas (Miller and Desimone, 1994), as well as the results of studies of EEG coherence during visual image interpretation (Petsche et al., 1992).

The second characteristic, while not directly addressed by the CSO model, is nevertheless included since this feature is expected to be a prominent feature of the observed EEG associations. The inverse relationship between correlation and inter-regional distance, is predicted on the basis of studies of the neuro-anatomy of the cortex. Braitenberg (1978) has pointed out that the cortex contains between 10 and 100 times as many short-axoned neurons that arborize in the immediate vicinity of the cell body, as long-axoned pyramidal cells whose axons reach lengths of upwards of several centimeters. This greater density of short connections within the cortex relative to longer connections, and the decreasing density of such connections with distance (Thatcher et al., 1986), should mean that correlations between electrode sites should decrease with increasing distance between the sites.

The third characteristic, the relatively larger increase in correlations between widely-spaced regions relative to closely-spaced regions, is suggested to be a result of the importance, for visual discrimination, of communication not only within, but also between cortical regions. Communication, and hence the level of coordinated activity, within and between closely spaced cortical regions is expected to be significant in the initial stages of the visual discrimination process, reflecting on-going analysis of perceptual elements of the central representation of the visual image by relatively local cortical regions. Over the duration of the discrimination process it is expected that the degree of coupling between these closely-spaced regions will not increase substantially. In contrast, communication, and hence the level of coordinated activity, among more widely-spaced cortical regions is expected to be relatively low at the start of the recording epoch, since visual elements have yet to undergo binding into more complex feature ensembles, and association with the correspondingly complex memory templates. At the moment of discrimination, image elements have been bound into feature ensembles that have been successfully matched with memory templates. This feature binding and memory matching are proposed to be indicated by an increase over time of the magnitude of correlations between relatively widely-separated cortical regions. The level of association between these regions is therefore proposed to increase substantially in order for discrimination to occur. As discrimination approaches, there should be a corresponding decrease in variability among interregional associations, since associations between closely-spaced regions change relatively little while associations between widely-spaced regions increase.

While these predictions all deal with the topological aspects of the neuronal processes associated with visual discrimination, the study of correlations addresses a complementary issue, the issue of the dynamics of the interregional signaling associated with discrimination. Implicit in the search for correlates of visual discrimination in an analysis of between-signal associations is the idea that the pattern of associations preceding discrimination is based on oscillatory components of the recorded potentials. These oscillatory components, it is

suggested, reflect the iterated nature of the neuronal processes associated with discrimination, that is, the successive iterations of the processes of feature analysis, transformation and memory matching that were outlined in the CSO model.

Reber (1985) suggests as a definition of perception, that it consists of "those processes that give coherence and unity to sensory input" (p. 527). In this study subjects will be presented with stimuli designed to encourage such perceptual unification to occur. Subjects will be shown high-contrast images containing target objects embedded within a visually-complex background. Subjects are expected to generally initially interpret such images only as patterns of random shapes. Over a short interval of time, and for at least a subset of the set of images presented, discrimination of the target objects from the background should occur. The neural processes associated with this discrimination should be reflected in EEG signals recorded concurrently with the stimulus presentation, particularly when these signals are analyzed in terms of features that reflect the proposed underlying inter-regional synchronization. These analyses will include cross-correlation, coherence and mutual information.

## II Linear Analyses

## **6 Cross-correlation analysis**

### **6.1 Introduction**

This experiment will try to find evidence in terms of cross-correlations related to the predictions that were made on the basis of the CSO model of object discrimination. First, the magnitude of correlations is expected to increase over time as the moment of discrimination approaches. Second, the size of correlations is expected to be approximately inversely proportional to inter-electrode distance. Third, correlations between distantly-spaced electrode sites are expected to increase more than correlations between closely-spaced sites. These predictions can be recast into the form of questions regarding the temporal and spatial features of the correlational structure of EEG signals recorded during a visual discrimination event. The primary question to be answered is, how do between-channel correlations change as a function of time, both before and after the moment of discrimination. A secondary question is, how do these correlations change as a function of distance between electrode sites.

An experiment has been designed in order to try to answer these questions. In overview, the pattern of correlations between signals recorded from pairs of electrode sites will be compared for two conditions. In the first, picture condition, subjects will be attempting to discriminate a target object that is visually embedded in a complex visual background. In the second, control condition, subjects will be looking at a fixation point on a neutral background. In the first case subjects will signal discrimination with an eye-blink. In the second condition subjects will blink at a time of their choosing. In both cases signals will be recorded and analyzed from two intervals of time. The first interval will be the 1 second interval preceding the eye-blink. The second interval will be the 1 second interval following the cessation of eye movements after the blink.

This procedure has been designed to provide subjects with the opportunity to engage in the neural events that are involved in the process of visual discrimination. Thus, the experimental stimuli consist of depictions of objects that are expected to be relatively familiar, such as birds and animals, and which if presented on their own, subjects could reasonably be expected to recognize immediately. These target objects, however, are embedded in a visual background that is intended to delay the onset of discrimination, by requiring subjects to attempt to discriminate between the background and the target objects. Until such target-background discrimination has taken place, the embedded targets will not be perceived or identified as discrete objects.

During the time that subjects are attending to the stimulus image, and when the outcome is successful discrimination, subjects are expected to automatically, that is, without conscious

awareness of the process, organize the visual elements which together comprise the depiction of the target object and the visual context in such a way that the target becomes salient and distinct from the context. It is this process that is the focus of the present investigation. It is suggested, on the basis of pilot work, that significant portions of this process will occur over a short interval of time, on the order of a second, preceding the actual moment at which the target object is perceived. Finally, it is assumed that whatever neurophysiological events are involved in this object discrimination process, such events will to a significant degree be reflected in the electrophysiological activity that will be measured across the scalp, to the extent scalp measurements are able to access the electrical activity of the relevant neural generating structures. Correspondingly, it is acknowledged that the activity of cortical processes will be represented in the scalp electrical measurements to a greater extent than that of subcortical processes.

In summary, the analysis of cross-correlations in the present section is intended to answer the following questions. First, is there a significant increase with time in the magnitude of correlation? Second, is there a significant decrease in the magnitude of correlation with increasing distance? Third, is there a greater increase in correlation between more widely-spaced cortical regions than between closely-spaced regions?

## **6.2 Method**

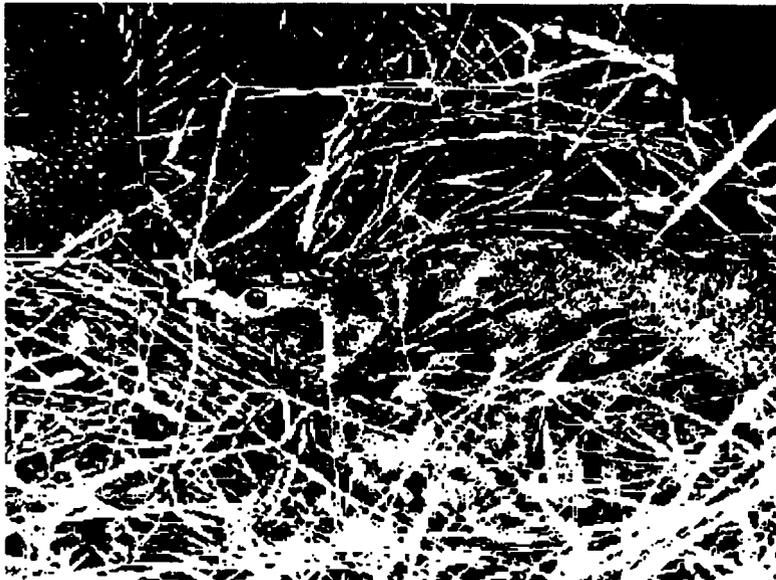
### **6.2.1 Subjects**

The subjects in this study consisted of 3 female and 3 male university undergraduates or graduates, ranging in age from 23 to 47, with a mean age of 33.5, and a standard deviation of 10.8. All subjects are right-handed with English as their first language. Subjects have no known neurological disorders. Two of the subjects (1 male and 1 female) were paid for their participation.

### **6.2.2 Materials and Procedure**

The stimuli are comprised of a set of 31 images, each of which depicts a target object, generally an animal or a bird, embedded within a complex visual matrix that is intended to have the effect of camouflaging the target. The example presented in Figure 6.1 shows a bird sitting in grass, with the shading and markings on the bird matching the pattern of the grass. Each image was constructed by first scanning a photograph of the appropriate scene, and then converting the scanned image into a monochromatic version with the original picture tones converted to 2 values, black and white. The resolution of the monochromatic image is 100 pixels

per inch. The result is a high-contrast version of the original picture. Scanning and conversion were done using PhotoStyler 2.0 © by Aldus Publishing, an image processing program. The result for each picture was a bit-mapped file that was then stored on disk. In total, 31 stimulus pictures were constructed in this way. Appendix 2 shows the remaining 30 stimulus pictures.



**Figure 6.1 Typical stimulus picture**

This picture is one of the stimulus images presented to subjects in the picture condition. It shows a bird positioned approximately in the center of the picture, sitting in dried grass. The original photograph was scanned to convert it into a bit-mapped file, and the file was then converted into a high-contrast 2-color black and white image (picture adapted from Frisch, 1973).

There were two conditions in this experiment, a picture condition and a control condition. In both conditions, subjects were seated in a darkened room facing a computer monitor on which the stimuli were presented. The distance between the subject's head and the display screen was approximately 60 cm, and the size of the stimulus picture on the screen was adjusted to 5 cm vertically, corresponding to a visual angle of approximately 4.8 degrees, by 6.7 cm horizontally, for a visual angle of approximately 6.4 degrees.

In the picture condition subjects were instructed to maintain focus on a fixation point that remained constantly in view in the center of the display. Subjects were instructed to blink when they felt that they had discriminated an object in the image. Subjects were presented with the entire sequence of 31 images, one image at a time. Each image was shown continuously for 8 seconds, with an inter-trial interval that varied randomly between 10 and 20 seconds. Subjects were instructed to keep looking at the discriminated object after they had blinked. The difficulty-of-discrimination level varied from very easy to very difficult, estimated on the basis of subject debriefings. Subjects were able to discriminate the easy-to-recognize objects within

approximately one second, while they were generally not able to discriminate the difficult objects at all. Subjects were presented with the entire set of 31 images twice in a continuous sequence, with only a normal inter-trial interval between sets. EEG recordings were made during the 8 second interval that the stimulus images were presented. The start of the recording was synchronized with image onset.

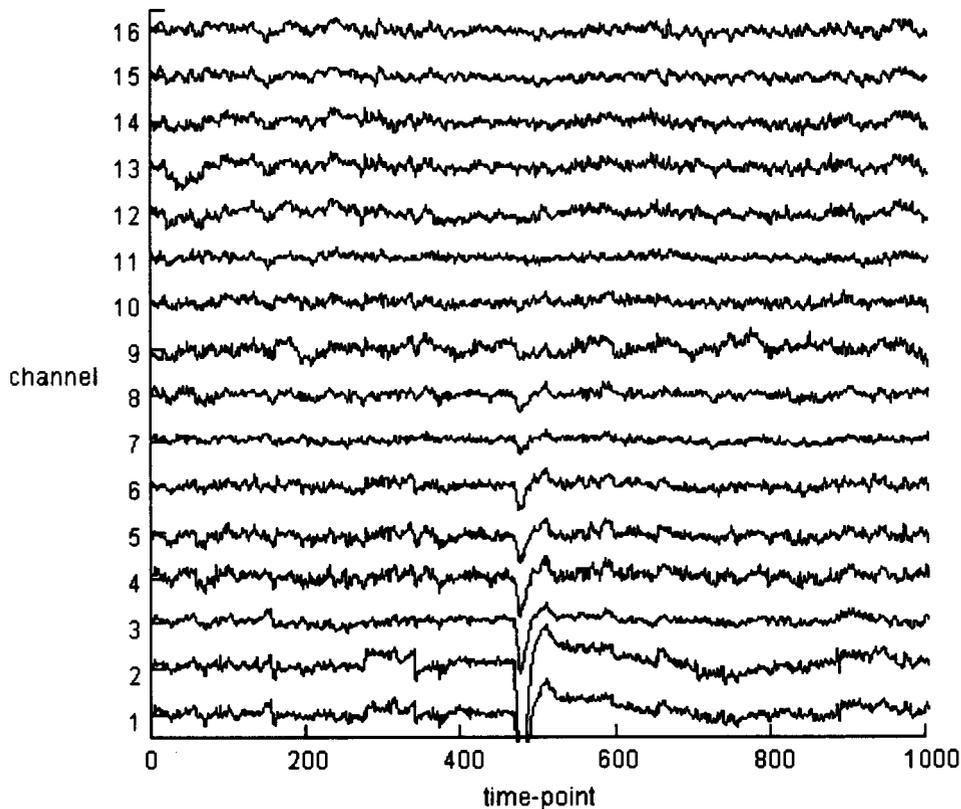
Following the presentation of the two sets of 31 stimulus images, a total of 20 control trials were recorded. No attempt was made to counterbalance the presentation schedule of control and picture trials across subjects. In the control condition, a diffusing screen was fixed over the monitor screen. This screen, consisting of a sheet of white paper large enough to cover the screen, had a fixation point marked on it in the same visual position as the fixation point on the monitor display. The diffuser did not allow any underlying shapes to be recognizable, but showed only a general brightening when the monitor display brightened. The length of this brightening interval was fixed at 8 seconds, the same duration for which the stimulus pictures were visible in the picture condition. Overall screen illumination was adjusted to create an approximately constant brightness level in the two conditions. In this control condition subjects were instructed to blink at any time of their choosing, but within the time during which the diffuser brightened. EEG recordings were made during this 8 second brightening interval. The start of the recording was synchronized with the onset of the 8 second interval. Using this arrangement, 20 control trials were recorded. The entire recording session of picture and control trials required approximately 45 minutes.

A total of 20 channels in the 1020 system were recorded. This electrode topography is shown in Figure 6.6. Electrode Fpz was used as the ground connection, and linked-ears were used as the reference. Recordings were made using the EEG amplifiers in a Nihon-Khodon model EEG-4217 EEG recording station. Amplifier outputs were routed to data collection software, Brainwave V1.1 ® written by Procet Engineering, through a National Instruments model ATM1064-F analog to digital conversion system. In the Nihon-Khodon machine, the high-pass filter setting was 3 Hz, the low-pass filter setting was 70 Hz (-3 dB points), and a 60 Hz notch-filter was used. The EEG data were digitized at 128 samples per second and each trial stored on disk as a separate ASCII file. Figure 6.3 shows a typical record of 8 seconds of EEG recorded at electrode Fp1 from a subject in the picture condition. Figure 6.4 shows a single channel from this ensemble, electrode Fp1, illustrating the typical EEG features associated with an eye-blink.

	Fp1	Fpz	Fp2	
F7	F3	Fz	F4	F8
T3	C3	Cz	C4	T4
T5	P3	Pz	P4	T6
	O1		O2	

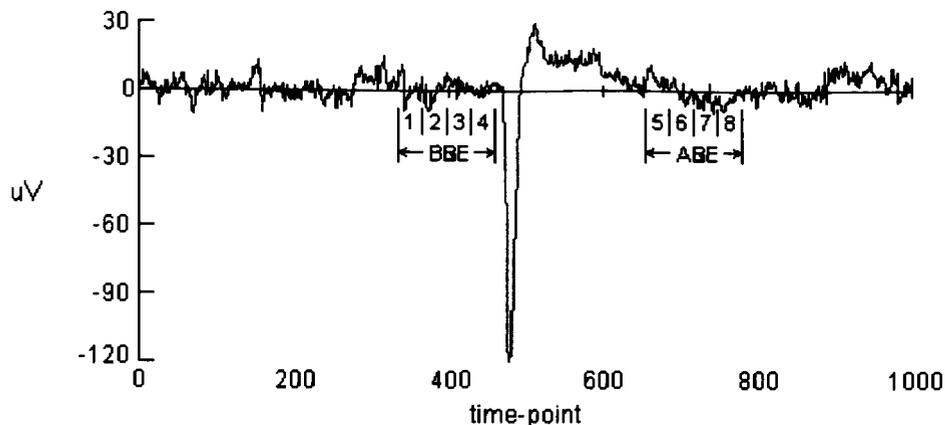
**Figure 6.2 Electrode topography**

Data were recorded from these 20 electrodes which were located using the international 1020 system. Electrode Fpz was used as the ground reference, and linked ears were used as the signal reference.



**Figure 6.3 Typical wave-form ensemble**

These waveforms were recorded from subject 7 in the picture condition. The eye blink can be seen at approximately the center of the interval. The horizontal scale represents a time-interval of 8 seconds, digitized at 128 points per second, resulting in a total of 1024 points.



**Figure 6.4 Typical eye-blink wave-form.**

Typical eye-blink signal (measured at electrode Fp1) from subject 7 in the picture condition, consists of a negative-spike, followed by a slow positive wave. The amplitude of the positive wave is typically minimal 1.5 seconds after the start of the spike. Thus, the eye-blink is defined as 1.5 seconds in duration. The 1 sec. intervals selected for analysis, labeled BBE (before-blink epoch) and ABE (after-blink epoch) are each subdivided into 4, 0.25 sec. time-windows. The end of the BBE, is 20 time points (0.156 seconds) before the start of the spike. The ABE starts at the end of the eye-blink. Sampling rate is 128 points/sec.

Data were recorded in 6 separate sessions, one session for each of 6 subjects. Data from one subject (6) were rejected after debriefing revealed that the subject had misunderstood and therefore not followed the instruction to blink only on discriminating the camouflaged object. The recording sessions were preceded by 3 pilot sessions, the aim of which had been to provide initial data to aid in hypothesis construction, as well as to perfect the recording paradigm and to verify the integrity of the recording equipment. Subject 5 had participated in one of the pilot sessions. In each experimental session there were 62 picture trials and 20 control trials. Trials were rejected if they were found to contain artifacts such as relatively sharp changes in potential that might be related to eye movements or general body movements, or if the start of the blink occurs too close to the beginning or too close to the end of the record. Specifically, trials were rejected if the blink occurred within 1.187 seconds of the start of the record, or within 2.5 seconds of the end of the record.

For both control and picture conditions, two intervals within the data record were analyzed. The first interval is referred to as the before-blink epoch (BBE). The end of the BBE is defined as 24 data points (187 milliseconds) before the start of the eye-blink. The start of the eye-blink was defined as a negative going voltage change with a rate of change of at least 50 microvolts in 31 milliseconds. The value of 24 data points was chosen with the intention of accounting for the reaction time between the discrimination event and the start of the blink. That is, the 24 data points corresponding to 187 ms were intended to exclude from the subsequent

analysis potentials related to preparation for the eye blink. The BBE, defined as the 1 second interval preceding the end-point of the BBE, is divided into 4 time-windows, each 32 data points (0.25 seconds) long. These time windows are labeled 1 through 4, and are shown in Figure 6.4. The length of a time-window was chosen as a compromise between adequate temporal resolution within a recording epoch on the one hand, and getting enough data points in each window for robust analytical results on the other hand. The primary type of analysis that will be performed on this data is correlation. If there are too few data points in a window then the between-channel correlations will be dominated by machine noise, at frequencies for which the period is much less than the number of points in the window. The lowest frequency that will be examined is 2 Hz. The window width of 0.25 seconds corresponds to a half cycle of a 2 Hz signal. While this window width should still allow a 2 Hz signal to be analyzed, it is estimated that a narrower window would probably not be usable. Trials are rejected if they do not contain at least 4 such time-windows. This leads to the criterion that the blink must occur at least 1.187 seconds after the start of the record (4 windows of 0.25 sec. each, plus 187 milliseconds for blink reaction time). The length of the before-blink epoch itself, 1 second, was chosen as a compromise between on the one hand including a sufficient number of time windows to get a good picture of what is happening before the blink, and on the other hand, not limiting the number of cases that would be available to be analyzed. A longer before-blink epoch would have precluded a greater number of cases from being analyzed, those cases in which the blink occurred too close to the start of the record.

The second interval that will be analyzed is referred to as the after-blink epoch (ABE). The ABE starts 1.5 seconds after the start of the blink, and has a duration of 1 second, and correspondingly 4 time-windows labeled 5 through 8, and shown in Figure 6.4. The start of the ABE was defined to be 1.5 seconds following the start of the blink. This definition was made on the basis of observations of multiple records of the blink waveform which indicated that significant electrical activity associated with the eye-blink, a low-frequency rolling wave, was generally minimal after 1.5 seconds following the start of the negative-going voltage spike.

In sum, each trial was qualified on the basis of being free from artifacts, and containing at least 128 data points, equivalent to 1 second, both before and after the blink. From trials qualified in this way, the three midline channels, Fz, Cz, and Pz, were removed. The motive for this is that it seems reasonable that if there are no generating brain structures underlying these channels then signals from these channels would represent the summation of signals from adjacent channels. As such, these midline channels would necessarily be correlated with adjacent channels, and would therefore carry no independent information. Subsequent analysis was therefore carried out on the remaining 16 channels of the international 1020 system: Fp1, Fp2, F7, F3, F4, F8, T3, C3, C4, T4, T5, P3, P4, T6, O1 and O2.

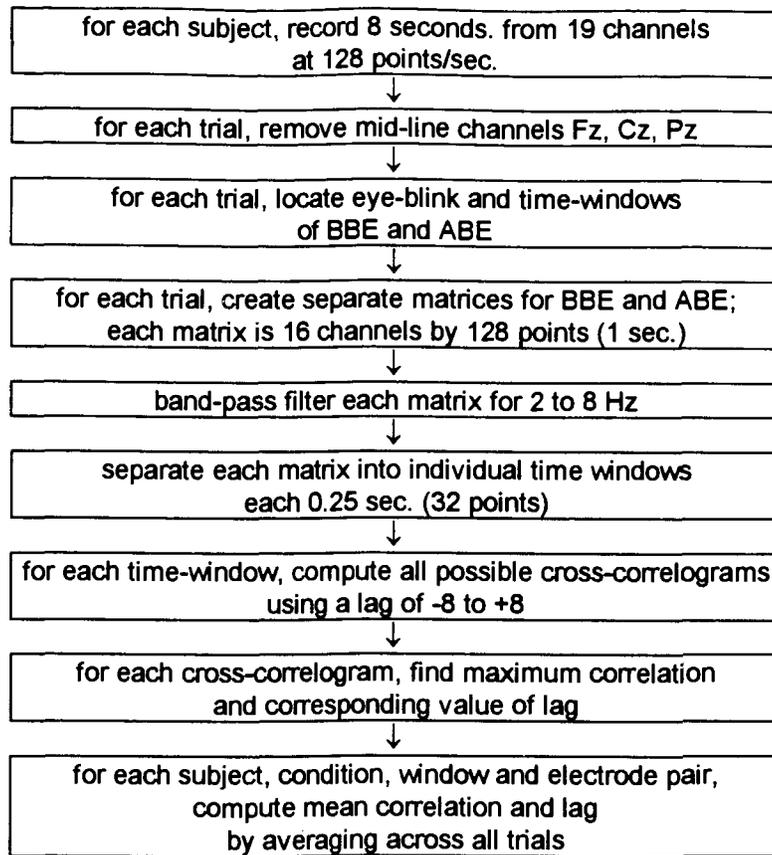
Next, the 1 second intervals of the BBE and ABE from each qualified trial were band-pass filtered to extract the 2 to 8 Hz frequency components. The lower frequency limit of 2 Hz was chosen in order to exclude from further analysis the effects of any low-frequency artifacts that might yet be present in the recordings, and which might be generated by sources such as slow eye or facial muscle movements, or by respiration. Additionally, the length of the time-window, 32 data points, is equal to the period of a 4 Hz signal. In the 32 point time window there will be one-half cycle of a 2 Hz signal, and proportionately smaller fractions of a cycle of lower frequencies. These small fractions of a cycle for such lower frequencies would convey little information about the behaviour, such as correlation, of these low frequencies. In any event, there would be relatively little signal at frequencies below 2 Hz since the high-pass filter setting used for recording the data was set at 3 Hz (-3 dB).

The higher frequency bands, alpha, beta and gamma, were not analyzed because of the findings of the pilot experiments, and because of the results of the numerical experiment outlined in Appendix 1. Pilot analyses on these higher frequency bands indicated that in fact there was little or no effect of picture condition on correlations when signals were filtered to extract alpha, beta and gamma band components. Furthermore, estimates of signal to noise ratio within the beta and gamma bands, as described in Appendix 1, indicated that the magnitude of correlation would be affected by the level of noise within these bands. The findings described in Appendix 1 are that correlations computed from data filtered to extract higher frequency components such as the beta and gamma band, are likely to be significantly affected by the signal-to-noise ratio within the recordings. Noise is defined here as electrical signals that are not generated by the subject, but rather originate either within the recording equipment or are picked up by capacitive and inductive coupling between the subject and the electrode wires connecting the subject with the recording equipment, and extraneous electrical wiring and equipment within the vicinity of the subject. If this noise has a time-varying component with a time-scale comparable to that of the data recording itself, then the magnitude of correlations computed from such data will include a component proportional to the noise level. This time-varying component in the computed correlations may mask or at least modulate any effect of the experimental manipulations. A conclusion of the results of Appendix 1 is that the effects, on magnitude of intercorrelation, of time-varying components of noise may not be separable from the effects of the experimental manipulation, at frequencies higher than the alpha band. Consistent with this finding, it has been pointed out gamma band signals may not be easily detectable on the scalp because first, such signals are greatly attenuated due to volume conduction through the skull, and second, these signals are confounded with scalp muscle activity (Bressler, 1990).

The theta frequency band, as defined in the present study at 2 to 8 Hz, was not subdivided into smaller segments because of a phenomenon associated with narrow-band

filtering: narrow band-pass filtering can create spurious signals when the data contain rapid changes or spikes, even if such changes are of low amplitude: the spike events tend to become smeared out in time in inverse proportion to the width of the filter pass-band, generating spurious periodic signals over the length of the time-interval. Such spurious signals may have relatively less effect if the resulting data were being analyzed for amplitude or power, but they would have a significant effect if correlational analysis was used: these spurious signals would have the effect of essentially adding noise to the data and, as a result, affecting the magnitude of correlations computed from the data.

The next step of the analysis was to compute cross-correlograms for each condition, for each subject, for each trial, for each time window of the BBE and ABE, and for each of the 120 possible pairs of time series using the recordings from the 16 electrodes. Each correlogram was computed by varying the lag between each pair of time series over the range of -8 to +8. One unit of lag corresponds to the inverse of the sampling rate, 7.8125 milliseconds. This range of lags was chosen on the basis of the results of analyses conducted on the pilot data. For each value of lag the correlation between the pair of time series is computed. An average cross correlogram was then computed for each subject, for each condition, for each epoch, for each window, and for each channel pair. This was done by averaging over the cross-correlograms for individual cases for a subject. For each of these mean correlograms, 2 values were determined, the maximum value of correlation, and the value of lag at which the maximum correlation occurred. These steps are shown in the following diagram:



For each electrode pair an estimate was computed of the physical distance between the pair of electrodes. A simple Euclidean distance was computed based on a flat scalp geometry, resulting in a distance measure in arbitrary units. This flat scalp geometry consists of a 5 by 5 matrix upon which electrode sites are positioned. Each electrode can thus be assigned a column number and a row number to indicate its location on this matrix. Distance is computed by taking the square root of the sum of the squared difference between the column numbers and the squared difference between the row numbers. The minimum distance of 1 corresponds to adjacent electrodes. The maximum distance between 2 electrodes is approximately 4.47 units. This flat scalp model is sufficient for the purposes of the present analysis, since the only variable that is computed from the distance is the ordinal ranking of electrode pairs in terms of distance between members of a pair.

The cross-correlation computations as well as all supporting functions were carried out using the data analysis program Simulnet™ version 2.3.

### 6.2.3 Analyses

The predictions that correlation would increase with time during the BBE, and decrease with increasing between-electrode distance, were tested using an analysis of variance of the intercorrelations. Correlation was analyzed, using a 2-factor within subjects analysis of variance. The two factors were time, and distance between electrodes. The variable time had 4 levels corresponding to the 4 time windows within each of the recording epochs. The variable distance had 120 levels corresponding to the 120 intercorrelations ranked according to between-electrode distance. The prediction, that over the interval preceding discrimination correlations between distantly spaced electrodes would increase more than correlations between more closely spaced electrodes, was tested by conducting separate 1-way within subject analyses of variance on the 12 electrode pairs separated by the shortest distances and the 12 electrode pairs separated by the longest distances.

Although no predictions were made at the outset regarding the behaviour of the value of lag, an exploratory analysis of lag was conducted. Lag was analyzed using a procedure identical to that employed for correlation, using a 2-way within-subjects analysis of variance. As for correlation, the two factors were time and distance, defined as for the analysis of cross-correlations.

A similar analysis of variance was conducted for recordings in what will be termed the no-blink condition. The no-blink condition consists of the subset of trials in the picture condition in which subject failed to discriminate the camouflaged target object embedded in each stimulus image during the 8 second recording interval. It is predicted that the results in this condition will be very similar to those obtained in the control condition, in which subjects were instructed to blink at will while looking at a fixation point on a neutral screen. If the results in these two conditions are similar then the validity of the control condition will be supported.

For each of these analysis an estimate of effect size was computed. The measure used was omega-squared, as suggested by Keppel (1991, p. 63). The values of effect size may be interpreted using guidelines suggested by Cohen (1977):

Effect Size	Interpretation
0.01	small effect
0.06	moderate effect
> 0.15	large effect

## 6.3 Results

### 6.3.1 Analysis of Correlation

The three predictions made on the basis of the CSO model were confirmed by the results. Confirming the first prediction, there was one significant effect of time, in the picture condition in the BBE ( $F = 3.16$ ,  $p = 0.027$ , effect size = 0.0003). Correlation increased over the duration of the BBE, from a value of 0.715 to 0.789. In contrast, in the control condition correlations remained relatively constant over the BBE, decreasing slightly and non-significantly from 0.740 to 0.719 ( $F < 1$ ). Table 6.2 lists the mean correlations for each time window, averaged over all distances, that is, over all 120 possible pairwise combinations of electrodes. These mean correlations as a function of time-window are graphed in Figure 6.5a.

Confirming the second prediction, there were significant effects of distance in both the BBE and ABE, and in both the picture condition (BBE:  $F = 20.9$ ,  $p < 0.0001$ , effect size = 0.11; ABE:  $F = 16.7$ ,  $p < 0.0001$ , effect size = 0.085) and in the control condition (BBE:  $F = 27.3$ ,  $p < 0.0001$ , effect size = 0.11; ABE:  $F = 25.5$ ,  $p < 0.0001$ , effect size = 0.10), with correlation decreasing with increasing distance in all cases. Correlation decreased from approximately 0.85 for adjacent electrodes to approximately 0.6 for electrode pairs spaced furthest apart. These mean correlations as a function of inter-electrode distance are graphed in Figure 6.6a for the before-blink epoch and Figure 6.6b for the after-blink epoch. Table 6.1 shows the results of the analysis of variance, listing the values of  $F$  along with the corresponding values of probability and effect size.

Confirming the third prediction, correlations between the 12 most closely-spaced electrode pairs increased from 0.836 to 0.856 over the BBE ( $F = 1.19$ ,  $p = 0.317$ , effect size = 0.0003). Over this same interval correlations between the 12 most distantly-spaced electrode pairs increased from 0.577 to 0.698 over the BBE ( $F = 2.89$ ,  $p = 0.038$ , effect size = 0.003). Since these secondary analyses of variance were conducted on only subsets of the data upon which the original analysis of variance was conducted, it is not expected that the significance probabilities may need the corrections normally required when multiple tests of significance are conducted on the same data. These results are shown in Table 6.3, and graphed in Figure 6.7a for the picture condition and Figure 6.7b for the control condition.

**Table 6.1 Results of analysis of variance of intercorrelation**

The table shows the results of a two-way within subjects analysis of variance of cross-correlogram maxima. Significant effects of distance occur for all conditions and epochs. A significant effect of time occurs in the before-blink epoch in the picture condition only.

Condition	Epoch	Item	Time	Distance	T x D
Picture	BBE	F	3.16	20.9	1.0
		p Effect	0.027	< 0.0001	0.53
	ABE	F	< 1	16.7	< 1
		p Effect		< 0.0001	0.085
Control	BBE	F	< 1	27.3	< 1
		p Effect		< 0.0001	0.11
	ABE	F	1.27	25.5	< 1
		p Effect	0.29	< 0.0001	0.10

**Table 6.2 Mean intercorrelations**

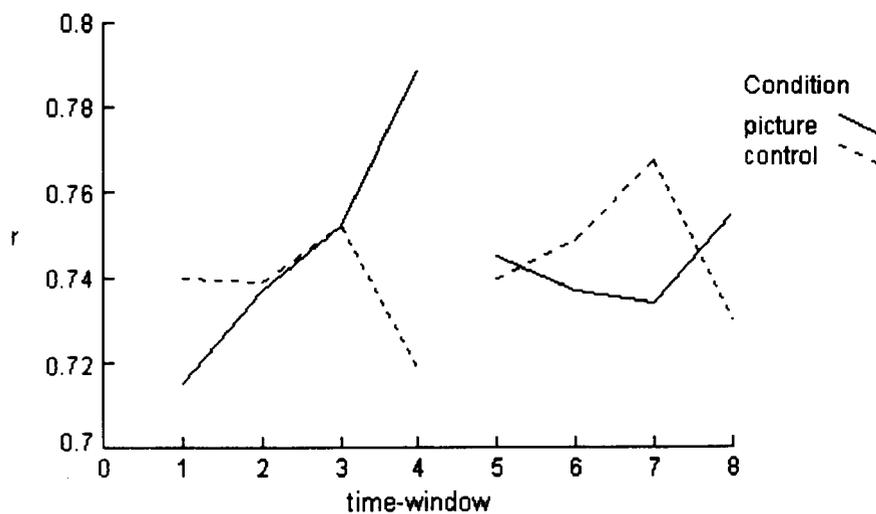
The table shows cross-correlation maxima averaged over all trials for both conditions, and for each time window.

Epoch	Window	Control	Picture
BBE	1	0.740	0.715
	2	0.739	0.737
	3	0.752	0.752
	4	0.719	0.789
ABE	5	0.740	0.745
	6	0.749	0.737
	7	0.768	0.734
	8	0.730	0.755

**Table 6.3 Short vs. long intercorrelations**

The table shows cross-correlation maxima, averaged over all trials, for the 12 shortest inter-electrode distances, labeled Short, and the 12 longest inter-electrode distances, labeled Long. Only long distance correlation increase significantly from windows 1 to 4.

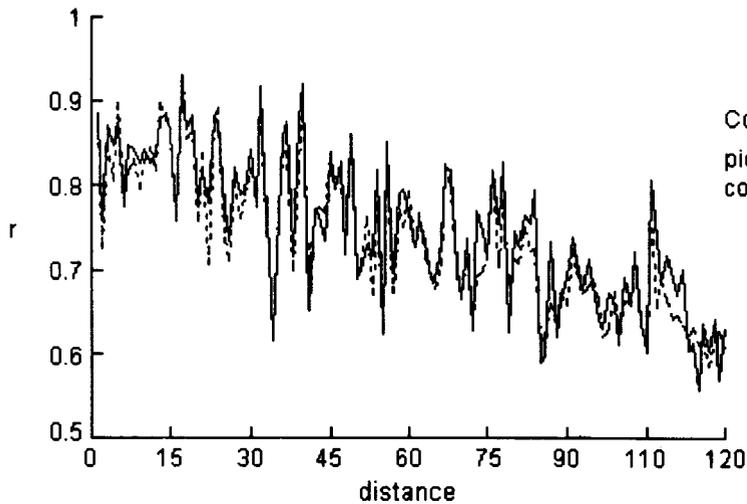
Window	Distance	
	Short	Long
1	0.836	0.577
2	0.822	0.619
3	0.833	0.651
4	0.856	0.698
F	1.19	2.89
p	0.32	0.038
Effect	0.0003	0.003



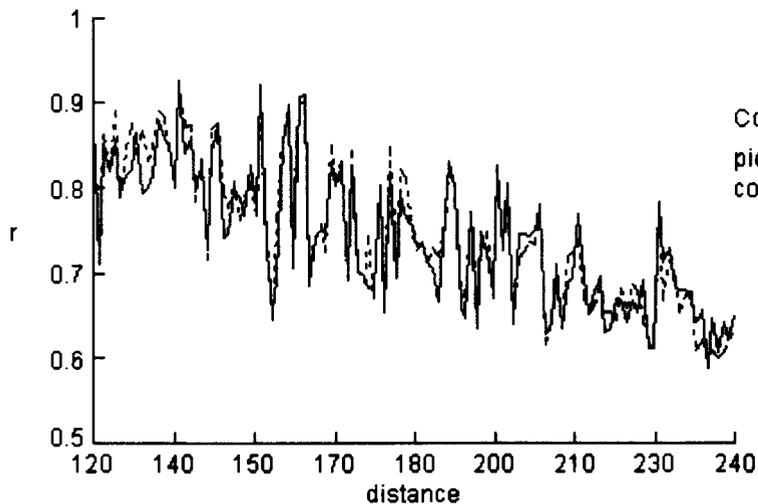
**Figure 6.5 Correlation vs. time**

Correlations are cross-correlogram maxima averaged across all subjects, and across all 120 possible pairwise correlations.

(a) before-blink epoch



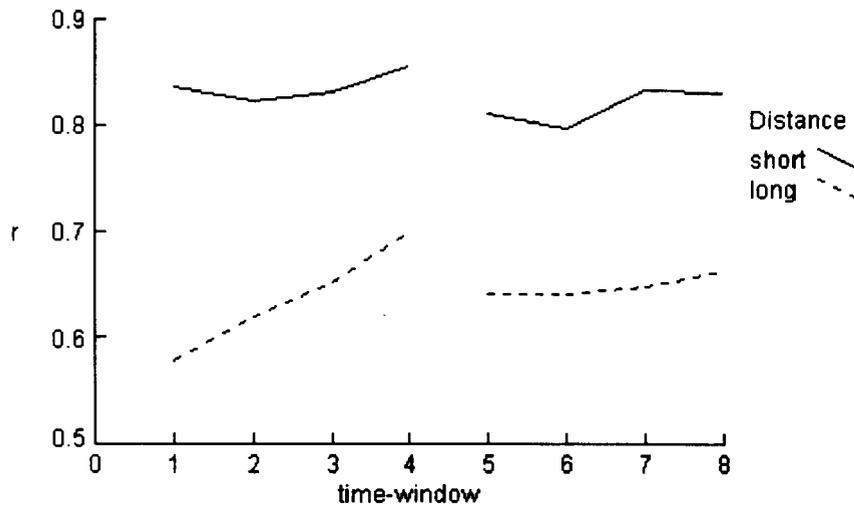
(b) after-blink epoch



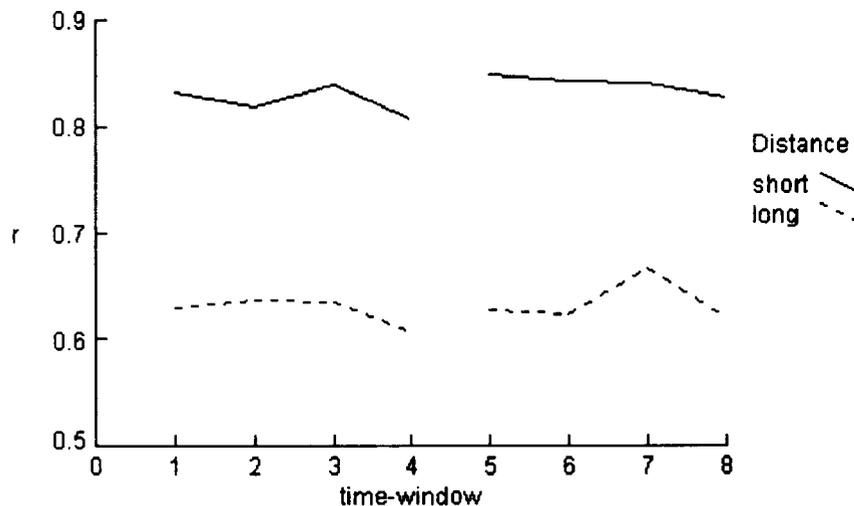
**Figure 6.6 Correlation vs. electrode spacing**

Correlations are cross-correlogram maxima, averaged across subjects and across the 4 time windows of the before-blink epoch (a) and the after-blink epoch (b). Correlations are shown for all 120 possible pairwise combinations of the 16 recorded electrodes. Each point on the graphs shows the correlation for a single pair of electrodes. Electrode pair 1 is part of the group of most closely spaced electrode pairs. Electrode pair 120 is part of the group of most distantly spaced pairs.

(a) picture condition



(b) control condition



**Figure 6.7 Short and long distance correlations**

Short distance correlations are averaged over correlations between the 12 electrode pairs separated by the shortest distance. Long distance correlations are averaged over correlations between the 12 electrode pairs separated by the longest distances. In the picture condition (a) long distance correlations increase significantly from time-window 1 to 4, while short correlations remain relatively constant. All other correlations remain relatively unchanged over the 4 time-windows of the respective epochs.

### 6.3.2 Analysis of Lag

There were no significant effects of time in the picture condition in either the BBE ( $F < 1$ ), or the ABE ( $F = 1.24$ ,  $p = 0.30$ , effect size = 0.0). There were no significant effects of time in the control condition in either the BBE ( $F < 1$ ), or in the ABE ( $F < 1$ ). Table 6.5 lists mean lags for each time window, averaged over all 120 possible electrode pairs. These means are graphed in Figure 6.8.

In the picture condition, there was a significant effect of distance in the BBE ( $F = 2.41$ ,  $p < 0.0001$ , effect size = 0.008), with an increasing value of lag with distance, from approximately 0.35 (2.73 ms) for short distances to 0.55 (4.30 ms) for long distances. A unit of lag corresponds to a time of 7.8125 milliseconds. Also in the picture condition, there was a significant effect of distance in the ABE ( $F = 1.67$ ,  $p < 0.0001$ , effect size = 0.004), with a decreasing value of lag with distance, from approximately 0.23 (1.80 ms) for short distances to 0.04 (0.31 ms) for long distances.

In the control condition, there was a significant effect of distance in the BBE ( $F = 1.46$ ,  $p = 0.001$ , effect size = 0.002), with an increasing value of lag with distance, from about 0.28 (2.19 ms) for short distances to about 0.44 (3.44 ms) for long distances. Also in the control condition, there was a non-significant effect of distance in the ABE ( $F = 1.09$ ,  $p = 0.243$ , effect size = 0.0004), with a decreasing value of lag with distance, from approximately 0.16 (1.25 ms) for short distances to -0.06 (-0.47 ms) for long distances. There was a significant interaction between time and distance in the control condition in the BBE ( $F = 1.25$ ,  $p = 0.001$ , effect size = 0.003), resulting from the value of lag being higher in time-window 2 for long distances than for short distances. Figure 6.9 shows the values of lag as a function of distance for the BBE (a) and ABE (b).

Summarizing the effects of time and distance on lag, there were no significant effects of time on the value of lag in either the control or picture conditions. In both the picture and control conditions, lag increased with distance during the BBE, and decreased with distance during the ABE. These results are listed in Table 6.4.

**Table 6.4 Results of analysis of variance of lag**

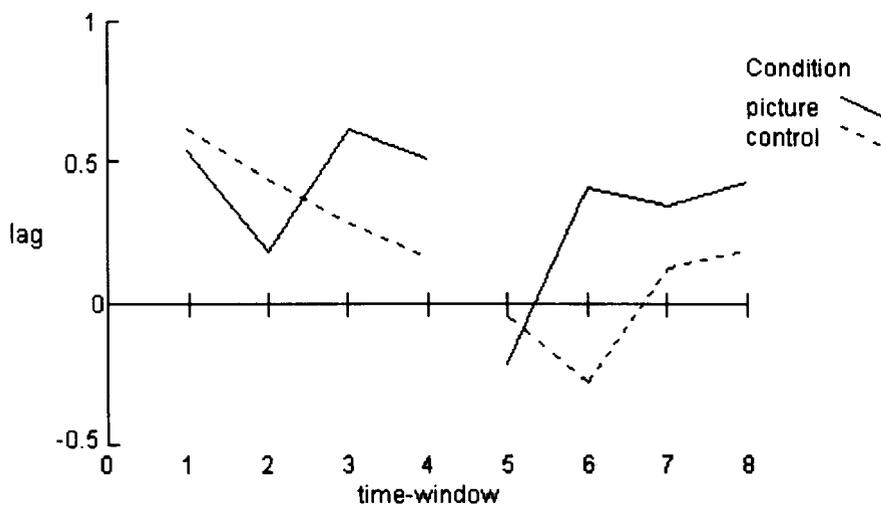
The table shows the results of a two-way within subjects analysis of variance of lags corresponding to cross-correlation maxima. Lag increases significantly with distance in both epochs in the picture condition, and decreases significantly in the before-blink epoch in the control condition. A significant interaction between time and distance occurs in the before-blink epoch in the control condition.

Condition	Epoch	Item	Time	Distance	T x D
Picture	BBE	F	< 1	2.41	< 1
		p Effect		< 0.0001 0.008	
	ABE	F	1.24	1.67	< 1
		p Effect	0.30 0.0	< 0.0001 0.004	
Control	BBE	F	< 1	1.46	1.25
		p Effect		0.001 0.002	0.001 0.003
	ABE	F	< 1	1.09	< 1
		p Effect		0.243 0.0004	

**Table 6.5 Mean lags**

The table shows lags corresponding to cross-correlation maxima averaged over all trials.

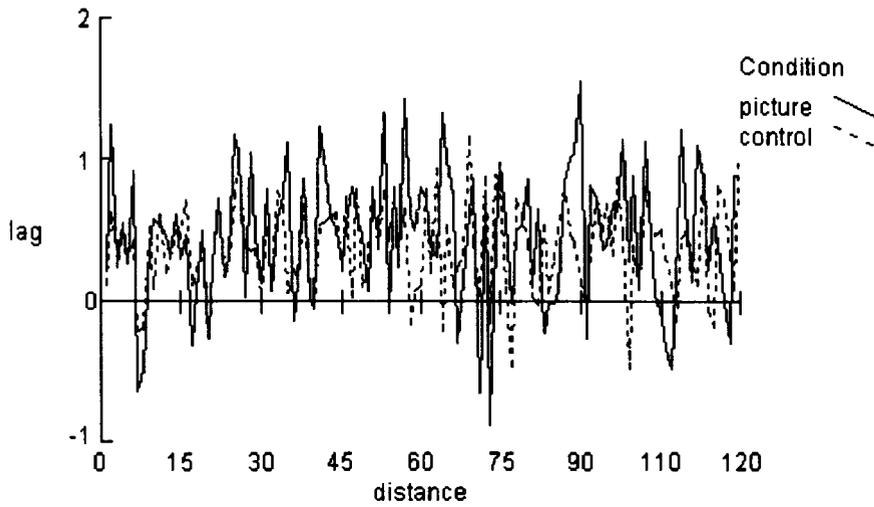
Epoch	Window	Control	Picture
BBE	1	.617	.536
	2	.439	.185
	3	.282	.618
	4	.162	.506
ABE	5	-.046	-.213
	6	-.277	.409
	7	.123	.344
	8	.193	.433



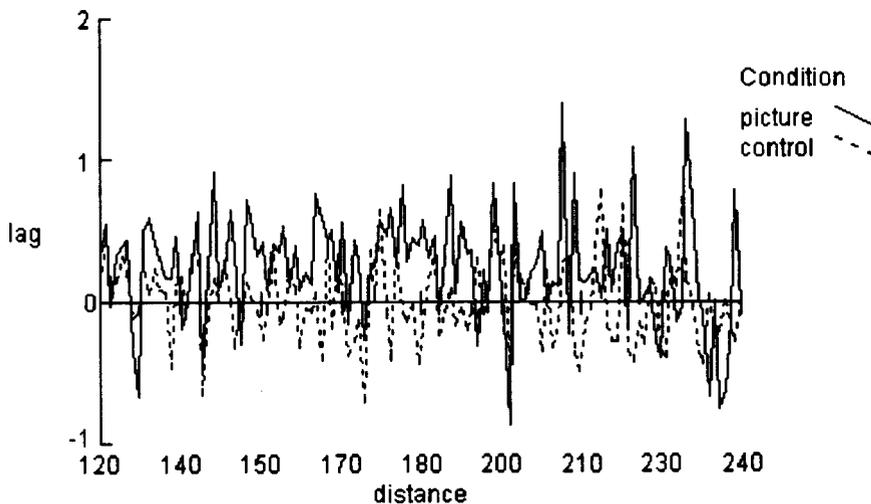
**Figure 6.8 Lag vs. time**

Lags are the values of lag corresponding to cross-correlation maxima for each condition, averaged across subjects and across all 120 possible pairwise correlations.

(a) before-blink epoch



(b) after-blink epoch



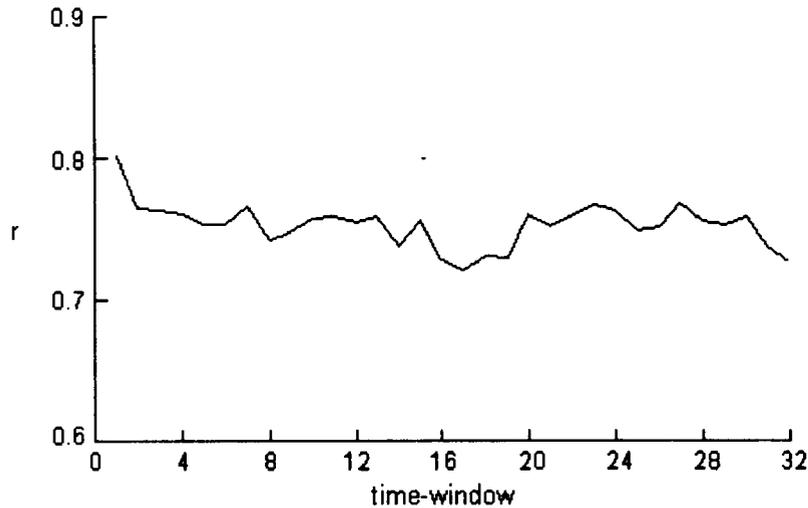
**Figure 6.9 Lag as a function of electrode spacing**

Lags correspond to cross-correlogram maxima, averaged across subjects and across the 4 time windows of the before-blink epoch (a) and the after-blink epoch (b). Lags are shown for all 120 possible electrode pairs of the 16 electrodes that were recorded. Each point on the graphs shows the lag for one pair of electrodes. Electrode pair 1 is part of the group of most closely spaced electrode pairs. Electrode pair 120 is part of the group of most distantly spaced pairs.

### 6.3.3 The No-blink Condition

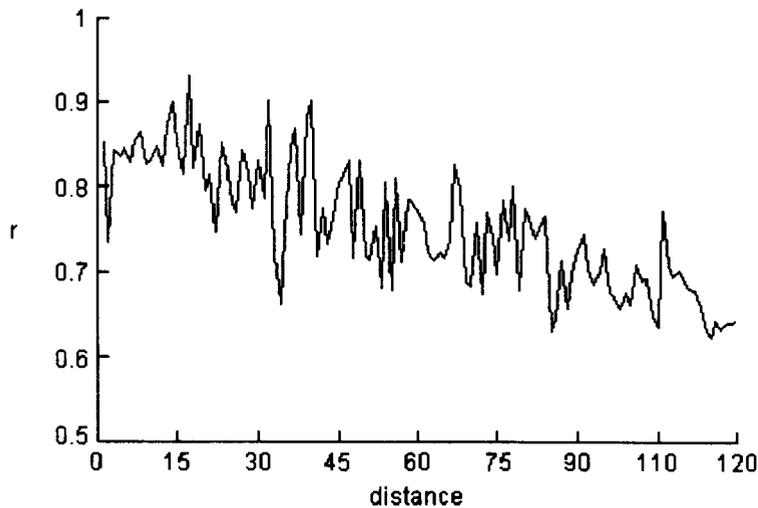
Magnitude of correlation in the no-blink condition is shown in Figure 6.10 as a function of time, and in Figure 6.11 as a function of distance. The magnitude of correlations, averaged over

all distances, was approximately 0.75. This value of correlation is consistent with the values of correlation measured in the control condition. Correlation varied from approximately 0.85 at short distances to approximately 0.65 at long distances. These values are consistent with long and short correlation magnitudes computed in the control condition. The results of the analysis of variance were a significant effect of distance only ( $F = 100.4$ ,  $p < 0.0001$ , effect size = 0.06).



**Figure 6.10 Correlations vs. time, no-blink condition**

Correlations are cross-correlogram maxima for the no-blink condition. Correlations vs. time window are computed for 2 groups of electrode pairs. Short and long correlations are computed between electrode pairs separated by 1, and more than 4 distance units respectively. Correlations are averaged across the 12 correlations within each distance group, across cases, and across subjects. Limits indicate the range of values.



**Figure 6.11 Correlations vs. distance, no-blink condition**

Correlations are cross-correlogram maxima for the no-blink condition, plotted as a function of inter-electrode distance, are averaged across all time-windows and across all subjects. Correlations are shown for all 120 possible electrode pairs of the 16 electrodes that were recorded. Each point on the graphs shows the correlation for one pair of electrodes. Electrode pair 1 is part of the group of most closely spaced electrode pairs. Electrode pair 120 is part of the group of most distantly spaced pairs.

## 6.4 Discussion

### 6.4.1 Correlations

The results of the cross-correlation analysis are consistent with the predictions that were made, on the basis of the CSO model, for the behaviour of correlations in the interval preceding the moment of discrimination. Correlations were found to decrease with increasing inter-electrode distance, to increase with time approaching the moment of discrimination, and to increase more between widely-spaced electrodes than between electrodes that were close together. These findings will now be discussed in more detail.

The observation that significant effects of time were found using theta band components but not, according to pilot analyses, in higher frequency bands, is consistent with similar findings in previous work. Bressler et al. (1993) and Gevins et al. (1987) both observed interregional coherences that decreased in magnitude with increasing frequency, with the largest coherences occurring in the theta band. In a picture interpretation task, Petsche et al. (1992) similarly found large associations in the theta band, between regions including the occipital and temporal cortices.

A number of significant effects of between-electrode distance on cross-correlation were found. In both the control and picture conditions correlations decreased with increasing inter-electrode distance. A similar effect was found by Thatcher et al., (1986) in their study of the spatial distribution of coherence and phase of resting EEG. The authors used their findings as evidence in support of a model of cortico-cortical connections involving short and long distance axonal connections. Relevant to the present finding, Thatcher et al. (1986) suggested that the decrease in coherence with increasing inter-electrode distance was the result of the decreasing density with distance, of short-axonal connections. At short inter-electrode distances, they suggest, coherence is mediated by short-distance axonal connections between Golgi type II cells, also referred to as interneurons, short-axoned neurons whose axons form dendrites in the immediate neighborhood of the cell body. The density of these connections between any two cortical regions decreases with the distance between the regions, causing a corresponding drop in level of coherence. At large inter-regional distances coherence, they suggest, would be mediated instead by long-axonal connections between Golgi type I cells, long-axoned neurons with projections to relatively distant regions. The combined effect, they predicted, was that level of coherence should be an approximately quadratic function of distance, first decreasing with distance to reflect the diminishing effect of interneurons with increasing distance, and later increasing to reflect the effect of long-axoned neurons. Confirming this suggestion, their finding was that coherence at first decreased rapidly with distance, and then leveled off at a non-zero magnitude. In the present study, while there is no evidence of a quadratic relationship as such between correlation and distance, it is nevertheless clear that correlation, while decreasing with distance, does not fall to zero at the largest distances. This finding might thus be interpreted as being generally consistent with the hypothesis that at the short distances correlation is mediated by short-axoned associations between Golgi type II neurons, while at long distances over which the effect of short axoned connections is presumably minimal, the non-zero value of correlation must be mediated by the effects of long-axoned associations between Golgi type I neurons. On the other hand, the non-zero value of correlation at long distances might reflect the effects of volume conduction. The results of the analysis of lag, discussed below, can be used to decide between these two possibilities.

A number of effects of time on cross-correlation were also found. In the present study, in the control condition there is a slight and non-significant decrease in correlation over the 1 second epoch preceding the eye-blink, with most of the decrease occurring during the final quarter second of the epoch. In the picture condition, there is a significant increase in correlation over the 1 second interval before the eye-blink that signals discrimination. Furthermore in the picture condition there is a greater increase in correlation between electrode sites separated by longer distances than between those separated by shorter distances. Thus, although in the

picture condition the patterns of changes with time are found for all inter-electrode distances, the most noticeable effects of the experimental manipulation are found in the correlations between the most widely separated sites. Summarizing this pattern, in the picture condition, there is an increase in correlations over the interval preceding discrimination that is greater between electrodes spaced far apart on the scalp relative to electrodes that are adjacent on the scalp. These findings are consistent with the hypothesis that, as the process of object discrimination proceeds towards the moment of discrimination, there is increasing coordination between the activities of ever more distantly separated brain regions, as a result, it is proposed, of reciprocal signaling between these regions, along cortico-cortical association tracts between Golgi type I neurons.

A neuroanatomical basis for such reciprocal, inter-regional signaling is known to exist in the form of long axonal pathways connecting multiple regions of the cortex. These pathways can be functionally grouped into three categories: association tracts, commissures, and projection tracts, that together constitute the white matter of the brain (Diamond, Scheibel and Elson, 1985; Barr and Kiernan, 1988; Nieuwenhuys, Voogd and van Huijzen, 1981). These connections will now be described in order to illustrate the high degree of connectivity that exists both within and between hemispheres, and between subcortical and cortical regions.

Association tracts consist of bundles of axons that connect various regions within each of the cerebral hemispheres. Association tracts themselves may be classified generally into two groups on the basis of the length of the tracts. Short association fibers, also referred to as U-fibers, connect adjacent cortical gyri. Long association fibers connect more distant regions, and include three major tracts. The cingulum connects the frontal and parietal lobes with parts of the temporal lobe and with the parahippocampal gyrus. The uncinate fasciculus connects the anterior regions of the temporal lobe with the orbital gyrus of the frontal lobe and with portions of the middle and inferior frontal gyri. A portion of the uncinate fasciculus known as the inferior occipitofrontal fasciculus connects the orbital and frontal gyri with the occipital lobe. Lastly, the arcuate fasciculus forms a path between the superior and middle frontal gyri with portions of the temporal lobe. A part of the arcuate fasciculus referred to as the superior longitudinal fasciculus connects areas of the frontal and occipital cortices. These association fibers are shown schematically in Figure 6.12.

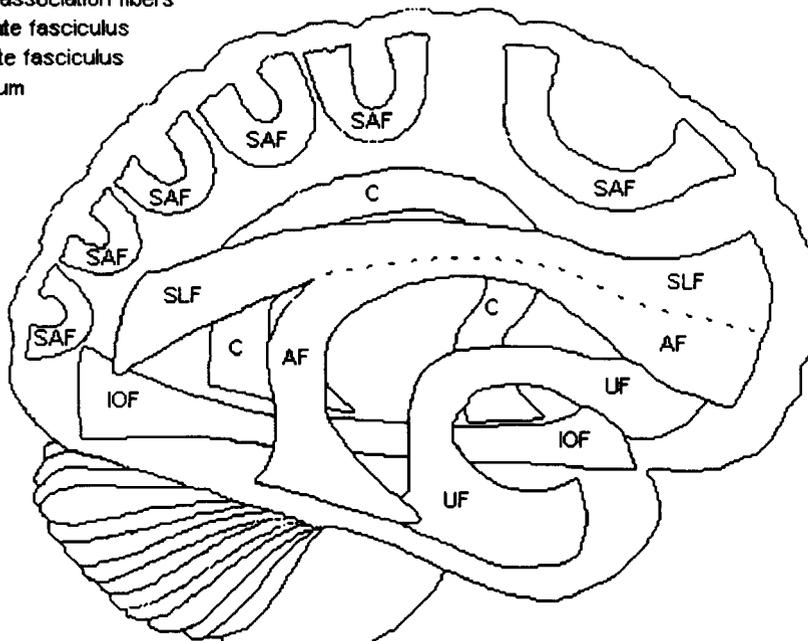
Commissures are groupings of axons that form paths between homotopic areas in the two cerebral hemispheres. The two major commissural fiber bundles are the corpus callosum and the anterior commissure. The corpus callosum is composed of the genu anteriorly, and the body and the splenium posteriorly. The genu connects corresponding anterior cortical regions in the two hemispheres. The body, which intersects association tracts and projection fibers in each hemisphere, provides one of the principal paths between corresponding regions of the left and

right hemispheres. The splenium interconnects posterior cortical areas including the left and right occipital cortices. The anterior commissure is one of the bundles of fibers which provides a path between the left and right temporal lobes. These commissures are shown schematically in Figure 6.13.

Projection fibers are bundles of axonal fibers that connect regions of the cortical sheet with subcortical nuclei. In the medullary center, these projection fibers form the corona radiata, which links with many areas of the pyramidal cell layers of the cerebral cortex. In the subcortical regions the fibers of the corona radiata congregate in the internal capsule, carrying fibers many of which function as a reciprocal signal pathway between the thalamus and the cerebral cortex. The internal capsule is divided into 5 parts: the anterior limb, the genu, the posterior limb, the retrolenticular fibers, and the sublenticular fibers. The anterior limb contains frontopontine fibers, as well as connecting the mediodorsal thalamic nucleus with the prefrontal cortex. The genu includes fibers originating in the ventral lateral nucleus of the thalamus, and projecting to motor and premotor areas of the frontal lobe. The posterior limb carries the middle thalamic radiation, which includes efferent fibers of the ventral posterior thalamic nucleus that project to the somesthetic region of the parietal lobe. The middle thalamic radiation carries as well other fibers that contribute towards creating a reciprocal signaling system between the thalamus and the association cortex of the parietal lobe. The retrolenticular fibers originate largely as efferent fibers of the lateral geniculate nucleus, and form the optic radiations which terminate in the primary visual area of the occipital cortex. The sublenticular fibers originate mainly in the medial geniculate nucleus, and continue as the auditory radiations to project to the auditory areas of the temporal lobe. These projection fibers are shown schematically in Figure 6.14.

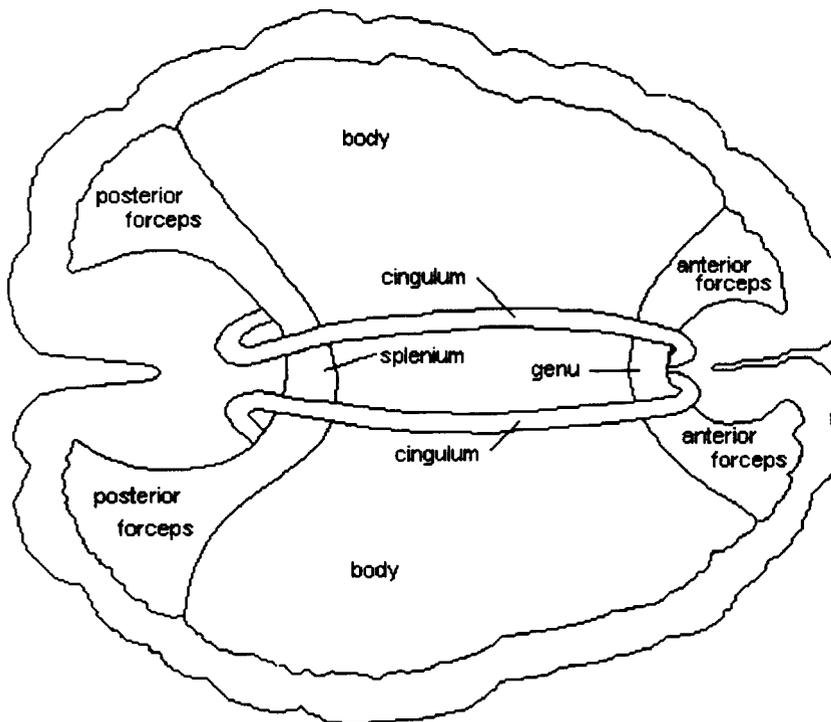
Together, these three types of neural pathways are reasonable candidates to form the neuroanatomical basis of the correlations that have been found in the present study to exist between signals recorded between all areas of the cortex. In terms of volume, the bundles of fibers connecting the various cortical and subcortical regions occupy a substantial if not the major portion of the volume of the brain. Connectivity would clearly appear to be of the essence in the functioning of the brain. It is suggested that the pattern of correlations found in the present study to be associated with visual discrimination, is directly made possible by this connectivity.

IOF inferior occipitofrontal fasciculus  
 SLF superior longitudinal fasciculus  
 SAF short association fibers  
 UF uncinete fasciculus  
 AF arcuate fasciculus  
 C cingulum



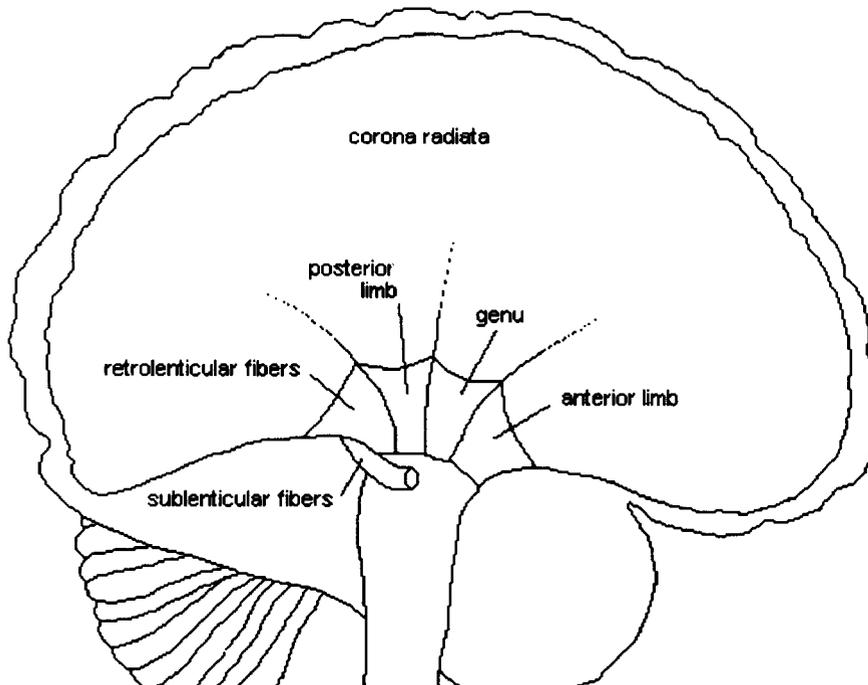
**Figure 6.12 Association Fibers**

The major association fibers, shown schematically in this figure, interconnect cortical and subcortical regions within each of the two hemispheres (Adapted from Diamond et al., 1985).



**Figure 6.13 Commissural Fibers**

Commissural fibers, shown in this schematic presentation, connect homotopic regions of the left and right hemispheres (Adapted from Diamond et al., 1985).



**Figure 6.14 Projection Fibers**

The projection fiber systems shown here schematically connect subcortical nuclei with multiple and widespread cortical regions (Adapted from Diamond et al., 1985).

As a final issue concerning methodology, in this experiment, the order of presentation was not varied between subjects. All subjects first received, twice over, the entire set of 31 images, and then received the 20 trials of the control condition. There was no attempt to counterbalance the order of presentation. A consequence of this could be some measure of practice effect. In general, such a practice effect would mean that subjects had learned something during the initial presentations of the stimulus images, and that such learning would then have influenced their performance in the control trials. It was initially conjectured however that probably all that subjects would learn during the picture trials would be how to maintain their visual focus on the fixation spot for the eight second recording intervals. It is possible nevertheless that during the picture trials subjects might learn the behaviour of scrutinizing for possible hidden images whatever image was presented. Such scrutiny could then have been applied to the blank screen of the control condition, perhaps resulting in representations of objects being 'discovered' when none were being presented. In other words, it is possible that there might have occurred some measure of the phenomenon seen with participants in sensory deprivation experiments, in which neutral visual screens can provide a background for hallucinatory episodes. Counterbalancing by presenting some subjects with the control trials before the picture trials would have been able to deal with this question.

### 6.4.2 Lags

At the outset there were no predictions made concerning the behaviour of lag. Lag was therefore analyzed from an exploratory perspective. There were no significant changes in lag with time in either condition or epoch. There were however significant changes in lag with inter-electrode distance. These effects are graphed in Figure 6.9a for the before-blink epoch and Figure 6.9b for the after blink epoch.

In both picture and control conditions, lag was found to increase with increasing inter-electrode distance in the epoch preceding the blink. Lag increased from approximately 0.35 (2.7 ms) to 0.54 (4.2 ms) in the picture condition, and from approximately 0.28 (2.2 ms) to 0.43 (3.4 ms) in the control condition. This increase in lag with distance before the blink is consistent with an explanation proposed by Thatcher et al. (1986). The authors computed, for resting EEG, coherence and phase, measures that are analogous to squared cross correlation and lag. The authors suggested that if volume conduction were responsible for coherence then phase should not vary with distance, since volume conduction mechanisms involves relatively short time delays whose rate of change with distance is correspondingly small. If, on the other hand, coherence is mediated by axonal transmission along association and other tracts, then it is to be expected that lag should vary relatively more as a function of distance, increasing with distance, since axonal signal transmission involves time-delays that do increase significantly with distance. The present findings are consistent with this latter position, that the computed patterns of correlation must be primarily the result of axonal transmission, along association fibers, commissures and projection fibers, rather than by volume conduction.

A contrasting result however was found in picture condition in the epoch following the blink, with lag decreasing with increasing inter-electrode distance. The value of lag decreased from approximately 0.23 to 0.04. This result might be interpretable in terms of a conjectured description of the relative level of synchronization within and between cortical areas. Thus, lags between closely-spaced regions remain relatively unchanged with time both prior to, and following discrimination. The magnitude of lags between more distantly-spaced regions, however, is different before and after discrimination, reflecting differences in the nature of interregional signaling between the before and after-blink epochs. Before discrimination, the visual analytic processes associated with the analysis of the image are associated with a relatively high level of interregional signaling between more widely-separated cortical areas, thus making manifest the time-delay effects associated with axonal transmission. Immediately following successful discrimination there is relatively less signaling between widely-separated cortical regions, so that the time-delay effects associated with axonal transmission would be relatively insignificant. Furthermore, the high degree of synchronization between distantly-

spaced regions immediately following discrimination could result in the lag between such regions decreasing to values that are lower than those for more closely-spaced regions. Thus, before discrimination, as the visual image is being analyzed, the high level of interregional communication would be associated with distance-dependent time-delays and therefore a relatively large value of lag. Once discrimination has occurred and synchronization between distant cortical regions has been established, the level of such inter-regional communication could decrease, lessening the impact of distance-dependent delays, while the now-synchronized activity between these distantly-spaced cortical regions would be reflected in a low value of lag.

The second significant finding with respect to lag was the interaction between time and distance following the blink in the control condition. The value of lag was higher for long distances than for short distances, but this difference was larger for early time-windows than for later time-windows in the before-blink epoch. Thus, for later time-windows, approaching the blink, the value of lag was approximately the same for short and long distances. For early time-windows however, the value of lag was higher for long than for short distances. This finding might be interpreted as reflecting some small amount of visual analysis occurring in the early time-windows preceding the blink in the control condition. The level of such visual analysis might be expected to decrease to even lower levels in a relaxation of visual attention immediately before the blink. Overall then, this interaction might, it is conjectured, indicate the changing level of visual attention directed towards the neutral target and fixation point presented to subjects in the control condition.

#### **6.4.3 Correlation Before and After Discrimination**

Examination of the mean correlation for each time-window (Table 6.2) shows that a part of the increase in correlation that occurs prior to discrimination is still evident immediately following discrimination, after which the level of correlations drop to approximately pre-discrimination levels. A comparison of the levels of correlation immediately before and after discrimination might provide some evidence for the kinds of processes involved in the process of recognizing camouflaged objects. Correlations immediately following discrimination, 0.745, were intermediate in value between their initial levels at the start of the BBE, 0.715, and their levels at the end of the BBE, 0.789. To the extent that level of correlation may be related in a general way to type of neural processing, this finding might indicate that at least some, although not all, of the processes that were occurring just before discrimination were still in operation immediately after discrimination. As a first approximation functional description of the process of object discrimination, it might be conjectured that this process involves both a graphical and a lexical component. The graphical component would include those neural sub-processes involved with

identifying visual elements of the image, transforming bundles of elements into feature ensembles, and associating those feature ensembles with pre-existing visual memories. These are a subset of the processes that could be reasonably be expected to be involved in visual discrimination. The lexical component of the process of visual discrimination would include those neural subprocesses involved in associating the emerging feature ensembles with pre-existing lexical memories, or in alternative terms, in retrieving a label for the discriminated object.

It seems reasonable that while either or both of these components might be active immediately before discrimination, only sub processes other than those associated with labeling of the object should be active immediately following discrimination: once a label for the object has been retrieved the associated neural activities would presumably no longer be required to be active. That is, in most instances of discrimination, if labeling is involved then it will occur immediately prior the moment of discrimination rather than immediately following discrimination. Since at least a portion of the increase in correlation just before discrimination is still present just after the blink, it might be concluded that not all of the increase in correlated activity occurring just before discrimination is associated with the lexical component. Rather, at least a portion of the increase in correlated activity just before discrimination should be expected to be associated with processes other than those associated with the lexical component of discrimination, and therefore associated with the graphical component. One tentative conclusion that might be drawn from these findings is that both a graphical and a lexical component may be involved in the discrimination process, or alternatively that both processes are involved in at least some trials and for some subjects. Other scenarios are possible, however. An alternative conclusion would be that only the graphical component is involved, but that the sub-processes involved in this component are more active immediately before than immediately after discrimination. Still a third possibility is that the graphical component is active immediately before, while the lexical component is active immediately following discrimination, as subjects blink upon visually recognizing the stimulus image, but then afterwards seek to associate the image with a label.

Follow-up debriefings of all subjects indicated that several of these possibilities, in fact, occurred. On some presentations subjects blinked after they had recognized the object and had a label for it, while on other presentations they blinked after only recognizing the visual image, and before a label was available. In those cases, subjects indicated that a label would sometimes become available after the blink.

It might be concluded that the process of visual discrimination of a relatively complex image cannot be readily associated with either only a visual or only a lexical component. Rather, as subject debriefings indicated, the graphical component might occur first to be followed by the lexical component, or alternatively both of these components might occur together. In the

present paradigm it was felt to be important that the attention of subjects not be directed towards labeling of the visual images, as it was felt that such direction would have an effect on the nature of the events preceding the eye-blink, and in fact might bias subjects towards wanting to label the stimulus images before-blinking to signify discrimination. In the procedure used in the present study, therefore, subjects were deliberately not debriefed after individual trials. With follow-up debriefings conducted after the end of the testing sessions, it was not possible to determine for each trial whether labeling had been involved: such debriefings suggested that subjects would have difficulty in recalling accurately the time-order of discrimination and labeling on individual presentations. In the present study, therefore, the neural processes associated with visual object discrimination will not be dissected into a graphical and a lexical component, but rather will be considered only as a whole, consisting of a conglomerate of component sub-processes which are assumed to include a requisite graphical component and an optional lexical component.

#### **6.4.4 The No-blink Condition**

The validity of the control condition is supported by the results of the analysis of the no-blink condition, those trials in which subjects viewed a camouflaged object but failed to recognize it. The values of correlation in this condition are approximately equal to the corresponding levels of correlation in the control condition. This finding would suggest that as far as level of correlation is concerned, there is little difference between subjects simply looking at a blank screen as in the control condition, and subjects looking at a camouflaged object and not recognizing it as in the no-blink condition. It seems reasonable to infer from this that the results would not have been different if a complex but undiscriminable object, rather than a blank diffusing screen, had been used in the control condition. Thus, the results measured in the control condition appear to be reasonably independent of the type of neutral image used, supporting the validity of the control condition.

## 7 Topographical Distribution of Correlations

### 7.1 Introduction

The findings of the correlation analysis were that intercorrelations increased significantly with time preceding the moment of discrimination, and that this increase with time was larger for electrode pairs separated by large distances than for pairs that were close together. The work of the present section is intended to identify the particular brain regions that were involved in this pattern of correlations. The present section therefore examines the topographical distribution over the scalp of the change in correlations, over the duration of the before-blink epoch. More particularly, the present section attempts to answer the following question: What is the topographical distribution of the differences, between the picture and control conditions, of the correlations between pairs of electrodes, in each of the 4 time-windows of the before-discrimination epoch? That is, having computed for each electrode pair the difference in correlations between the picture and control conditions, how are these differences distributed over the scalp?

Correlation differences will be investigated, rather than absolute correlation values, in each of the conditions, because what is wanted is a reflection of the effect of the experimental manipulation on the distribution of correlations. Within the record of the before-blink epoch, there may be evidence of other effects, such as preparation for the eye-blink. By examining correlation difference rather than absolute correlation it is intended that such effects, that are not related to the experimental manipulation, will be subtracted out. A second motivation for using correlation differences is the results of pilot work, which indicated that examining such absolute correlation values did not provide a clear picture of the topology of the correlated activity. Absolute correlation values in both conditions were found to be large, at approximately 0.6 to 0.8, in relation to the corresponding between-condition differences, which were typically approximately 0.1, in a small subset of the 120 possible electrode pairs.

### 7.2 Method

A 2-dimensional projection or map of the physical electrode positions over the scalp is used to display the between-channel correlations. This map is used to display the difference in magnitude of correlation between control and picture conditions, with a separate map for each of the 4 time windows of the before-blink epoch. This correlation difference  $\delta r_{i,k}$  is computed for each pair of electrodes,  $i$  and  $k$ , as

$$\delta r_{i,k} = r(\text{picture})_{i,k} - r(\text{control})_{i,k}$$

where  $r(\cdot)_{i,k}$  represents the intercorrelation between electrodes  $i$  and  $k$ . The magnitude of this correlation difference for a pair of electrodes is coded in terms of the thickness of a line joining the two electrodes. Negative correlation differences are shown as gray lines, positive changes in correlation are shown as black lines. A positive difference indicates a higher correlation in the picture condition. A negative difference correspondingly indicates a higher correlation in the control condition.

### 7.3 Results

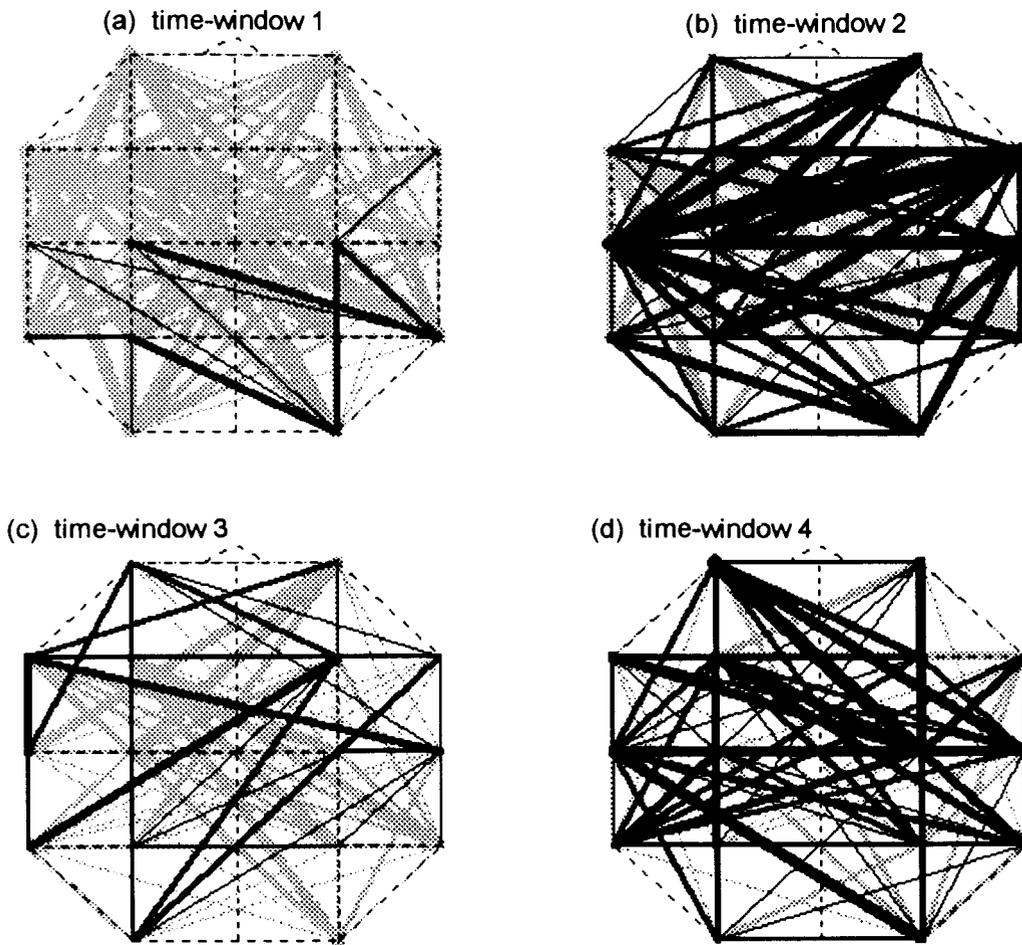
The topographical distributions of  $\delta r$ , the changes in cross-correlation between the picture and control conditions are shown in Figure 7.1 through Figure 7.5 for each subject, and in Figure 7.6 averaged across subjects. Table 7.1 lists the values of  $\delta r$  averaged across subjects and sorted in order of increasing distance between electrode pairs.

A number of general features may be immediately noted by examining Figure 7.6. First, between the 4 time-windows there are clear changes in the topography of the values of  $\delta r$ . Second, these changes progress from relatively little difference in time-window 1, to a large positive difference in time-window 4 indicating a greater degree of correlation, in time-window 4, in the picture than in the control conditions. Third, each of the 4 time-windows displays a unique pattern. These patterns are summarized as follows:

1. Time-window 1: there is relatively little difference between the picture and control conditions, with minor positive values over the right frontal (F8) and central (C4) areas, indicating little difference in correlation in all areas between the picture and control conditions.
2. Time-window 2: the pattern now shows a greater positive value of correlation difference over the left hemisphere. Largest positive values occur over left anterior temporal (T3) and right occipital (O2) areas. Somewhat smaller positive values are found over left occipital (O1), left central (C3), left and right posterior temporal (T5, T6) and right parietal (P4). This pattern indicates a larger correlation in the picture than the control conditions between bilateral occipital areas and bilateral posterior temporal areas, and the left anterior temporal area.
3. Time-window 3: the distribution of  $\delta r$  values now shows a more bilateral distribution of positive values, localized over bilateral frontal (F7, F8) and right anterior temporal (T4) areas, with somewhat smaller values over the left frontal area (Fp1). This pattern indicates larger

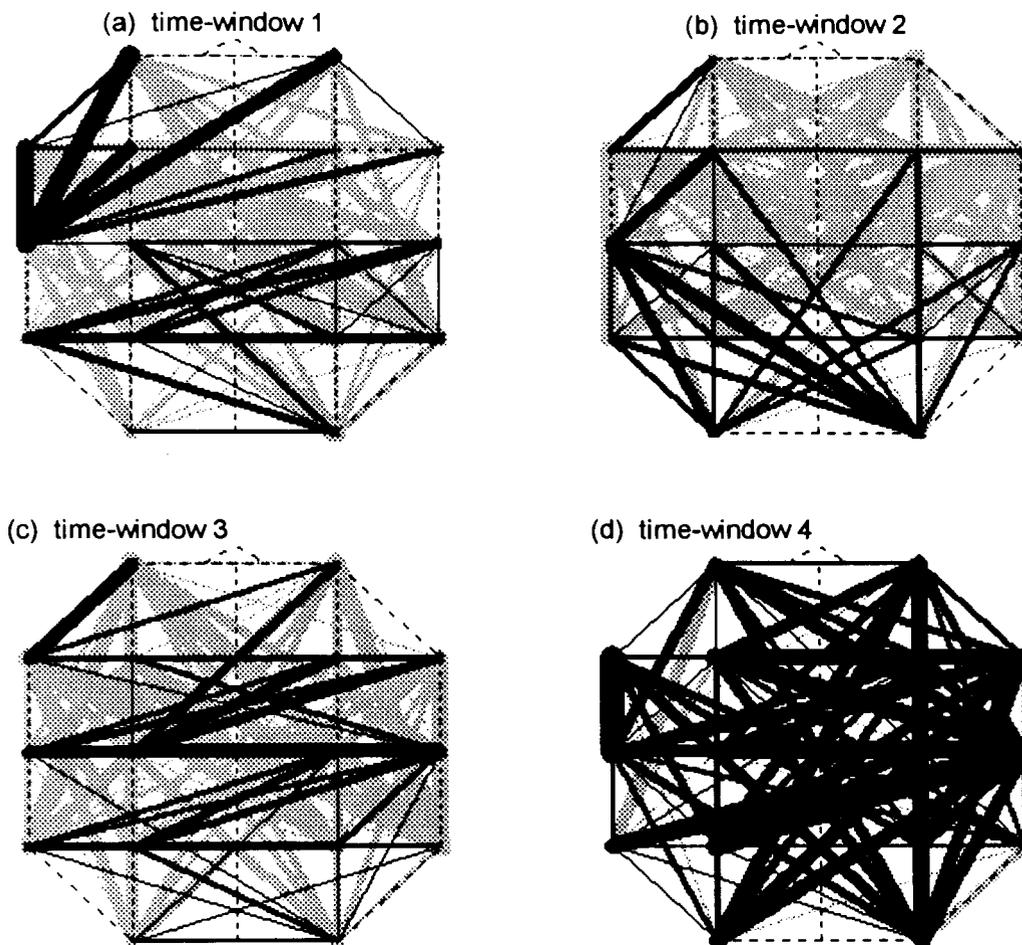
interhemispheric correlations, in the picture than in the control conditions, between bilateral frontal and temporal areas.

4. Time-window 4: the overall pattern now shows the involvement of all areas in large positive values of correlation difference, indicating a relatively larger correlation between all areas in the picture than the control conditions. There is a somewhat greater degree of correlation between left and right hemispheres than in the anterior-posterior direction.



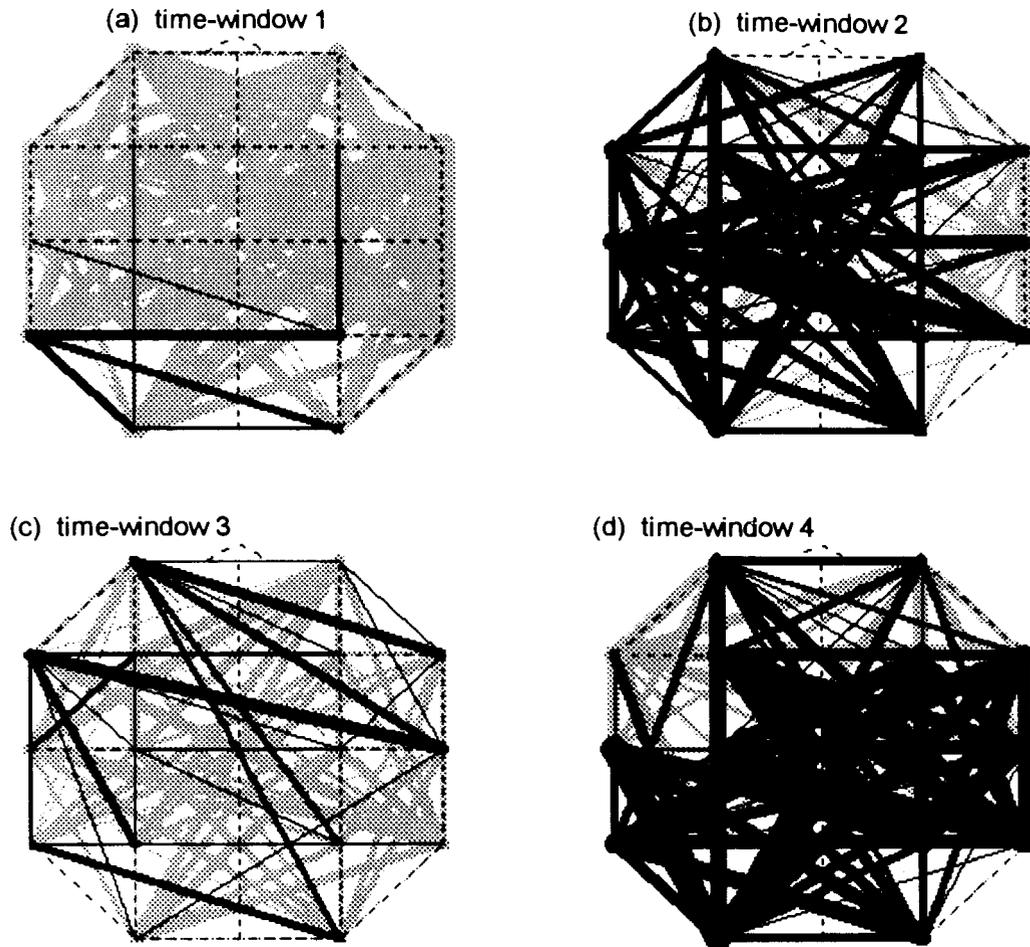
**Figure 7.1 Topography of correlation differences, subject 4**

The graphs show the magnitude of the difference in correlations, between the picture and control conditions. Correlation differences are computed as  $r(\text{picture}) - r(\text{control})$ . Correlations are cross-correlogram maxima, averaged across all trials for this subject. Line thickness is proportional to magnitude of correlation difference. The thickest lines correspond to a correlation difference of 0.15. Correlation differences less than 0.03 are not shown. Black lines indicate a positive difference, with a higher correlation in the picture than in the control conditions. Gray lines indicate a negative difference, with a higher relative correlation in the control condition.



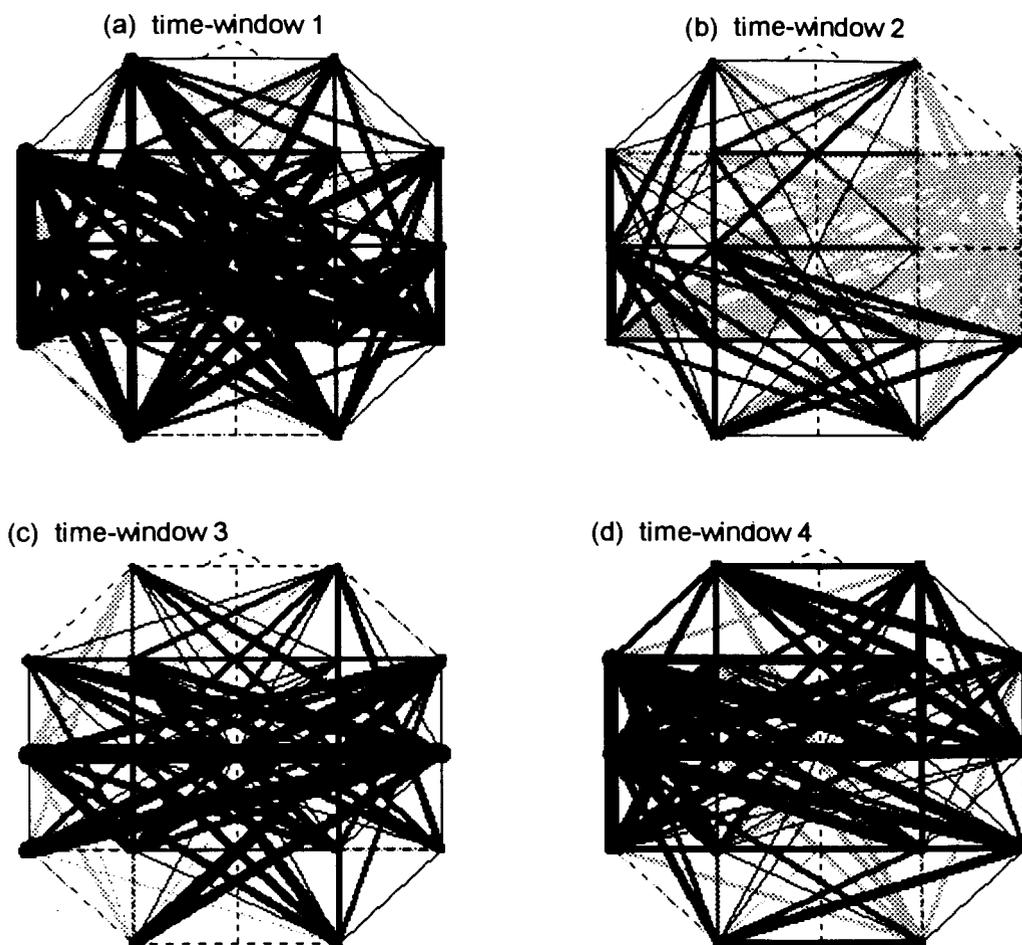
**Figure 7.2 Topography of correlation differences, subject 5**

The graphs show the magnitude of the difference in correlations, between the picture and control conditions. Correlation differences are computed as  $r(\text{picture}) - r(\text{control})$ . Correlations are cross-correlogram maxima, averaged across all trials for this subject. Line thickness is proportional to magnitude of correlation difference. The thickest lines correspond to a correlation difference of 0.15. Correlation differences less than 0.03 are not shown. Black lines indicate a positive difference, with a higher correlation in the picture than in the control conditions. Gray lines indicate a negative difference, with a higher relative correlation in the control condition.



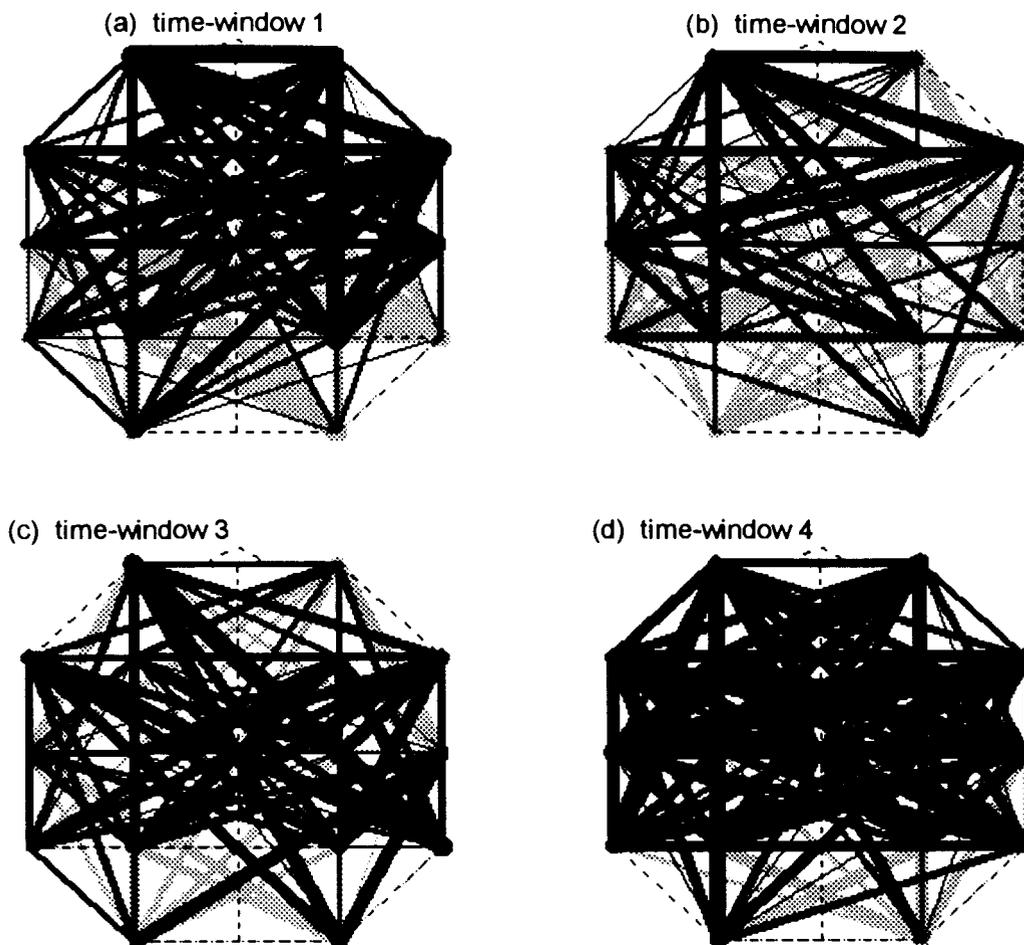
**Figure 7.3 Topography of correlation differences, subject 7**

The graphs show the magnitude of the difference in correlations, between the picture and control conditions. Correlation differences are computed as  $r(\text{picture}) - r(\text{control})$ . Correlations are cross-correlogram maxima, averaged across all trials for this subject. Line thickness is proportional to magnitude of correlation difference. The thickest lines correspond to a correlation difference of 0.15. Correlation differences less than 0.03 are not shown. Black lines indicate a positive difference, with a higher correlation in the picture than in the control conditions. Gray lines indicate a negative difference, with a higher relative correlation in the control condition.



**Figure 7.4 Topography of correlation differences, subject 8**

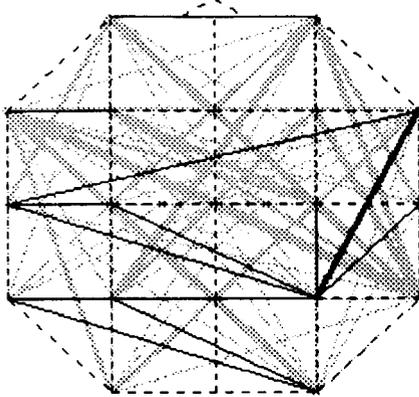
The graphs show the magnitude of the difference in correlations, between the picture and control conditions. Correlation differences are computed as  $r(\text{picture}) - r(\text{control})$ . Correlations are cross-correlogram maxima, averaged across all trials for this subject. Line thickness is proportional to magnitude of correlation difference. The thickest lines correspond to a correlation difference of 0.15. Correlation differences less than 0.03 are not shown. Black lines indicate a positive difference, with a higher correlation in the picture than in the control conditions. Gray lines indicate a negative difference, with a higher relative correlation in the control condition.



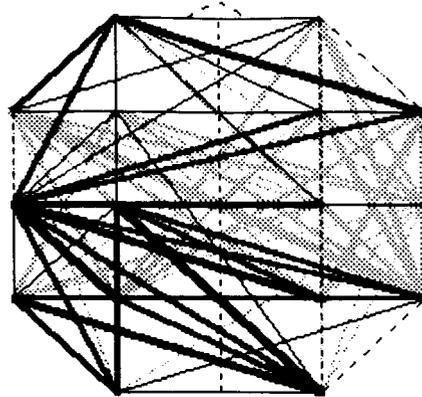
**Figure 7.5 Topography of correlation differences, subject 9**

The graphs show the magnitude of the difference in correlations, between the picture and control conditions. Correlation differences are computed as  $r(\text{picture}) - r(\text{control})$ . Correlations are cross-correlogram maxima, averaged across all trials for this subject. Line thickness is proportional to magnitude of correlation difference. The thickest lines correspond to a correlation difference of 0.15. Correlation differences less than 0.03 are not shown. Black lines indicate a positive difference, with a higher correlation in the picture than in the control conditions. Gray lines indicate a negative difference, with a higher relative correlation in the control condition.

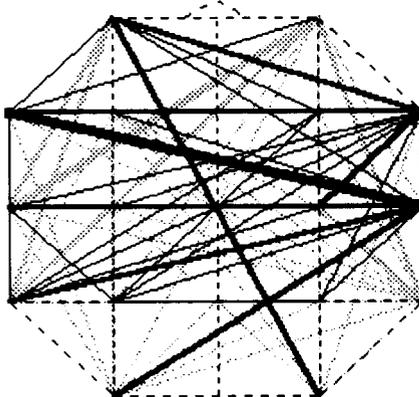
(a) time-window 1



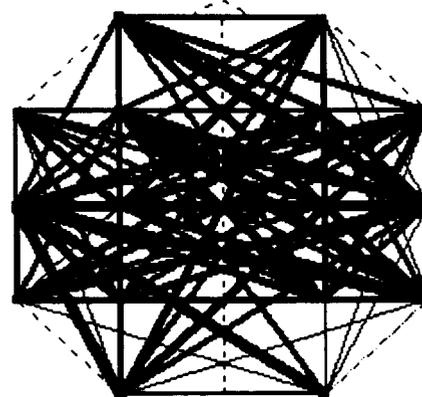
(b) time-window 2



(c) time-window 3



(d) time-window 4



**Figure 7.6 Topography of correlation differences, average across subjects**

The graphs show the magnitude of the difference in correlations, between the picture and control conditions. Correlation differences are computed as  $r(\text{picture}) - r(\text{control})$ . Correlations are cross-correlogram maxima, averaged across all trials and all subjects. Line thickness is proportional to magnitude of correlation difference. The thickest lines correspond to a correlation difference of 0.15. Correlation differences less than 0.03 are not shown. Black lines indicate a positive difference, with a higher correlation in the picture than in the control conditions. Gray lines indicate a negative difference, with a higher relative correlation in the control condition.

**Table 7.1 Correlation change**

The correlation change, computed as  $r(\text{picture}) - r(\text{control})$ , is shown for all possible pairs of the 16 electrodes and for each time-window of the before blink epoch.

Channel 1	Channel 2	Time-window			
		1	2	3	4
C4	P4	0.035	-0.0303	-0.0277	0.0451
T3	T5	-0.0238	0.0409	0.00972	0.103
T5	P3	0.0482	0.0542	-0.0263	0.00894
T3	C3	0.024	0.0659	0.0154	0.0705
P4	O2	-0.0237	0.00824	-0.0326	-0.0396
T4	T6	-0.0208	-0.106	-0.0101	0.0457
C4	T4	0.0159	0.0182	0.0135	0.0761
F4	F8	-0.00586	-0.00558	0.0046	0.0582
F7	F3	0.0343	0.0472	-0.0168	0.0606
Fp2	F4	0.0173	-0.0211	0.00464	0.0639
F8	T4	-0.032	-0.0547	-0.0111	0.0259
Fp1	F3	-0.00503	0.0339	-0.00884	0.0622
F4	C4	-0.054	-0.0519	0.0134	0.0232
F3	C3	-0.0392	-0.00236	-0.00888	0.0749
C3	P3	-0.0484	0.0304	0.00768	0.0369
F7	T3	-0.0296	-0.0368	0.0184	0.0664
P4	T6	-0.0241	-0.0268	-0.00491	-0.0163
P3	O1	0.00663	0.059	-0.0379	0.0225
T5	O1	0.00007	0.0565	-0.00732	0.0117
F3	T3	-0.00815	0.0553	-0.0164	0.0655
C4	T6	-0.0343	-0.0582	-0.06	0.0508
T3	P3	0.0124	0.13	0.0477	0.107
Fp2	F8	0.0158	-0.0204	0.00154	0.00688
T6	O2	-0.0456	0.00706	-0.022	-0.0337
F8	C4	0.0156	-0.014	0.0605	0.086
F7	C3	-0.0162	0.0112	-0.0156	0.108
F4	T4	0.00257	-0.0772	0.0176	0.125
T4	P4	0.0517	-0.0515	0.0343	0.0737
C3	T5	-0.0357	0.0192	-0.00599	0.0426
Fp1	F7	0.00372	0.0677	0.026	-0.00272
F4	P4	-0.0229	-0.0842	-0.00807	0.0521
O1	O2	0.0178	0.0533	-0.00242	0.0645
C3	O1	-0.0498	0.099	-0.0861	0.0381
F7	T5	-0.031	-0.043	0.0337	0.0576
Fp2	C4	-0.0436	-0.0647	-0.0348	0.0865
C3	C4	-0.0276	0.023	0.00387	0.0867
F3	F4	-0.00009	0.0373	0.0382	0.0683
Fp1	C3	-0.0674	0.0375	-0.0107	0.112
Fp1	Fp2	0.0514	0.0433	0.0151	0.12
P3	P4	0.0141	0.0231	0.0228	0.0122
F8	T6	-0.0535	-0.145	-0.0244	0.0303
F3	P3	-0.0546	-0.0442	-0.00099	0.106
C4	O2	-0.0101	-0.00826	-0.0283	0.0144
F4	T6	-0.0972	-0.0975	-0.0281	0.0866
P4	O1	-0.0102	0.009	-0.0395	-0.0122
F3	C4	-0.0973	-0.0316	-0.0244	0.102

**Table 7.1 Correlation change**  
(continued)

Channel 1	Channel 2	Time-window			
		1	2	3	4
F4	C3	-0.00953	0.00661	0.00599	0.067
F3	T5	-0.0126	-0.069	-0.0227	0.0733
C4	P3	-0.0123	-0.0248	0.00959	-0.0133
Fp1	T3	-0.0849	0.0586	-0.0652	0.0679
T4	O2	-0.0268	-0.0251	-0.00188	0.0277
Fp2	T4	-0.0678	-0.0857	-0.0307	0.0427
T3	O1	-0.0176	0.0879	-0.0369	0.158
C3	P4	0.0206	0.0824	-0.00977	0.0746
F7	P3	-0.0494	-0.0596	0.0114	0.0752
P3	O2	0.0363	0.0567	-0.00157	-0.0139
F8	P4	0.0861	-0.107	0.0217	0.0944
Fp2	F3	-0.00174	-0.00185	0.0125	0.0708
Fp1	F4	0.00087	0.0205	0.0439	0.118
F4	P3	-0.0685	-0.062	0.0335	0.00532
F3	P4	-0.0483	-0.0446	-0.0304	0.13
C4	O1	-0.00918	-0.0152	-0.0334	0.0331
C3	O2	-0.00445	0.11	-0.0265	-0.00191
Fp2	C3	-0.0645	-0.00685	-0.0509	0.107
Fp1	C4	-0.085	0.0197	0.00104	0.0995
F7	F4	0.00007	-0.0159	0.0529	0.0778
P3	T6	-0.0206	0.00385	-0.0047	0.00561
T5	P4	0.0288	0.0298	0.00792	0.0235
F3	O1	-0.0229	0.00238	-0.0644	0.117
Fp2	P4	-0.00016	-0.111	-0.0205	0.102
F3	F8	-0.0612	0.0105	0.0601	0.091
Fp1	P3	-0.0875	0.00788	0.00137	0.042
C3	T4	0.0154	0.0151	0.0848	0.193
T3	C4	-0.0183	0.0972	0.0229	0.0896
F4	O2	-0.0605	0.00691	-0.0171	0.0396
T6	O1	-0.0272	0.0194	-0.0301	0.0491
F4	T3	0.0133	0.0635	0.00908	0.12
T5	O2	0.0348	0.13	0.00202	0.0232
Fp2	T6	-0.105	-0.13	-0.0564	0.0558
T4	P3	-0.0065	-0.00125	0.0391	0.0969
F3	T4	-0.0543	-0.0875	0.0451	0.189
C4	T5	0.0119	-0.0568	0.0268	0.0746
C3	T6	-0.00639	0.0691	-0.0288	0.099
Fp1	F8	0.0125	0.0947	0.0677	0.106
F7	O1	-0.075	-0.0369	-0.0281	0.0364
Fp1	T5	-0.0537	0.0137	-0.0655	0.0173
T3	P4	0.038	0.122	0.0141	0.125
F8	O2	-0.0293	-0.0451	-0.0188	0.0311
F7	C4	-0.0619	-0.032	-0.0116	0.0918
F8	C3	-0.0274	0.00583	0.0453	0.115
Fp2	F7	-0.0366	0.0556	0.0518	-0.0174
F4	T5	-0.0468	-0.0473	0.0372	0.112
F3	O2	-0.0332	0.0268	-0.0483	0.0743
T4	O1	0.00576	-0.0383	0.0645	0.0918
F3	T6	-0.0863	-0.0607	-0.0482	0.171

**Table 7.1 Correlation change**  
(continued)

Channel 1	Channel 2	Time-window			
		1	2	3	4
Fp2	P3	-0.103	-0.0581	-0.0368	0.0848
Fp1	P4	-0.0382	-0.016	0.00464	0.0931
T3	O2	-0.031	0.0781	-0.0102	0.138
Fp2	T3	-0.0441	0.0237	-0.1	0.0606
F7	P4	-0.0472	-0.0795	-0.0124	0.0672
Fp1	T4	-0.0524	-0.0528	0.0446	0.169
F8	P3	-0.0321	-0.0746	0.0347	0.0634
F4	O1	-0.0364	-0.00472	-0.00471	0.0932
Fp2	O2	-0.0652	-0.0268	-0.01	0.0546
Fp1	O1	-0.0263	0.0122	-0.00973	0.0553
T5	T6	-0.00472	0.0925	-0.0389	0.119
T3	T4	-0.0139	0.0529	0.0931	0.2
F7	F8	-0.075	0.0471	0.0634	-0.056
T4	T5	0.00997	-0.0404	0.0897	0.148
T3	T6	-0.0304	0.088	-0.0365	0.175
F7	T4	-0.0367	-0.0668	0.142	0.0616
F8	T3	0.0237	0.0887	0.0467	0.102
Fp1	T6	-0.104	-0.0805	0.00677	0.0966
Fp2	T5	-0.0503	0.007	-0.0788	0.0464
F7	O2	-0.128	-0.0836	-0.0154	0.0117
F8	O1	-0.0607	0.0046	0.00886	0.118
Fp1	O2	-0.031	0.0142	0.0716	0.0185
Fp2	O1	-0.0317	0.0183	-0.0297	0.12
F7	T6	-0.137	-0.15	-0.0289	0.0806
F8	T5	-0.0424	-0.0089	0.0287	0.124

## 7.4 Discussion

In the picture condition correlations are more extensive, connecting more electrode sites, both within and between hemispheres, than in the control condition. This observation is consistent with the proposed description of correlated activity during visual discrimination, that the pattern of correlations should include an increasing number of cortical areas as the process of discrimination progresses to the moment of discrimination. Thus, substantial changes in correlation between the picture and control conditions occur between electrodes over almost all cortical regions. A number of observations might be made regarding the topographical distribution of the values of  $\delta r$ .

A general observation is that the distribution of  $\delta r$  varies profoundly between time-windows. Accepting the hypothesis that the neuronal activity underlying object discrimination is periodic in nature, then the time-scale of this neuronal activity can be roughly estimated as being no greater than the time interval from one time window to the next, 250 ms. This interval can be

compared with the finding in pilot work that significant changes in correlation were found only in the lowest frequency range of 2 to 8 Hz. These pilot results suggest that the period of the processes accessed by the correlation measurements is not much less than 125 ms, the period of an 8 Hz signal. While these pilot results should be confirmed by replication, it would appear that the periodicity of the neuronal events associated with discrimination lie within the range of from 125 ms to 250 ms.

A second observation can be made with respect to the possible duration of the events associated with object discrimination. Examination of Figure 7.6 suggests that the significant activity occurs between time-windows 2 and 4. This observation suggests that the subset of the neuronal events associated with object discrimination that are accessed by the correlation measurements occur within an interval of less than 1 second. On many trials, however, subjects did not blink to signal discrimination until some seconds after stimulus onset. At least 2 possibilities might account for this extra time between stimulus onset and discrimination. First, it might be that on some trials subjects needed the extra time to orient to the part of the image containing the camouflaged object. It was clear from debriefings that subjects were generally able to comply with the instruction to attend to the fixation point at all times. Subjects might therefore have been shifting their attentional focus to different parts of the image, while staying on the fixation point. A second possibility is that neuronal events contributing to the eventual discrimination, a fuse as it were, occur prior to time-window 2, but are characterized by relatively low values of correlation. This might be the case if, for example, such events involved relatively independent activity in neuronal populations smaller in cortical extent than the several square centimeters accessed by a single electrode. This second possibility is consistent with the CSO model. According to this model, a component of the process of discrimination involves analysis of low-level stimulus features, in correspondingly local cortical regions, and with a correspondingly low level of interregional signaling.

A third general observation is that the specific features of these distributions of  $\delta r$  vary to some extent with subject. This might be interpreted as indicating that subjects are complying with experimental instructions to varying degrees, or alternatively that subjects are following instructions, but that in so doing are nevertheless exhibiting to some extent unique patterns of correlated neural activity. In spite of this, to some extent, subject-specific nature of the patterns of correlations, the overall findings are that striking commonalities in the differences between the control and picture conditions do exist.

Examining Figure 7.6, a progression of events is clearly indicated. In the first time-window, 1, there is relatively little difference between picture and control conditions in terms of level of correlated activity.

In time-window 2 the activity evolves to include occipital and posterior temporal areas bilaterally, as well as the left anterior temporal area. This pattern in time-window 2 might be attributable to two effects. The first is the operation of the ventral visual pathway identified by Ungerleider and Mishkin (1982) as the primary cortical system involved with the process of object recognition. This pathway involves the occipital and inferotemporal cortices. The second effect is the accessing of a lexical associate or label for the object. As stated above, debriefing indicated that such naming prior to discrimination occurred for most subjects on some of the trials. This naming process would then be expected to involve speech and language, and therefore the participation of the sensory language area consisting of the auditory association cortex located in the left temporal region and adjacent parietal areas. Subjects were all right-handed and thus their language areas were most probably situated in the left temporal regions. These observations are consistent with the findings of the study by Petsche et al. (1992) which also found significant theta band coherences involving the left temporal region.

In time window 3, the pattern of correlations shifted to involve interhemispheric connections between frontal and temporal areas. This shift towards the involvement of frontal areas might be conjectured to indicate memory access processes associated with the frontal regions. At the same time, the correlations between occipital and temporal areas prevalent in time-window 2 are relatively diminished. The change in topography between time-windows 2 and 3 suggests a progressive sequence of events within the discrimination process. There might be an alternative explanation, however, in terms of the hypothesized recurrent nature of the discrimination process. This recurrent activity is proposed to involve successive cycles of a process that consists of feature analysis, feature transformation, and memory matching. On the basis of pilot analyses, the frequency of this periodicity is estimated to lie between 4 and 8 Hz, the frequency band that was used in the present analysis. The corresponding range of periods is therefore between 125 and 250 ms. It might be that the duration of the time-windows in the present analysis, 250 ms, is interacting with the periodicity of the neuronal activity involved in the discrimination process. In effect, these time-windows act like a stroboscope, preferentially presenting glimpses of portions of underlying neuronal activity that is synchronized with the period of the time-windows. Since the timing of the time-windows is not synchronized with such neuronal activity, the overall effect might be a 'beat' effect, an apparent frequency to the observed activity equal to the difference between the frequencies of the time-windows and the neuronal processes. While clearly this conjectured scenario is itself based on conjectures, the result of these effects would be that the apparent neuronal events within one time window would appear to be anomalously distinct from the activity in other time-windows.

In the final time window, 4, the pattern of correlations involves, as predicted by the CSO model, most cortical regions, and with somewhat more profuse correlations between than within

hemispheres. These observations of time-window 4 suggest that successful discrimination of complex visual images, depicting real-world objects, eventually needs to involve the correlated activity between all cortical regions, including occipital, frontal, temporal, parietal and central areas, and require significant levels of interregional signaling both within and between hemispheres.

In summary, these findings are generally consistent with the CSO model. During the task of object discrimination using visually complex, non lexical images, an increase in the degree of coupling is observed between almost all regions of the brain, both between and within hemispheres. The evolving pattern of interregional associations can be interpreted as an index of the rate of information interchange or signaling between multiple and relatively local cortical centers of coordinated activity, which over the course of the discrimination process increase in extent to include increasingly greater proportions of the cortex, in an iterated process of matching increasingly complex feature ensembles with the stored results of prior learning. When the elementary features of the stimulus are bound and transformed into a construct with which a sufficiently accurate memory match is possible, the elementary and discrete features of the stimulus may be said to have been transformed and bound into a unitary percept.

Examining Figure 7.1 through Figure 7.6, it is clear that there is a great deal of variability in the pattern of correlation differences across subjects. While all regions of the brain are connected by substantial levels of correlation, the time window at which the overall maximum amount of intercorrelation occurs varies between window 2 and window 4. There are at least two possible explanations for these observations, one having to do with the experimental paradigm, and the other having to do with deeper issues of inter-individual variability.

One possibility is that some portion of the variability is due to the fact that for each subject, the correlations represent a mean across trials with the underlying EEG signals aligned on the onset of the blink. That is, the location of the time windows is referenced to the eye blink. It appears possible that subjects, while attempting to comply with instructions, nevertheless did not always blink immediately after the target object was discerned. In that case, the peak of the correlations would appear at different time before the blink itself. Any individual variability in this reaction time would then show up as differences in the location of the maximum intercorrelations with respect to the blink.

A more interesting possibility is that the observed differences are attributable to inter-individual differences in neural function during the task of camouflaged object discrimination. Thus, it is possible that the differences in the pattern of intercorrelations observed in Figures 7.1 through 7.5 reflect corresponding differences in the way that different subjects' brains are wired. While at the behavioural level a common level of performance is observed, this performance may be subserved by significantly different neural organization, both in terms of 'hard-wired'

interregional connections, and in terms of how these connections transiently organize during task performance.

This issue recalls that one of Edelman's (1989) foundational postulates for his Theory of Neuronal Group Selection is that there does not exist, in any sufficiently complex neuronal system, a precise point-to-point wiring scheme. Rather, he observes, there is evident a significant degree of individual variability in the configuration of neuronal interconnections. This variability in turn provides the substrate that allows the process of neuronal group selection to "differentially amplify" particular variants within neuronal populations. As stated earlier, Edelman (e.g., 1989) suggests that neuronal groups, circuits composed of multiple interconnected neurons, are an appropriate level to consider neuronal systems. Such neuronal groups are seen to be highly variable in terms of their internal wiring configuration. Further, such variability extends to inter-group, and eventually to interregional connection patterns. On a general level it is obvious that commonalities in interregional connectivity do exist, as evidenced by anatomical structures such as commissures, tracts and projections. When patterns of connectivity are examined on ever more detailed levels, correspondingly greater degrees of inter-individual variability are observed. Edelman (e.g., 1989) suggests that there exists competitive activity among neuronal groups, activity that leads to a process of selection in which some groups survive and function at the expense of other groups. Importantly, it is the nature of an individual's interactions with the environment that determines the evolution of such neuronal groups. Those groups are selected for whose activity is reinforced as a result of such environmental interactions. As such interactions can be guaranteed to be, in detail, highly variable between individuals, the pattern of connectivity within neuronal groups will correspondingly be variable. Furthermore, and as suggested by the present results, such variability may extend to interregional signaling configurations. Again, while on a gross level, patterns of anatomical connections are relatively, although perhaps not absolutely, constant across individuals, patterns of interregional signaling may be strongly dictated by intra-group wiring. For different individuals performing a common task, the same pattern of large-scale anatomical connections may thus support a wide range of patterns of signalling between cortical regions, because of the individually-specific patterns of small-scale connectivities within and between neuronal groups.

In sum, the present findings are consistent with Edelman's (e.g., 1989) foundational proposition that environmental interactions determine the selection of neuronal groups and therefore the fine structure of neuronal connectivity.

## 8 Topographical Distribution of Net Correlations

### 8.1 Introduction

This section will attempt to answer the following question. For each of the 4 time-windows of the before-blink epoch, to what extent is the oscillatory activity at each of the electrode sites related to the oscillatory activity at all other electrode sites in terms of cross-correlation? In other words, for each time-window and for each electrode site, what is the average correlation between the signal from that electrode, and the signals from all other electrode sites? According to the predictions made at the outset of this study, successful discrimination should ultimately involve communication between all cortical regions, and therefore a high level of correlation between the signal at any one electrode site and the signals at all electrode sites. The existence of the numerous tracts, commissures and projections interconnecting all cortical regions provides a physiological substrate for such communication to occur.

### 8.2 Method

The average correlation difference  $\delta r_k'$  for each electrode  $k$  is computed, for each of the 4 time windows of the before-blink epoch, by averaging over the correlations between that electrode and all 15 other electrodes:

$$\delta r_k' = 1 / n \{ \sum [ r(\text{picture})_i - r(\text{control})_i ] \}$$

where  $n$  is the total number of electrodes - 1, and the sum is taken over  $i (\neq k) = 1$  to  $n$ . For each of the 4 time-windows the result is a set of 16 numbers, each of which specify the mean correlation a single electrode and the 15 other electrodes.

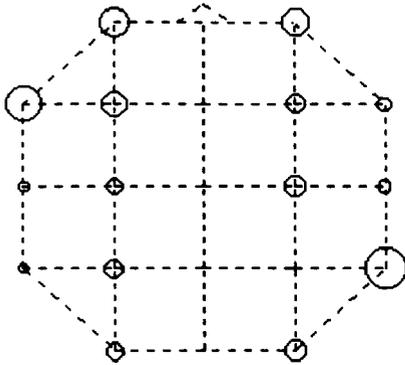
### 8.3 Results

Figure 8.1 shows the distribution of  $\delta r'$  for a typical subject, subject 5. Figure 8.2 shows the distribution of  $\delta r'$  averaged over subjects. In these figures the diameter of the circles at each electrode position is proportional to the value of  $\delta r'$ . A circle diameter equal to the grid spacing in the figures corresponds to a value of  $\delta r'$  of 0.12. A filled circle at an electrode position represents a positive value of  $\delta r'$ , and thus a greater average correlation in the picture than in the control conditions between that electrode and all other electrodes. Empty circles represent corresponding negative values of  $\delta r'$ . All figures are drawn to the same scale. These correlations are listed in Table 8.1. A substantial difference can be seen in the distribution of  $\delta r'$  between each of the 4 time windows. This distribution will be outlined for each time window.

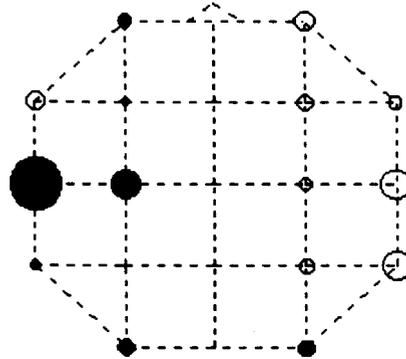
1. Time-window 1: there are net negative values of  $\delta r'$  over most areas, with only small positive values over the left temporal area (T3, T5).
2. Time-window 2: positive values of  $\delta r'$  occur mainly in the left hemisphere. The largest positive values are found over the left fronto-temporal (T3), and the left central (C3) areas. Smaller positive values are found over the occipital areas bilaterally (O1, O2), and the left posterior temporal area (T5). Negative values occur over all other regions.
3. Time-window 3: Positive values of  $\delta r'$  are distributed more bilaterally in this time-window, with the largest positive values over left frontal (F7), and the right fronto-temporal (F4, F8, T4) areas.
4. Time-window 4: Positive values of  $\delta r'$  occur over all regions. Largest positive values are found over the left fronto-temporal and central (F3, T3, C3), and right fronto-temporal (T4) areas.



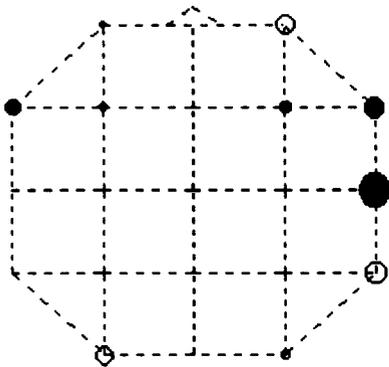
(a) time-window 1



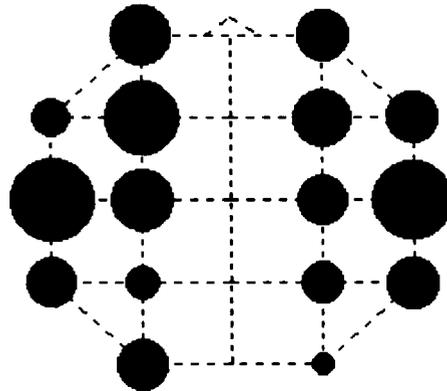
(b) time-window 2



(c) time-window 3



(d) time-window 4



**Figure 8.2 Topography of net difference correlations, average across subjects**

The graphs show the distribution over the scalp of the net difference correlations for each electrode site, averaged across subjects. Difference correlation is computed for each pair of electrodes as  $r(\text{picture}) - r(\text{control})$ . The net difference correlation for any one electrode is the sum of the correlation differences between that electrode and all 15 other electrodes, divided by 15. Circle diameter indicates the magnitude of the net difference correlation. A diameter equal to the size of the map grid corresponds to a value of 0.12. Empty circles indicate a negative net difference correlation, and filled circles indicate a positive value.

**Table 8.1 Average correlation difference**

The average correlation difference is computed for each electrode and each time window by calculating the difference correlation between an electrode and all other electrodes, and dividing the result by 15. Difference correlation are computed by subtracting the correlation in the control condition from the correlation in the picture condition.

Channel	Time-Window			
	1	2	3	4
Fp1	-0.0378	0.0183	0.00818	0.0784
Fp2	-0.0353	-0.0252	-0.0242	0.0669
F7	-0.0457	-0.025	0.018	0.0479
F3	-0.0327	-0.00857	-0.00896	0.097
F4	-0.0245	-0.0222	0.0135	0.0741
F8	-0.0177	-0.0149	0.026	0.0664
T3	-0.0127	0.0678	0.00077	0.11
C3	-0.0224	0.0377	-0.00535	0.0816
C4	-0.025	-0.0153	-0.00458	0.0631
T4	-0.014	-0.0401	0.041	0.104
T5	-0.0112	0.0119	-0.00066	0.0658
P3	-0.0244	0.00272	0.00664	0.0426
P4	0.00396	-0.0184	-0.00536	0.055
T6	-0.0531	-0.0383	-0.0277	0.0677
O1	-0.0224	0.0218	-0.0225	0.0664
O2	-0.0267	0.0202	-0.0108	0.0272

## 8.4 Discussion

The results of this analysis of the distribution of average correlation differences,  $\delta r'$ , are generally consistent with the results of the previous analysis, the distribution of intercorrelation differences. The present results demonstrate that object discrimination involves, in a short time interval preceding the moment of discrimination, the progressive involvement of various cortical areas, with clear differences in the distribution of  $\delta r'$  between each of the 4 time windows. In time-window 1, there is little difference in average correlation between the picture and control conditions, at any of the electrodes. In time-window 2 the left fronto-temporal and central areas, along with the occipital areas bilaterally and the left posterior temporal areas, all show the largest positive values of  $\delta r'$ , indicating a relatively larger correlation in the picture than in the control condition, between these areas and all other cortical areas. This observation, consistent with the results of the previous analysis, the distribution of correlation differences, suggests that within time window 2 the neuronal systems that are operating may include the ventral occipital-inferotemporal visual processing pathway, and the sensory language area in the left temporal region. In time-window 3, a more generally bilateral pattern of small positive values of  $\delta r'$  indicates an increasing level of communication between frontal and temporal areas in the left

and right hemispheres. In time-window 4, and also consistent with the results of the analysis of correlation differences, the topographic distribution of average correlation differences similarly indicates that visual discrimination eventually involves increases in correlation between all areas of the cortex. This evolving pattern of average correlation differences supports the prediction of the CSO model that object discrimination should involve progressively larger cortical extents, beginning with the occipital and temporal regions, and ultimately spreading to include most cortical regions.

## **9 Mutual Information Analysis**

### **9.1 Introduction**

There are two motives for considering mutual information as a measure to be investigated in the present study of visual discrimination. The first motive involves the conceptualization, briefly mentioned earlier, of the neural processes during perception as involving a changing pattern of information interchange between cortical systems. Mutual information more directly than correlation addresses this issue of information interchange. Mutual information, like correlation, is a measure which is defined for a pair of variables. In informational terms, mutual information is a measure of how much information about one variable can be predicted by making a measurement of the second variable. In terms of the EEG time series that are the subjects of analysis in the present study, the mutual information between two time series, each recorded from one electrode, is an estimate of how much information about one of these time series is available from the second time series. The suggestion is now made that by extension, mutual information, calculated for each of the 4 time windows of the before and after-blink epochs, for each pair of electrodes, is related to the rate of information transfer between the cortical areas accessed by those electrodes.

The second motive involves the nature of mutual information as a statistic. Pearson product-moment correlation, used in the computation of cross-correlations, estimates the strength of linear relationship between two variables. In the present context, when computed for the signals from a pair of electrodes, this measure of correlation, or more precisely the square of the correlation, is a measure of the extent to which the signal at one of the electrodes can be predicted by means of a linear function of the signal measured at the second electrode. In contrast, the mutual information function estimates the strength of a general relationship between two variables, without the restriction of linearity. Again in the present context, when computed for the signals from a pair of electrodes, mutual information would be a measure of the degree to which the signal at one electrode could be predicted by means of an arbitrary, and not necessarily linear, function of the signal from the second electrode. By comparing the results of the mutual information analysis with the results of the correlation analysis, it may be possible to estimate the extent to which the relationship between the activity of the different cortical regions can be considered to be reasonably well modeled by a linear process. A practical limitation in this respect is that the two measures, correlation and mutual information, are not equivalent in terms of the number of samples of data they require to produce stable results. In particular, because of the way in which mutual information is calculated, as described in the following section, it would appear that a greater number of data points is needed in order to

calculate a stable estimate of mutual information, as compared with the number needed for a stable estimate of correlation.

## 9.2 Method

The concept of mutual information can be developed in terms of the concept of entropy. Entropy is a measure of the average amount of information that is available from a single measurement of a variable. Entropy, and mutual information, have their origins in information theory (Shannon, 1948), and for this reason discussions of these quantities involve the concept of a message. A message, in terms of measurements made of a dynamical system, can be considered to be equivalent to the range of a set of values, a range within which the measurements made on such a system may lie. Consider a system a variable of which produces any one of  $n$  different messages, or equivalently  $n$  ranges of values. Furthermore, each of these messages, or ranges of values of the variable, has a probability  $p_i$  of occurring. The entropy of such a system, as estimated by the measurements made of the variable, is defined as

$$H = - \sum p_i \log p_i$$

In concrete terms, and applied to the time-series of the EEG, each of these messages is some range of values of the voltage measurement. Consider for example that the total range of voltage measurements is -80 microvolts to +80 microvolts. This total range can be divided into a number of sub-intervals, such as for example 8, 10 microvolt intervals, starting with -80 to -70 microvolts, and ending with +70 microvolts to +80 microvolts. The EEG time-series is then binned by assigning each data point of the time-series to one of these 10 intervals or bins. This process essentially constructs a discrete frequency distribution from the time-series data samples. In terms of the definition of entropy, each one of these 10 intervals is one possible message from the system being measured. Next, a probability  $p_i$  is assigned to each of these intervals or messages. This probability is, for each interval, the probability that the time-series has a voltage value within the interval. The result now is the discrete probability distribution for the original time-series. Next, each of these probabilities is multiplied by the logarithm of the probability, forming the products  $p_i \log p_i$ . Finally, these products are summed over the total number of intervals, in this example, 10. The result is an estimate of the entropy of the system, the average amount of information derived from a single measurement made on the system. When the logarithm is taken to base 2, the units of entropy are bits.

Mutual information is defined in terms of entropy (Fraser and Swinney, 1986; Gray, 1990). Consider 2 systems  $S$  and  $Q$ , each generating messages  $s_i$  and  $q_k$  as above. A value of entropy can be defined for both of these systems,  $H(S)$  and  $H(Q)$ . Next, the concept of entropy

can be extended to include the case where a pair of measurements,  $(s_i, q_k)$  is made simultaneously from these two systems. The joint entropy,  $H(S, Q)$  is the amount of information available from this single pair of measurements of systems S and Q. The notion of joint entropy can be developed in terms analogous to those used for the entropy of a single system. The pairs of measurement  $(s_i, q_k)$  are first binned. Continuing the example above, each of the individual measurements from systems S and Q are placed into intervals of 10 microvolts, beginning with the -80 to -70 microvolt interval and ending with the 70 to 80 microvolt interval. These separate intervals are combined to form a discrete joint frequency distribution of  $n$  by  $n$  bins. The first bin for example contains those measurements for which messages  $s_i$  and  $q_k$  both fall within the range of -80 to -70 microvolts, and so on. As before, for each of these bins a probability  $p_{i,k}$  is computed, creating a discrete joint probability distribution for the pair of time series. Finally, the sum of the products of these probabilities and their logarithms is accumulated. The joint entropy of systems S and Q is then

$$H(S, Q) = - \sum p_{i,k} \text{Log } p_{i,k}$$

Mutual Information is then defined in terms of the individual entropies of systems S and Q, and their joint entropy, as the sum of the individual entropies minus their joint entropy:

$$I(S, Q) = H(S) + H(Q) - H(S, Q)$$

In the present analysis, all logarithms are taken to base 2, so that the resulting values of mutual information represent the number of bits of information that can be predicted about one time series, from a measurement made on a second time series.

A partial limitation in applying mutual information to the present data is that, in order to form a reasonable estimate of the discrete frequency distributions for the variables, more data points may be needed than are required in order to compute the corresponding Pearson correlation.

The mutual information computations as well as all supporting functions were carried out using the data analysis program Simulnet™ version 2.3.

### 9.3 Results

The three predictions made on the basis of the CSO model were generally confirmed by the results. Partially confirming the first prediction, there was one, marginally significant, effect of time, and this occurred in the picture condition in the BBE ( $F = 2.24$ ,  $p = 0.08$ , effect size = 0.0002). Mutual information increased over the duration of the BBE, from 0.690 bits to 0.728 bits. In contrast, in the control condition mutual information remained relatively constant over the BBE, decreasing slightly and non-significantly from 0.685 bits to 0.673 bits ( $F < 1$ ). Table 9.2

lists the mean mutual information values for each time window, averaged over all distances, that is, over all 120 possible pairwise electrode pairs. These values as a function of time-window are graphed in Figure 9.1a.

Confirming the second prediction, there were significant effects of distance in both the BBE and ABE, and both in the picture condition (BBE:  $F = 15.5$ ,  $p < 0.0001$ , effect size = 0.08; ABE:  $F = 18.1$ ,  $p < 0.0001$ , effect size = 0.09) and in the control condition (BBE:  $F = 24.6$ ,  $p < 0.0001$ , effect size = 0.10; ABE:  $F = 19.6$ ,  $p < 0.0001$ , effect size = 0.08) conditions, with mutual information decreasing with increasing distance in all cases. Mutual information decreased from approximately 0.78 bits for adjacent electrodes to approximately 0.61 bits for electrode pairs spaced furthest apart. These values of mutual information as a function of inter-electrode distance are graphed in Figures 9.2a for the before-blink epoch and Figure 9.2b for the after-blink epoch. Table 9.1 shows the results of the analysis of variance, listing the values of  $F$  along with the corresponding values of probability and effect size.

Confirming the third prediction, in the picture condition, the values of mutual information between the 12 most closely-spaced electrode pairs increased from 0.776 bits to 0.808 bits over the BBE ( $F = 1.71$ ,  $p = 0.17$ , effect size = 0.001). Over this same interval the values of mutual information between the 12 most distantly-spaced electrode pairs increased from 0.596 bits to 0.652 bits ( $F = 2.25$ ,  $p = 0.086$ , effect size = 0.002). Since these secondary analyses of variance were conducted on subsets of the data upon which the original analysis of variance was conducted, it is not expected that the significance probabilities may need the corrections normally required when multiple tests of significance are conducted on the same data. These results are shown in Table 9.3, and graphed in Figure 9.1b.

**Table 9.1 Results of analysis of variance of mutual information**

The table shows the results of a two-way within subjects analysis of variance of mutual information. Significant effects of distance occur for all conditions and epochs. A marginally significant effect of time occurs in the before-blink epoch in the picture condition only.

Condition	Epoch	Item	Time	Distance	T x D
Picture	BBE	F	2.24	15.5	< 1
		p Effect	0.08	< 0.0001	0.08
	ABE	F	1.79	18.1	< 1
		p Effect	0.15	< 0.0001	0.09
Control	BBE	F	< 1	24.6	1.01
		p Effect		< 0.0001	0.11
	ABE	F	1.22	19.6	< 1
		p Effect	0.31	< 0.0001	0.08

**Table 9.2 Mean mutual information**

The table shows mutual information, averaged over all trials.

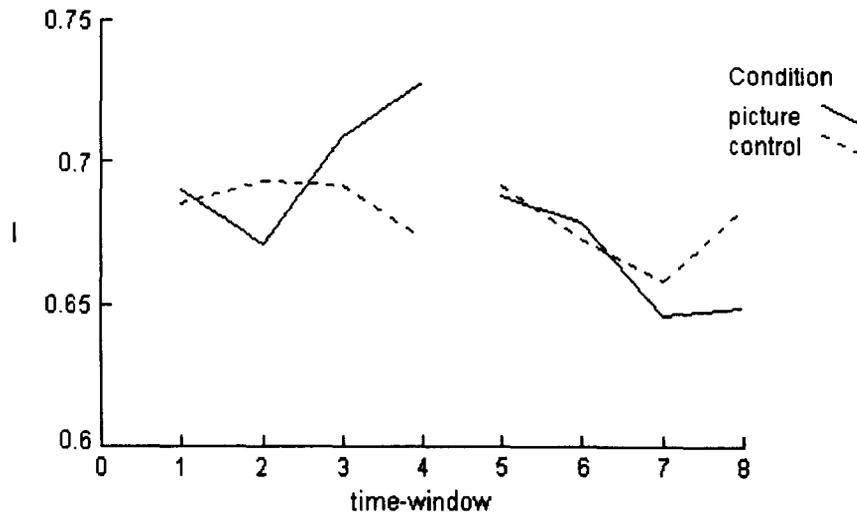
Epoch	Window	Control	Picture
BBE	1	0.685	0.690
	2	0.693	0.671
	3	0.692	0.709
	4	0.673	0.728
ABE	5	0.692	0.688
	6	0.673	0.679
	7	0.658	0.646
	8	0.684	0.649

**Table 9.3 Short vs. long mutual information**

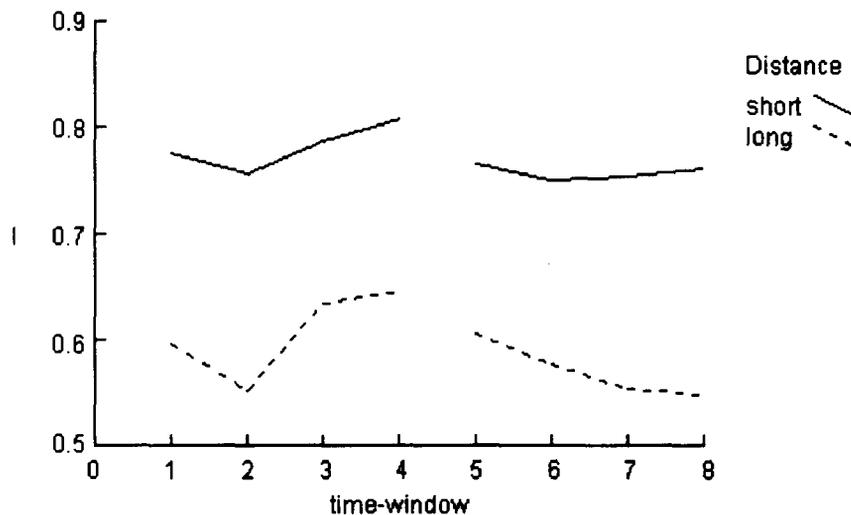
The table shows mutual information, averaged over all trials, for the 1 shortest (Short) and the 12 longest (Long) between electrode distances. Mutual information between long-distance electrode pairs increase more from windows 1 to 4 than mutual information between short-distance electrode pairs.

Window	Distance	
	Short	Long
1	.776	.596
2	.756	.552
3	.788	.634
4	.808	.645
F	1.71	2.25
p	0.17	0.08
Effect	0.001	0.002

(a) mutual information averaged over all distances



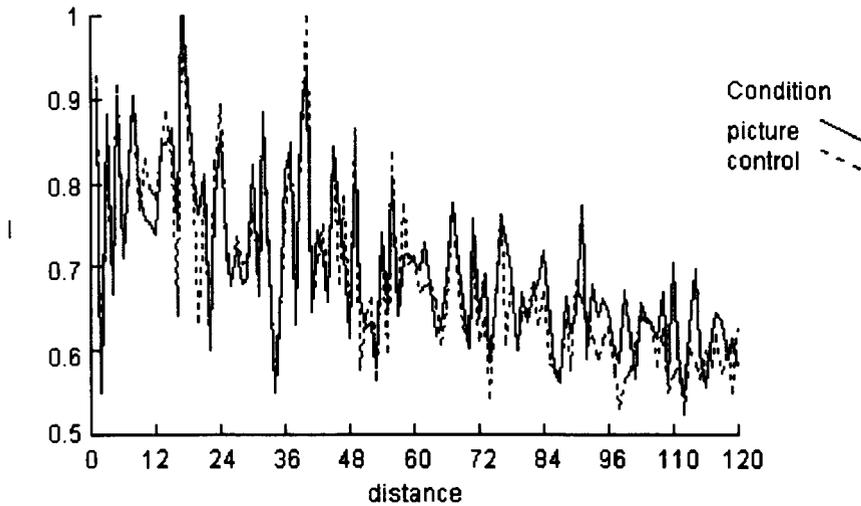
(b) mutual information for short and long inter-electrode distances, picture condition



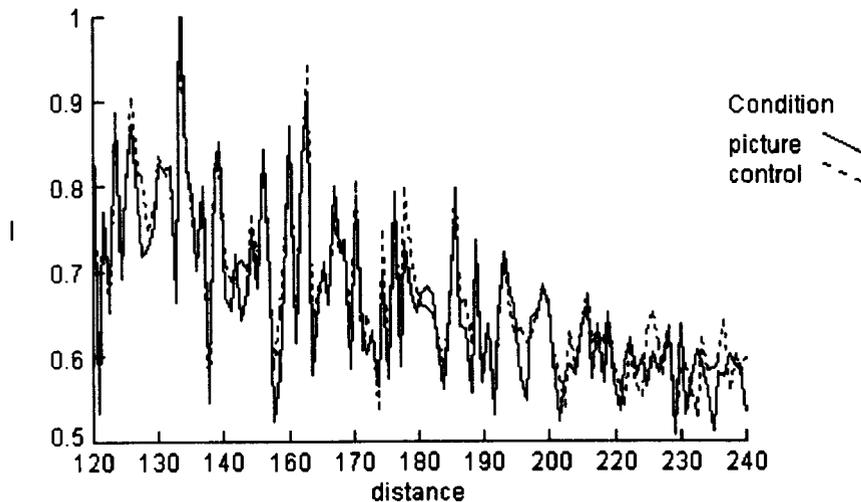
**Figure 9.1 Mutual information vs. time**

(a) Mutual information vs. time window in the before and after-blink epochs. Mutual information values are averaged across all 120 possible electrode pairs, across cases and across subjects. There is a marginally significant increase in mutual information between time windows 1 and 4 in the picture condition. (b) Short distance mutual information values in the picture condition are averaged over the 12 electrode pairs separated by the shortest distance. Long distance values are averaged over the 12 electrode pairs separated by the longest distances. Long distance mutual information values increase significantly from time-window 1 to 4, while short distance values remain relatively constant.

(a) before-blink epoch



(b) after-blink epoch



**Figure 9.2 Mutual information vs. inter-electrode distance**

Mutual information decreased significantly with distance for both the before-blink epoch (windows 1 to 4) and after-blink epoch (windows 5 to 8), and in both the picture and control conditions. Mutual information values are averaged across subjects and across the 4 time windows of the before-blink epoch (a) and the after-blink epoch (b). Values are shown for all 120 possible electrode pairs of the 16 electrodes that were recorded. Each point on the graphs shows the mutual information for one pair of electrodes. Electrode pair 1 is part of the group of most closely spaced electrode pairs. Electrode pair 120 is part of the group of most distantly spaced pairs.

## 9.4 Discussion

The values of mutual information varied in much the same way, as a function of time and inter-electrode distance, as the values of correlation. First, mutual information increased, although only marginally significantly, over the 1 second interval preceding discrimination. In terms of the interpretation of mutual information as an estimate of the number of bits of information that can be predicted about one process from a measurement on a second process, these results could be restated as an increase, over the before-blink epoch, in the rate of information exchange between the cortical systems accessed by the recorded electrodes. Second, mutual information decreased with increasing inter-electrode distance. Third, mutual information not only increased with time up to the moment of discrimination, but increased by a greater amount for more distantly spaced electrode pairs than for more closely spaced pairs. All of these findings are in accord with the predictions made on the basis of the CSO model for the expected behaviour of interregional associations.

Given this similarity between the results of the cross-correlation analysis and the mutual information analysis, one conclusion that can be tentatively drawn is that the relationship between the signals recorded from pairs of electrodes appears to be well modeled as a linear function. That is, the findings of the mutual information analysis are reasonably consistent with the conclusion that the activity of the cortical system accessed by any one electrode is at least approximately linearly related to the activity of any other such cortical system.

## 10 Coherence Analysis

### 10.1 Introduction

The intent of this section is to repeat the analysis that was carried out using cross-correlation and lag, using alternative measures of association, coherence and phase. Coherence, in general terms, is a measure of association, which can be computed for a pair of time histories; that is, the frequency domain analog of squared cross-correlation. More particularly, coherence is the cross-correlation between two complex Fourier power spectra that are computed for two time series; that is, the cross spectral density (CSD), averaged over some range of frequencies. As such, coherence is an estimate of the amount of shared power, or variance, within that frequency range, between the two time series. In alternative terms, coherence represents the proportion of the power, or variance, within some specified frequency band, in one time series that can be accounted for by a linear function of the other time series (Otnes and Enochson, 1972, 1978). In these terms, the analogy with squared cross-correlation, the proportion of the variance of one variable that can be accounted for by a linear function of a second variable, becomes evident. While cross-correlation is defined in the time domain, coherence is analogously defined in the frequency domain. The second statistic that will be computed is phase, a measure analogous to lag that represents an estimate of the difference in phase angles between the periodic components within the two time-series, averaged over some range of frequencies.

One motivation for using coherence and phase analysis in the present study, since cross-correlation and phase have already been computed, is to attempt to connect some of the present results with the results of earlier studies that have used coherence and phase analysis of the EEG.

EEG amplitude and coherence changes related to the different thinking processes involved in the visualization of an abstract concept and the interpretation of a painting were found in a study by Petsche, Lacroix, Lindner, Rappelsberger and Schmidt-Henrich (1992). This study investigated the question of whether changes in brain function would be found corresponding to the difference between a self-generated mental image and a mental image generated from a prior perception. In one task, subjects were asked to generate a mental image corresponding to an abstract concept, a task expected to involve thinking with images. In a second task, subjects were asked to interpret a painting viewed before the recording session, a task expected to engage thinking with language. EEG recordings were analyzed using measures of amplitude and coherence.

A complex pattern of coherence changes was found. In the abstract visualization task, coherence increases included the left frontal and central regions, and right frontal, central, and temporal regions in the beta bands. Coherence decreases included the right frontal and temporal areas in the theta and alpha bands. In the painting interpretation task, coherence increases included the left frontal, central, temporal and parietal areas in the theta band, the left central area in the alpha band, and left frontal areas in all beta bands. Coherence decreases included the right anterior region in the alpha band, and the right posterior area in all beta bands. These results were interpreted as suggesting that the differences in mental processes associated with self-generated and perceptually-inspired mental images were reflected most consistently in activity over frontal regions, and that mental imagery involves connections between multiple, widespread, cortical regions.

Rappelsberger and Petsche (1988) had similarly found that EEG coherence and amplitude changes were affected by a mental visualization task, cube rotation. Subjects were shown a cube which they were then asked to visualize rotating. A complex pattern of coherence increases was found that included all cortical regions, but that was to some extent different for males and females. However, a finding common to both females and males was a coherence increase in all frequency bands between left and right parietal areas. The authors suggest that degree of coherence between cortical regions may be related to functional couplings between these areas.

Thatcher et al. (1986) computed coherence and phase for resting EEG's recorded from a sample of 189 children with ages from 5 to 16 years. They found that coherence decreased approximately quadratically with increasing inter-electrode distance, while lag increased approximately quadratically with increasing distance, results which led the authors to conclude that EEG coherences were determined mainly by axonal rather than volume conduction.

In terms of the present study, it is expected that coherence and phase should show the same dependence on inter-electrode distance, since again, the inter-regional associations are proposed to be determined primarily by the effects of axonal conduction, and not by volume conduction. It is expected as well that the behaviour of coherence and phase should generally echo that of cross-correlation and lag, since cross-correlation and lag analyses, like coherence and phase analyses, are based on the behaviour of oscillatory signal components. In particular, it is expected that coherence should increase with time over the before-blink epoch, reflecting increasing synchronization, and hence increasingly similar power spectra, between signals recorded over multiple cortical regions. It is expected that phase should increase with inter-electrode distance, reflecting the fact that the observed inter-regional associations are the result of signal transmission along axons, with the associated relatively large change in level of

association with distance, rather than as a result of volume conduction with correspondingly small rate-of-change of association with distance.

One partial limitation in applying coherence and phase analysis to the present data is the relatively limited number of data points, 32, available in each time-window. This small number of points results in a power spectrum with a correspondingly limited number of discrete frequency points. It is known (Otnes and Enochson, 1972) that cross spectral density estimates are distributed approximately as chi-squared variables, with a standard error of estimate given by  $e = 1 / \sqrt{n}$ , where  $n$  is the number of individual frequency values averaged over in computing the CSD. Thus, in the present case, the 32 data points in the original time-series are used to generate 16 discrete frequency values in the CSD. Standard error is then 25%. In the previous analysis of cross-correlation, correlation varied as a function of time by approximately 10% (from approximately 0.5 to 0.61). Using these values as a rough guide, and assuming that the magnitude of the coherence effect is of the same order as the magnitude of the cross-correlation effect, it is predicted that it may not be possible to detect the time-related variation in coherence, and consequently, in phase. Under the same assumption, the distance-related variation in coherence should be detectable. Cross-correlation varied as a function of distance by approximately 60% (from approximately 0.8 to 0.2).

## 10.2 Method

The procedure used in this coherence and phase analysis duplicated exactly the procedure that was used earlier in the cross-correlation and lag analysis, except that coherence and phase computations were substituted for the cross-correlation and lag computations.

The coherence computation first involves computing complex Fourier spectra for the two time series,  $X(f)$  and  $Y(f)$ . Next, from these Fourier spectra, the following quantities are computed:

$$\text{Power spectral density of } X(f): G_x(f) = (2 / n) |X(f)|^2$$

$$\text{Power spectral density of } Y(f): G_y(f) = (2 / n) |Y(f)|^2$$

$$\text{Cross power density: } G_{xy}(f) = (2 / n) [X^*(f) Y(f)]$$

where  $| \cdot |$  denotes the absolute value and  $*$  denotes the complex conjugate. From the cross power density, the cospectra,  $C_{xy}(f)$ , and quadspectra,  $Q_{xy}(f)$  are computed using the relation

$$|G_{xy}(f)|^2 = C_{xy}(f)^2 + Q_{xy}(f)^2$$

The cospectra and quadspectra represent, respectively, the real and imaginary components of the cross power density. Next, coherence,  $\gamma$ , and phase,  $\phi$ , are computed for

each frequency component. Coherence is computed by dividing the squared absolute value of the cross power density by the power spectral densities of the two time series, a normalizing operation. Phase is computed by calculating the inverse tangent of the ratio of the quadspectrum to the cospectrum.

$$\gamma(f) = |G_{xy}(f)|^2 / G_x(f) G_y(f)$$

$$\phi(f) = \arctan (Q_{xy} / C_{xy})$$

Finally, smoothed values of coherence and phase are computed, by averaging over a range of  $n$  frequency components.

$$\text{Coherence} = (1 / n) \sum \gamma_i$$

$$\text{Phase}(f) = (1 / n) \sum \phi_i$$

In the present case, the average was computed over the 2 to 8 Hz frequency range. The resulting value of coherence represents the average cross-correlation between the power spectra of the two time series, normalized by dividing by the respective power spectral densities for the two individual time series. The values of phase are specified in degrees.

The coherence and phase computations as well as all supporting functions were carried out using the data analysis program Simulnet™ version 2.3.

### 10.3 Results

In terms of coherence, there were significant effects of distance in both the BBE and ABE, and in both the picture (BBE:  $F = 22.8$ ,  $p < 0.0001$ , effect size = 0.11; ABE:  $F = 22.6$ ,  $p < 0.0001$ , effect size = 0.11) and control (BBE:  $F = 32.6$ ,  $p < 0.0001$ , effect size = 0.13; ABE:  $F = 30.0$ ,  $p < 0.0001$ , effect size = 0.12) conditions, with coherence decreasing with increasing distance in all cases. Coherence decreased from approximately 0.75 for adjacent electrodes to approximately 0.5 for electrode pairs spaced furthest apart. These mean correlations as a function of inter-electrode distance are graphed in Figures 10.1a for the before-blink epoch and 10.1b for the after-blink epoch. Table 10.1 shows the results of the analysis of variance, listing the values of  $F$  along with the corresponding values of probability and effect size. Table 10.2 shows the corresponding values of mean coherence for each condition and time window.

In terms of phase, there were a number of small but significant effects of distance. In the picture condition phase decreased in the before-blink epoch from 2.33 degrees for short distances to 0.34 degrees for long distances ( $F = 1.92$ ,  $p < 0.0001$ , effect size = 0.005). In the control condition phase decreased in the before-blink epoch from 1.90 degrees to 1.09 degrees ( $F = 1.41$ ,  $p = 0.002$ , effect size = 0.002), and increased in the after-blink epoch from 1.69

degrees to 2.80 degrees ( $F = 1.49$ ,  $p = 0.0005$ , effect size = 0.002). There were no significant effects of time on the magnitude of phase. These results are shown in Table 10.3.

**Table 10.1 Results of analysis of variance of coherence**

The table shows the results of a two-way within subjects analysis of variance of coherence. Significant effects of distance occur for all conditions and epochs.

Condition	Epoch	Item	Time	Distance	T x D
Picture	BBE	F	< 1	22.8	1.0
		p Effect		< 0.0001 0.11	0.21 0.0
	ABE	F	1.93	22.6	< 1
		p Effect	0.13 0.0001	< 0.0001 0.11	
Control	BBE	F	< 1	32.6	1.2
		p Effect		< 0.0001 0.13	0.013 0.002
	ABE	F	< 1	30.0	1.1
		p Effect		< 0.0001 0.12	0.19 0.0009

**Table 10.2 Mean coherence**

The table shows mean coherence for each time-window, averaged over all trials.

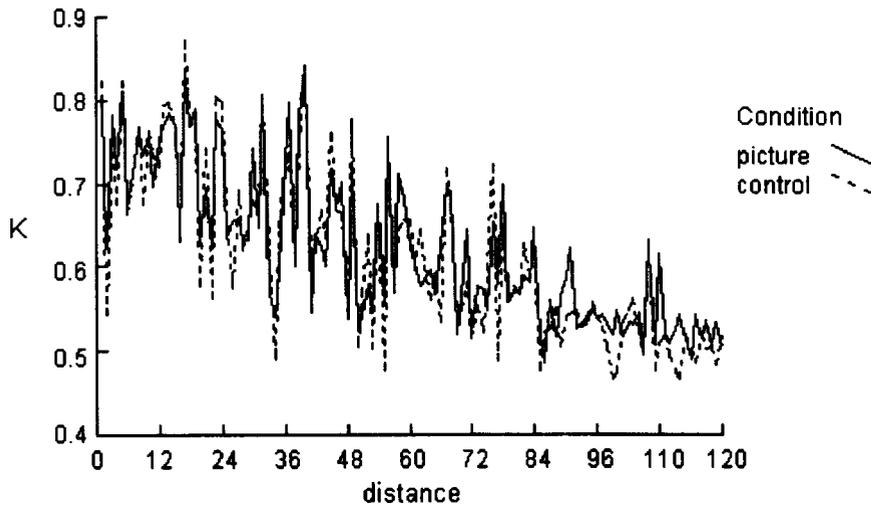
Epoch	Window	Control	Picture
BBE	1	0.626	0.625
	2	0.604	0.628
	3	0.612	0.608
	4	0.616	0.629
ABE	5	0.618	0.614
	6	0.621	0.633
	7	0.632	0.586
	8	0.62	0.594

**Table 10.3 Results of analysis of variance of phase**

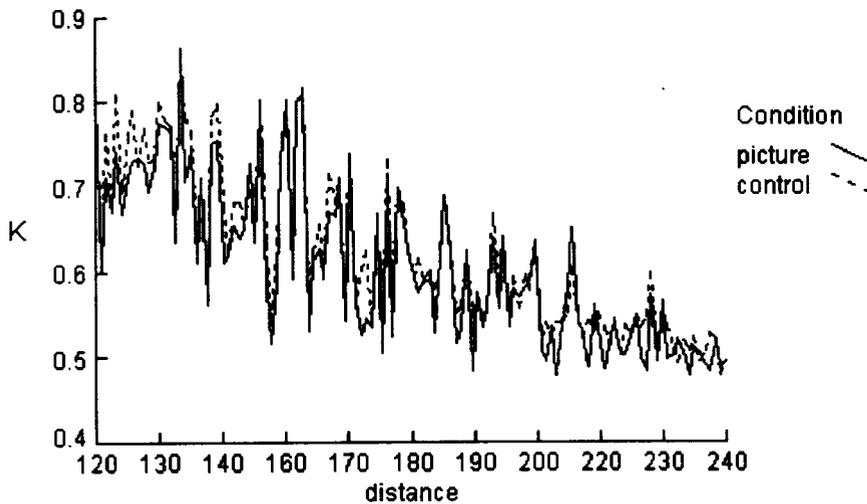
The table shows the results of a two-way within subjects analysis of variance of phase. Phase decreased significantly with increasing distance in the picture and control conditions in the before-blink epoch, and increased significantly with distance in the control condition in the after-blink epoch.

Condition	Epoch	Item	Time	Distance	T x D
Picture	BBE	F	< 1	1.92	1.05
		p Effect		< 0.0001 0.005	0.24 0.0009
	ABE	F	1.05	1.10	< 1
		p Effect	0.37 0	0.22 0.0006	
Control	BBE	F	1.25	1.41	< 1
		p Effect	0.29 0	0.002 0.002	
	ABE	F	< 1	1.49	< 1
		p Effect		0.0005 0.002	

(a) before-blink epoch



(b) after-blink epoch



**Figure 10.1 Coherence vs. Inter-electrode distance**

Coherence values are averaged across subjects and across the 4 time windows of the before-blink epoch (a) and the after-blink epoch (b). Values are shown for all 120 possible electrode pairs of the 16 electrodes that were recorded. Each point on the graphs shows the coherence for one pair of electrodes. Electrode pair 1 is part of the group of most closely spaced electrode pairs. Electrode pair 120 is part of the group of most distantly spaced pairs.

#### 10.4 Discussion

Averaged over all time-windows within an epoch, coherence varied significantly with inter-electrode distance. For both the before and after-blink epochs, at short inter-electrode

distances, corresponding to adjacent electrode positions, coherence magnitude was approximately 0.75, while at the longest inter-electrode distances coherence dropped to approximately 0.5. These findings are consistent with the findings of Thatcher et al. (1986), who similarly found, in resting EEG, an inverse relationship between between-electrode distance and coherence. Examining Figures 10.1a and 10.1b the change in coherence with distance appears to be approximately quadratic. This observation is only approximate however, since a simplified distance metric was used, based on a flat-scalp model, and intended only for ordinal ranking of electrode pairs in terms of distance. Nevertheless, this observation is at least generally consistent with Thatcher et al.'s (1986) results, similarly showing a quadratic relationship between coherence and distance. The present results are thus generally consistent with Thatcher et al.'s (1986) view of EEG associations, based on axonal rather than volume conduction.

Phase was found to decrease with distance in the before-blink epoch in both conditions, and to increase in the after-blink epoch in the control condition. It was expected that phase should increase with distance regardless of whether interregional correlations were founded on axonal signaling or volume conduction. An explanation that can be suggested for the observed decreases in phase with distance is that the accuracy of the phase computation was compromised because of the limited number of data points that was available within each time window. A replication of this study should attempt to increase the number of available data points. One way in which this might be done is to increase the sampling rate of the EEG signals from the present value of 128 points per second to a value of 256 points per second. This strategy should help to the extent that it provides new data points that are sufficiently independent of the existing data points.

## 11 Topographic Distribution of Coherence Differences

### 11.1 Introduction

Inspection of Figure 10.1a and 10.1b reveals differences between coherences in the picture and control conditions, but mainly for electrode pairs separated by the longer inter-electrode distances, and with a higher value of coherence in the picture than in the control conditions in the before-blink epoch, and the reverse in the after-blink epoch. A t-test was computed to test these observations. The t-tests were computed for the longest-distance 10% of the coherences in both the before and after-blink epochs. The results confirmed the observations. In the before-blink epoch, the mean coherences for picture and control conditions were 0.521 and 0.500 respectively, with a corresponding value of  $t$  of 2.95 ( $p = 0.004$ ). In the after-blink epoch, the mean coherences for picture and control conditions were 0.499 and 0.510 respectively, with a corresponding value of  $t$  of -1.72 ( $p = 0.05$ ). On the basis of the significance of these differences, topographic distributions were plotted, of the difference in coherence between picture and control conditions.

### 11.2 Method

These topographic distribution plots were constructed in exactly the same way as those for intercorrelation in Section 10. Coherence differences were displayed using a 2-dimensional projection or map of the physical electrode positions over the scalp. This map was used to display the difference in magnitude of coherence between control and picture conditions, with a separate map for each of the 4 time windows of the before-blink epoch. This coherence difference is computed for each electrode  $k$ , as

$$\delta\gamma_k = \gamma(\text{picture})_k - \gamma(\text{control})_k$$

The magnitude of this coherence difference for a pair of electrodes is coded in terms of the thickness of a line joining the two electrodes. Negative coherence differences are shown in gray, positive changes in coherence are shown in black. A positive difference indicates a higher value of coherence in the picture condition. A negative difference correspondingly indicates a higher coherence in the control condition.

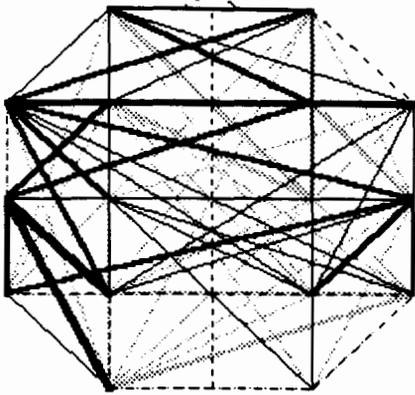
### 11.3 Results

The topographical distributions of  $\delta\gamma$ , the changes in coherence between the picture and control conditions, are shown in Figure 11.1 averaged across subjects. There are clear

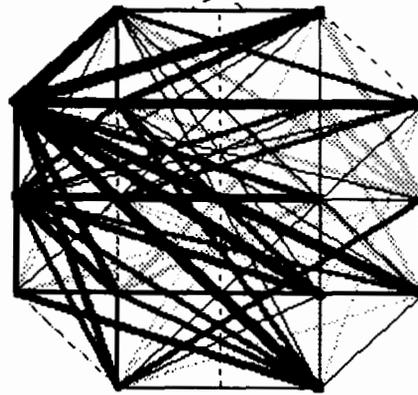
differences in the topography of the values of  $\delta\gamma$  between the 4 time windows, with unique patterns in each window. These patterns may be summarized as follows:

1. Time-window 1: Relatively moderate level associations exist between left fronto-temporal regions and a range of other areas including frontal and temporal regions bilaterally. Associations are also found between left temporal and left occipital areas.
2. Time-window 2: Widespread and strong levels of association exist between left fronto-temporal areas, and bilateral prefrontal, occipital and parietal areas.
3. Time-window 3: Moderate to strong associations occur between left frontal and anterior temporal areas, and bilateral prefrontal, frontal and occipital areas.
4. Time-window 4: Moderate level associations are relatively localized in extent, between left and right frontal and anterior temporal areas.

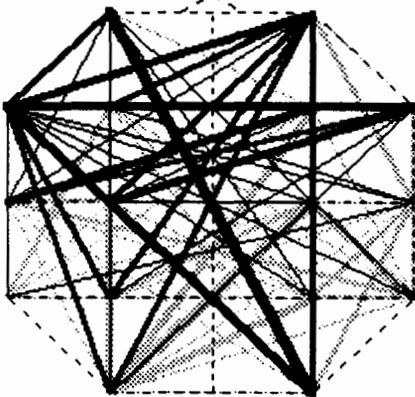
(a) time-window 1



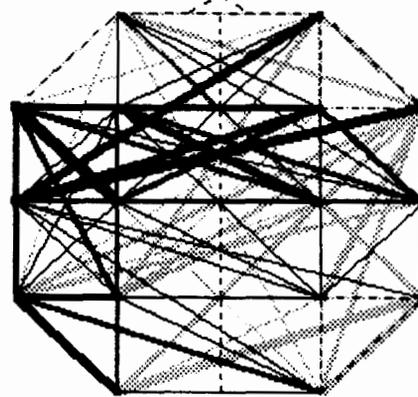
(b) time-window 2



(c) time-window 3



(d) time-window 4



**Figure 11.1 Topography of coherence differences, average across subjects**

The graphs show the magnitude of the difference in coherence, between the picture and control conditions. Coherence differences are computed as  $\gamma(\text{picture}) - \gamma(\text{control})$ . Coherences are averaged across all trials and all subjects. Line thickness is proportional to magnitude of coherence difference. The thickest lines correspond to a coherence difference of 0.15. Coherence differences less than 0.03 are not shown. Black lines indicate a positive difference, with a higher coherence in the picture than in the control conditions. Gray lines indicate a negative difference, with a higher coherence in the control condition.

## 11.4 Discussion

The pattern of coherence over the 4 time-windows of the before-blink epoch shows a greater involvement of the left hemisphere, and in particular of the left fronto-temporal region, which appears to be a focus of association: In general, other cortical areas appear to be preferentially associated with the left fronto-temporal region than with each other. The beginnings of this pattern are evident in time window 1, and become particularly evident in time window 2. In the following time windows 3 and 4 the pattern of association becomes more bilateral, with a less evident focus on the left fronto-temporal area, as connections develop

between bilateral prefrontal, frontal and occipital areas in time window 3, and between bilateral frontal and anterior temporal regions in time-window 4.

In general, the topographic distributions of coherence differences somewhat resembles those of intercorrelation differences, in showing a strong association between bilateral occipital and left temporal areas. Again, these findings are consistent with the ventral visual pathway proposed by Ungerleider and Mishkin (1982). A clear difference between the coherence distribution and the intercorrelation distribution occurs in time-window 4. In time-window 4 coherence shows a relatively moderate, relatively localized level of association, while intercorrelation showed a widespread and high level of association. Cross-correlation is proportional to the cross-product between sample voltage values, while coherence is proportional to the cross-product between sample frequency components. A high value of cross-correlation together with a low value of coherence for a pair of time-series might imply that, while multiple frequency components were present in common in both of the time-series and at the same phase angles, leading to a large value of correlation, the relative amplitudes of these components were sufficiently different between the two time series to result in a low value of coherence. This is of course only a conjecture, and other possibilities are possible, for example involving some sort of interaction with the relatively high standard error of estimate of the coherence. A replication using a higher sampling rate might provide results which could distinguish between these possibilities. A higher sampling rate will reduce the standard error of estimate of coherence, since more data points will be available for analysis within each time window.

## **12 The Discrimination Index**

### **12.1 Introduction**

This section will deal with the question of whether the information regarding the changing pattern of correlations with time in the before-blink epoch can be summarized as an index that can indicate the extent to which an object within a target object has been discriminated from its background.

Such an index, which will be referred to here as the Discrimination Index, will be constructed by making use of two sources of information about the correlations in the before-blink epoch. First, this index would need to take account of the overall value of correlation between all cortical areas. The mean correlation, computed over all electrode pairs for each time window, can be used to summarize this effect. Second, this index would need to be sensitive to the fact that, during successful discrimination, not only does the magnitude of correlations, averaged over all cortical regions, increase, but also the proportion of the cortex connected by these correlations increases. The variance in the correlations, computed over all electrode pairs, can be used to summarize this second effect. To elaborate this connection between correlation variance and proportion of cortex connected by significant levels of correlation, the present findings show that at the start of the before-blink epoch, correlations between closely-spaced regions are substantial relative to correlations between more distantly spaced areas. Thus, only a portion of the cortical regions is connected by substantial correlations. As time proceeds towards the moment of discrimination, the correlations between the closely-spaced regions remain relatively constant, while correlations between distantly spaced regions increase in magnitude, and in this way increasing the proportion of the cortex that is connected by correlations of substantial magnitude.

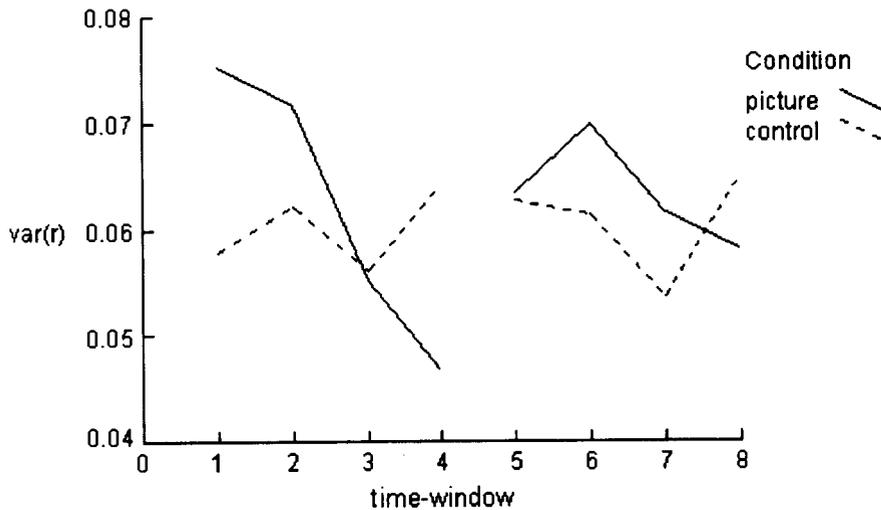
### **12.2 Method**

For both the picture and control conditions, and for each time-window, two statistics are computed. The first statistic is the mean correlation, calculated over all 120 electrode-pairs, for each time window. This value is then averaged over all trials for all subjects. The second statistic is the corresponding variance in the correlations over these 120 electrode pairs, also computed for each time window. Again, the resulting value is averaged over all trials for all subjects. Finally, for each time window and for both conditions, the ratio of these two statistics is computed. This ratio is the Discrimination Index, defined as the ratio of the mean correlation to correlation variance computed over all possible electrode pairs.

### 12.3 Results

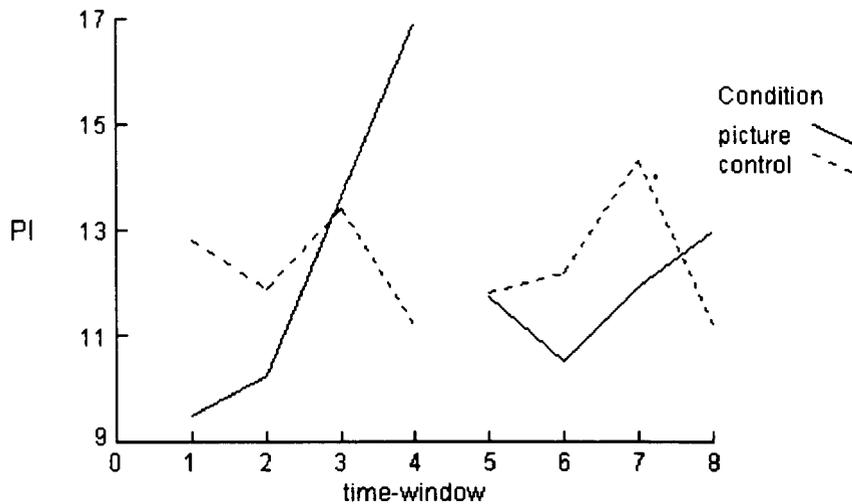
In the before-blink epoch, the Discrimination Index increased from 9.51 to 16.9 (77.7%) in the picture condition, and decreased from 12.8 to 11.2 (12.5%) in the control condition. In the after-blink epoch, the Discrimination Index increased from 9.51 to 16.9 (11.1%) in the picture condition, and decreased from 11.8 to 11.2 (5.1%) in the control condition. Correlation mean, correlation variance, and the Discrimination Index are listed for each time window in Table 12.1. Figure 12.1 shows correlation variance for each time window, averaged across subjects. Figure 12.2 shows the Discrimination Index for each time-window, averaged across subjects.

The change in both the mean and variance of the correlations appears to be an approximately quadratic change with time, with the rate of change of this index increasing with time window. As Figure 6.5 and Figure 12.1 illustrate, in the picture condition mean correlation increases while correlation variance decreases. In contrast, in the control condition mean correlations decrease while correlation variance increases. As shown in Figure 12.2, the ratio of these two quantities, the discrimination index therefore increases, approximately quadratically with time.



**Figure 12.1 Intercorrelation variance**

Intercorrelation variance is computed for each condition and time-window by calculating the variance over all 120 possible intercorrelations, and then averaging the result across all cases and subjects.



**Figure 12.2 Discrimination Index**

The Discrimination Index is computed for each condition and time-window by calculating the ratio of intercorrelation mean to intercorrelation variance.

**Table 12.1 Discrimination Index**

The table shows correlation mean and variance, computed over all 12 electrode pairs, and the Discrimination Index for each time window.

Epoch	Window	Mean		Variance		Discrimination Index	
		Control	Picture	Control	Picture	Control	Picture
BBE	1	0.740	0.715	0.578	0.0752	12.8	9.51
	2	0.739	0.737	0.622	0.0717	11.9	10.3
	3	0.752	0.752	0.562	0.0552	13.4	13.6
	4	0.719	0.789	0.642	0.0467	11.2	16.9
ABE	5	0.740	0.745	0.627	0.0635	11.8	11.7
	6	0.749	0.737	0.615	0.070	12.2	10.5
	7	0.768	0.734	0.537	0.0616	14.3	11.9
	8	0.730	0.755	0.652	0.0581	11.2	13.0

## 12.4 Discussion

The Discrimination Index increases approximately quadratically with time in the picture condition, in the 1 second interval preceding discrimination. In view of this clearly defined change, the index appears to be able to effectively summarize the information contained in the mean and variance of the intercorrelations, and therefore offers what would appear to be an effective measure of the degree to which a target has been discriminated from the background.

The Discrimination Index uses information about both the magnitude and spatial variability of intercorrelations. These two sources of information in turn access two general

features of brain function that are observed to occur during visual discrimination. In brief, correlation mean accesses information regarding the finding that intercorrelations increase over a short interval of time preceding the moment of discrimination. Correlation variance in turn accesses information regarding the finding that within this time interval intercorrelations between widely-spaced electrode pairs increases more than correlations between more closely-spaced electrode pairs. The magnitude of the correlations is suggested to be related to the degree to which oscillatory components of brain activity, averaged across the entire brain, are synchronized. This magnitude increases over the 1 second preceding discrimination, implying a corresponding increase in the degree of synchronization of the oscillatory activity, averaged across all cortical regions. The spatial variability of the correlations is a consequence of the finding that, over the before-blink epoch, intercorrelations between closely-spaced regions increase relatively little while intercorrelations between more distantly-spaced regions increase relatively more. Thus the values of distantly-separated intercorrelations approach the values of intercorrelations between closely-spaced regions, with the result that the variability in correlations computed across all distances, decreases over the before-blink epoch. An alternative way of conceptualizing the variance in correlations is as a measure of the proportion of the cortex that is engaged in significant, mutually-correlated activity. While in this situation correlation mean and variance are thus only partially independent, these two measures each contribute information that is useful in constructing the Discrimination Index.

### **III Nonlinear Analyses**

## 13 Neural Network Analysis

### 13.1 Introduction

As a matter of terminology, the phrase neural network is used in this thesis to refer to computational models of biological networks. To distinguish between these two cases, the biological networks will be referred to as neuronal networks, while the artificial networks will be referred to simply as neural networks.

Two related questions are addressed in this section. First, can the neuroelectric activity recorded in the picture and control conditions be discriminated using a neural network-based classifier? Specifically, can a neural network classifier distinguish between the signals recorded in the picture condition from those recorded in the control condition? A related question is, can an automated process be found that can reliably distinguish between these two conditions and in this way provide an objective indicator of whether an individual has been able to successfully discriminate a target from a camouflaging background?

One such indicator has already been discussed, the Discrimination Index. The Discrimination Index, the ratio of cross-correlation mean to cross-correlation variance, was shown to increase significantly over an interval spanning the 1 second prior to recognition, in the picture condition, but not in the control condition. The present section will attempt to demonstrate an alternative to the Discrimination Index. This alternative, a neural network classifier, has a number of advantages.

The first advantage is that a neural network-based classifier operates without the constraints of an a priori model of the basis upon which the classification is to be performed. The Discrimination Index on the other hand is designed to make use of information related to intercorrelations. The model implicitly adopted by the Discrimination Index is thus one which posits that differences between the picture and control conditions are to be found in the between-channel correlations. In contrast, the neural network classifier to be applied in the present section will operate without such restrictions. The classifier will make use of all possible features in the data in attempting to carry out the classification of the neuroelectric signals into one of the two categories, picture and control.

The second advantage is computational efficiency. Once the neural network has been trained by being presented with a sufficient number of exemplars of the two categories to be discriminated, a relatively small amount of computation is required to test and classify a novel exemplar. This advantage would be particularly important in possible uses of this approach in real-time applications where the network could rapidly discern that an individual has successfully discriminated a given target.

### 13.1.1 Neural Networks as Analytic Techniques

Neural Network programs were inspired in part by network theories of how storage of information and learning occur in the brain. One such theory is that of Donald Hebb (Hebb, 1949). According to Hebb's theory, learning and memory are phenomena which result from the strengthening of the synaptic connections between simultaneously active neurons: "When the axon of cell A is near enough to excite a cell B ... A's efficacy, as one of the cells firing B is increased." (Hebb, 1949). Repeated stimulation of some particular network of neurons eventually results in permanent changes in the strengths of the interconnections between the neurons within this network. The result is that a memory trace has been laid down, or relatedly, that learning has occurred.

The development of neural network models was inspired in part by such learning and memory functions of the brain. In general terms, the brain receives data from its environment. On the basis of this data, the brain is then able to induce rules pertaining to that environment, or to form internal representations of some of the features of that environment. These rules, or representations, then allow the brain to make predictions about a future state of affairs of the environment, on the basis of data about current conditions. Predictive ability in turn is related to probability of survival, conferring an evolutionary advantage on an organism which is able in this way to anticipate future conditions based on past experience. Neural networks, as computer simulations of such a rule-inducing system, similarly learn the rules, or features, embedded in examples presented to the network in training.

In a typical application, a neural network configuration may consist of three layers of nodes. An input layer containing a number of nodes equal to the number of elements in a predictor vector, provides a connection point, allowing the predictor vectors to be supplied to the network. Variable-strength couplings, the network weights, connect the input layer with the second, hidden layer. These input-to-hidden layer weights are modified over the course of the network training phase using some rule, such as the back-propagation algorithm. As a result, after training, the hidden layer nodes represent a set of features abstracted by the network from the bolus of training exemplars. A second set of weights, also modified during the training process, connects the hidden layer with the output layer. The nodes of the output layer thus each receive a unique weighted combination of the internal features stored in the hidden layer.

In the training phase, neural network is trained by presenting it with a series of exemplars. Each exemplar consists of a pair of vectors, a predictor vector, and a criterion vector representing the outcome or outcomes associated with that predictor. The predictors will, typically, be the values of a number of independent variables. The criterion, or target outcomes, associated with each such predictor will correspondingly represent the values of one or more

dependent variables. As a concrete example, predictors might represent recordings of brain electrical activity. The associated target outcomes could then be a code representing a corresponding behavioural response, or experimental condition such as a cognitive or perceptual task. Using a training rule such as back-propagation, the neural network attempts to minimize the difference between the actual outcome or output of the network, and the target outcomes coded within each of the training facts. In order to accomplish this goal, over the training session, the network develops an internal representation of the features present in the training examples. One limitation that can also be an advantage in applying neural networks to such pattern analysis tasks is that these internal feature representations do not necessarily correspond to obvious features of the data. The positive side of this behaviour is that these internal representations may, given sufficient training, come to represent features of the predictor vectors which are more efficient in performing the pattern analysis task than those features presumed to be significant on the basis of beforehand assumptions. In any event, these internal representations are distributed in the network weights, and are summarized in terms of the activation values of the hidden nodes: For each hidden node, these activation values are the weighted sums of the outputs of the previous layer, the weightings being the network weight values.

As training proceeds, the learning progress of the network can be periodically tested by presenting it with a series of test exemplars, while recording the resulting test errors. The set of test exemplars is generally created by sampling without replacement from the initial pool of training exemplars. Test error is the difference between the actual network outcomes and the target outcomes coded in the test exemplars. The size of this test error is an indication of how well the network has abstracted the significant features in the training exemplars, or in other words, how well the network has learned. When test error is seen to have reached a minimum, the network is considered to have been optimally trained. At this point, the network can be put to work, by presenting it with a set of exemplars for which there is no known outcome. The network will then generate an output for each of these unknown exemplars, on the basis of the information that the network has abstracted over the course of the training phase. These outputs are the network's predicted outcomes for each of the exemplars.

An advantage that neural networks have over other signal classification techniques is that no a priori model needs to be adopted. With Fourier analysis, for example, the a priori model is that the sought-after discriminability is present in the frequency and phase components of the data. Similarly, with classification techniques based on cross-correlation or coherence the initial assumption is made that the features upon which successful classification can be made involve correlations or coherences. With neural networks, there need not exist any preconceived notion about what aspects of the data are important for doing the classification. The data are

presented to the network, and the network takes on the problem of determining what dimensions or features in the data hold the key to discriminability. The network's internal, distributed representation of the data, coded in terms of the network weights, contains the discrimination criteria. As stated earlier, this feature of neural networks is a two-edged sword. While this internal representation of the data may be more effective than preconceived criteria in terms of performing the data analysis, these discrimination criteria that the network has developed can not be easily accessed. That is, a network user may not be able to relate the network's weights or hidden node activation values to physical features in the data.

The power of neural networks as function approximators, or as classifiers, derives from the ability of neural networks to function as nonlinear analyzers. This ability in turn is in part the result of the nonlinear transfer function generally adopted for the simulated neurons. Without a nonlinear transfer function, neural networks would be reduced to performing only as linear analyzers. We might consider the behaviour of a network with only an input and an output layer of simulated neurons or nodes, and with only linear transfer functions for each node. With such linear transfer functions the output of any node is linearly proportional to the input to the node. Such a network would only be able to present to the output nodes linear combinations of the values present at the input nodes. Next, consider added to this network a third, hidden layer containing nodes that also have linear transfer functions. The input to each hidden node is a linear combination of the inputs. The output of each hidden node is still this linear combination of inputs, albeit in general scaled by some numeric factor. Each output node in turn receives inputs from these hidden nodes, and thus receives several of the linear combinations of network inputs represented by each of the hidden nodes. Having a linear transfer characteristic, each output node linearly combines these incoming combinations. This 'linear combination of linear combinations' of the input signals is, by the definition of linearity, simply another linear combination, of the original network inputs. No matter how many layers a neural network had, if all nodes were linear, the overall network could do no more than generate outputs which were linear functions of the inputs.

When the hidden layer nodes are given nonlinear, rather than linear, transfer functions, the output of a hidden node becomes a nonlinear function of the signals received from the input nodes. The implication of this statement is that the way that a hidden node responds to any particular signal from any one input node now depends on the signals coming to that hidden node from all other input nodes. The nonlinear transfer function implies that interactions between the effects of the input signals are now possible. It is the complexity inherent in these interactions that is responsible for the rich, and sometimes chaotic behaviour that neural networks have been shown to have. In terms of their performance as classifiers, neural networks owe to this nonlinear transfer function that ability to discriminate between classes of input vectors

that are non-linearly related. Straightforwardly, a neural network can discriminate between two sets or classes of inputs that are separated by a nonlinear decision boundary in the variable space of the inputs, because the network can form a correspondingly nonlinear decision boundary, by generating at its outputs nonlinear combinations of the data presented to its inputs. Experience has shown that real-world data sets more often than not are characterized by such nonlinear relationships among the sub-groups within the data.

### **13.1.2 EEG Signal Analysis Using Neural Networks**

In this section are reviewed several of the studies that have been carried out to date that have demonstrated the effectiveness of neural network systems in the classification and categorization of EEG data. Although much of the theoretical work underlying neural networks predates the work leading to the development of other nonlinear methods such as chaos-analytic techniques, the application of neural networks to the classification of EEG signals is relatively more recent. To date, relatively few studies have been carried out in this area.

Gabor and Seyal (1992) applied a multi-layer back-propagation neural network to the problem of recognition of interictal epileptiform spike-wave patterns in the EEG, in a study using 5 epileptic subjects. EEG recordings were made from 8 pairs of channels including all scalp areas. The neural network used for the data analysis included 1 input node for each of these 8 EEG channel pairs, 8 hidden nodes and 1 output node. In order to decrease the computational load on the network, EEG data was preprocessed by calculating and using only the slopes of the spike events for each of the 8 channels. The training and testing vectors corresponded therefore to the spatial distribution of the rates of change of spike voltage. An average of 94.2% of the waves were classified correctly, with 20.9% false-positive classifications. These results, the authors suggest, provide evidence that a neural network-based pattern classifier can perform effectively in the identification of epileptiform transients in the EEG. They conclude that the network's inherent properties of being able to learn features across training examples, and of being able to generalize this learning to novel instances, were properties that allowed the network to identify waveforms which differed from the training patterns but that still maintained the spatio-temporal characteristics of epileptiform waveforms.

Jando, Siegel, Horvath and Buzaki (1993) similarly used a multi-layer neural network to classify epileptiform EEG activity. In their study, spike-wave activity was recorded over 12 hours from the neocortex of rats that had been bred to exhibit epileptic symptoms. Neural network configuration was optimized by conducting a parametric study of numbers of input and hidden neurons. It was found that a network of 16 input and 19 hidden neurons was most efficient in terms of classification error rate achieved after a fixed number of iterations. One output neuron

was used, corresponding to a discrimination between two conditions, epileptiform versus non-epileptiform input. The authors analyzed both the raw time-series, as well as Fourier transformations of the raw data including both amplitude and phase. Each time series consisted of 12 seconds of recorded EEG data, digitized at 100 samples per second, and selected visually to represent one of the two data conditions. The training set consisted of 469 time-series corresponding to spike-wave activity, and 1,133 time-series of non-spike-wave activity. The authors chose to analyze each time series using a time-window containing 16 data points, and sliding across the time-series in steps of one data point. At each step, the 16 data points were presented to the neural network input nodes.

The network correctly classified 96% of epileptiform events, and misclassified 30% of non-epileptiform events. The authors suggest that this performance demonstrates the power of a non-linear analytic technique such as a neural network to find correspondingly non-linear relationships between dependent and independent variables. Techniques limited to utilizing only linear relationships are prone, they suggest, to committing false-positive misclassification. Training speed was found to be higher when the network was supplied with Fourier transformed data than when it was supplied with raw data. This finding suggests that the distinguishing features between the epileptiform and control conditions consisted of amplitude and phase differences of periodic components within the data. As Smith (1993) has pointed out, doing some of the work that the network would otherwise have to do by preprocessing the data usually results in improved training.

EEG waveforms were classified according to sleep stage using a multi-layer network (Grozingler, Kloppel and Roschke; 1993). The goal in this study was to train the network to classify samples of EEG recorded during sleep as corresponding to REM or NREM periods. EEG recordings at electrode Cz were made from subjects during sleep. The EEG records were digitized at 100 samples per second, and separated into 6 frequency ranges, 0.5 to 3.5 Hz, 3.5 to 7.5 Hz, 7.5 to 15 Hz, 15 to 25 Hz, 25 to 45 Hz and 0.5 to 45 Hz. Power within each band was computed and used as one component of the network input vector, forming training and testing vectors of 6 components. Correspondingly the network was composed of 6 input nodes, one for each component, 4 hidden nodes and 1 output node. Raw data for each exemplar consisted of 2048 data points, with data from one night used for the training exemplars and data for a subsequent night used for the testing exemplars. A total of 1300 training and testing exemplars were used. Following a training phase, testing results were an average of 89% of exemplars correctly classified as either REM or NREM. The authors point out that conventional EEG analysis to determine sleep stage requires additional recordings of electrooculographic and electromyographic potentials. The results using the neural network demonstrate that such classification can be accomplished using only scalp EEG recordings.

The event related P300 response in patients with multiple sclerosis (MS) was distinguished from the P300 in a normal control group using a multi-layer neural network (Slater, Wu, Honig, Ramsay and Morgan, 1994). Characteristically, amplitude of the P300 response has been shown to be sensitive to stimulus probability, increasing for rare stimuli, while latency has been found to be linked to task difficulty, increasing for example with increasing difficulty of stimulus discrimination. While P300 characteristics have been found to be altered in patients with MS, with, in particular, an increase in latency, such alterations are too subtle, the authors point out, to allow them to be used in clinical diagnosis. An oddball paradigm was used to elicit the P300, with infrequent target audio tones interspersed with frequent standard tones of a different pitch. Recordings of P300 components were made to the target tones at electrode sites Fz, Cz and Pz. Averages were formed over 100 recordings each containing 256 data points. Training and testing exemplars were then formed by uniformly sampling 25 points from each of these averaged recordings. A set of 101 training exemplars was formed in this way for each of the three electrodes, 51 from the MS group and 50 from the control group. Testing data consisted of 10 MS and 10 control exemplars for each electrode. Three identical neural networks were used in the analysis, one for each electrode. Each network consisted of 25 input nodes, corresponding to the 25 exemplar data points, 8 hidden nodes and 2 output nodes. Final scoring was done on the basis of a 2 out of 3 majority rule. An exemplar was categorized according to whatever classification was assigned to that exemplar by at least 2 of the 3 networks. Classification accuracy when the performance of each network was considered separately was found to be 85% at Cz, and 80% at both Fz and Pz. Using the majority rule, classification accuracy was 90%. The authors point out the difficulty with using neural networks that the basis of the classifications is generally not easily available. On the other hand, they suggest, if network performance indicates an effective ability to distinguish between disease and control conditions, then the network is nevertheless useful as a tool for clinical diagnosis.

The diagnostic capability of neural networks has also been used to classify subjects as depressive, psychotic or normal (Kloppel, 1994b). EEG recordings were made from 18 subjects, 6 depressives, 6 psychotics and 6 normal controls. Recordings were made from 16 scalp electrodes, over an interval of 30 minutes. Preprocessing consisted of reducing 4 second segments of the EEG record to 6 values, representing spectral power levels in the delta, theta, alpha 1, alpha 2, beta 1 and beta 2 frequency bands. Artifacts were eliminated by setting to zero those data points identified as artifactual, and linear interpolation was then used to bridge the surrounding data points. Neural network was examined after training on data from varying numbers of subjects. After training on data from only one subject, the network was able to classify unlabelled data segments from that same subject with an accuracy of 80%. After training on two or more subjects, classifying data segments belonging to any one subject,

dropped to 66%. The network was thus able to recognize data belonging to the subject on which the network had been trained. However, the network was only marginally well able to generalize this knowledge to the classification of data from other subjects.

A different type of network, a Learning Vector Quantizer (LVQ), has also been used in a number of studies involving EEG classification. The LVQ is a self-organizing network, comprised of mutually interconnected nodes, in comparison with the multi-layer back-propagation network which is organized in hierarchically connected layers. The LVQ does not require feedback in order to learn. Such feedback is required by networks using the back-propagation training algorithm, in the form of the target criterion values that form part of each of the training exemplars. The LVQ is thus in the class of unsupervised networks, in comparison with multi-layer networks which are classed as supervised networks. Vectors representing sets of data points are presented to the LVQ. After a number of such presentations the LVQ is able to place a vector into one of several categories. This function is similar to that performed by traditional statistical cluster-analysis techniques.

An LVQ applied to the classification of EEG waveforms was used to predict laterality of hand movement (Pfurtscheller, Flotzinger, Mohl and Peltoranta, 1992). A total of 30 channels of EEG data were recorded from 3 subjects prior to voluntary right or left hand movements. Each channel was referenced to the weighted average of 5 surrounding channels. Subjects were asked to press a microswitch with either a left or right finger and for a specified duration. Direction and duration were indicated by cues presented in succession on a computer monitor. Subjects were asked to initiate the movement following the duration cue, and EEG recordings were made during the interval between the direction and duration cues. Recording terminated approximately 0.5 seconds before the start of movement. Training data consisted of single trial records of event-related desynchronization, in the 8 to 10 Hz and 10 to 12 Hz frequency bands. Trials were selected manually to use only those which showed clear and artifact-free alpha activity. In a training phase, the LVQ was allowed to self-organize, a process analogous to the pattern-formation which occurs in multi-layer neural network weights. After training, the LVQ could significantly well predict side of hand movement, with an accuracy of 85%, 74% and 64% respectively for the three subjects. The authors claim that this finding is the first demonstration that EEG signals can be classified without the use of averaging. They point out that by not using averaging, the problem of dealing with a statistically nonstationary signal is avoided. The results do however confirm the findings of averaged potential studies that signals related to the preparation for finger movement are available for several seconds prior to the start of movement.

An LVQ network was also used to perform on-line EEG classifications (Flotzinger, Kalcher and Pfurtscheller, 1993). Subjects were instructed to press a microswitch with either left

or right index fingers as indicated by a direction cue presented visually on a computer monitor, and upon an initiate-movement cue presented 1 second later. Training vectors were formed from EEG recordings made during this 1 second interval between the presentation of the direction cue, and prior to actual movement, from channel pairs C3-Cz and C4-Cz. The LVQ was trained to predict laterality of movement prior to the actual movement itself. The goal of the study was to have the network predict laterality without any movement actually occurring. The network was first trained, and then used to predict side of movement. The network's prediction was displayed on the computer monitor as feedback to the subject. In the later stages of the prediction phase the microswitch was removed so that network's predictions were made solely on the basis of pre-movement EEG's. After training, the LVQ was found to be able to correctly predict side of movement between 59% and 86% of the time, both with the microswitch present, and without the microswitch. Such on-line prediction is made possible, the authors suggest, by the operating speed of the LVQ, which they suggest is higher than that of a multi-layer neural network.

### **13.1.3 Summary and Discussion**

Neural networks, both of the multi-layer and LVQ types, have been demonstrated to be effective as pattern classifiers. Using multi-layer networks, this classification ability has been applied to the problem of recognizing epileptic spike-waves in humans (Gabor and Seyal, 1992) and in rats (Jando, Siegel, Horvath and Buzaki, 1993), classifying stages of sleep (Grozinger et al, 1993), recognizing the effects of multiple sclerosis on event-related potentials (Slater et al., 1994), and distinguishing normal, depressive and psychotic subjects (Kloppel, 1994b). LVQ networks have been used to classify movement-related potentials to predict the laterality of finger movement (Pfurtscheller et al, 1992; Flotzinger et al, 1993).

One reason that neural networks are effective in dealing with EEG data may be that neural network classifiers are examples of nonlinear techniques. In particular, it has been suggested that the success of neural networks as categorizers is probably due in part to the ability of neural networks to function as nonlinear discriminant analyzers. Webb and Lowe (1990) report on theoretical results involving layered nonlinear feed-forward adaptive networks that demonstrate why such networks are effective at performing classification tasks. The authors show that this discriminatory ability is a result of the first half of the network, from input nodes to hidden nodes, performing a nonlinear transformation of the input data into a feature space, defined by the hidden units, in which the discrimination should be easier. The second half of the network, from the hidden to the output nodes, then executes a linear transformation aimed at minimizing the mean-square error to a set of given output patterns. In short, neural networks are

capable of performing nonlinear discriminant analysis. The brain, as a distributed, nonlinear dynamical system, is probably not effectively describable or analyzable using purely linear methods. Reasonably, the analysis of such a nonlinear system should require the application of correspondingly nonlinear methods, examples of which have been discussed in this section. A comprehensive overview of the application of neural networks to the analysis of EEG data can be found in Klöppel (1994a).

## **13.2 The Hypothesis**

It is suggested that the exemplars representing the control and picture conditions of the present study will be distinguishable to varying degrees, depending on the time-window. In particular, and on the basis of the results of the cross-correlation analysis, it is suggested that there should be an increasingly effective discrimination over the 4 time-windows of the before-blink epoch. Furthermore, and again on the basis of the cross-correlation analysis results, it is proposed that the data from the 4 time-windows of the after-blink epoch will be relatively less discriminable.

The ability of the probabilistic network to discriminate between picture and control condition using data from the time-windows of the before-blink epoch will depend, it is suggested, on features of the data that are related to the increasing level of organization of the inter-cortical signaling over the course of this epoch.

## **13.3 Method**

The approach that will be used in this analysis is neural network classification of the data within a moving time window, swept across two 1 second intervals of the recorded EEG. The first interval is the 1 second epoch preceding the eye-blink by which subjects signal the target discrimination event. The second interval is the 1 second epoch which begins after the cessation of artifacts associated with the blink, 1.5 seconds after the start of the eye-blink artifact. EEG recordings made during these epochs in both the picture and control conditions will be used to construct network exemplars. The network will then be asked to learn to distinguish between exemplars corresponding to these two conditions.

### **13.3.1 The Neural Network**

The general type of neural network used in this study is a form of the probabilistic network, developed by Specht (1990), the generalized regression neural network (GRNN) (Specht, 1991; Wasserman, 1993). In general terms, the GRNN, like the back-propagation neural network, is able to approximate any functional relationship between input and output. The

following description will be based on the network being used as a classifier; that is, to learn to place test exemplars into one of 2 or more categories.

Structurally, the GRNN resembles the back-propagation neural network. The GRNN has a number of inputs equal to the number of predictor values in the training or testing exemplars. The input nodes of the GRNN, like those of a back-propagation network, are merely connection points to which the elements of the test exemplars are applied, one at a time. The GRNN has a number of hidden units equal to the number of training exemplars. There is one hidden unit for each training exemplar. Unlike the back-propagation network then, the GRNN does not require an estimate of the number of hidden units to be made before training can begin. Finally, the GRNN has a number of outputs equal to, if the GRNN is used as a classifier, the number of categories being discriminated. More generally, the number of outputs of the GRNN is equal to the number of criterion variables being predicted.

The GRNN however differs functionally from the back-propagation neural network. First, there is no counterpart to the iterated back-propagation network training phase. Instead, the entire training matrix is installed in the GRNN, as the weights between the input and hidden layers. In more detail, the weights between the input nodes and each hidden node represent a single training exemplar. Thus, the weights between the input layer and hidden node 1 are the components of the predictor part of training exemplar 1. Recall that each exemplar consists of two parts. The first part consists of the predictor values representing the values of the variables being used to predict some outcome, while the second part consists of the criterion values representing the values of the variables being predicted. The equivalent of training with the GRNN thus takes no more time than is required to load the training file into working memory. This scheme is in direct contrast with back-propagation networks which must iteratively apply a heuristic, such as the method of steepest descent, to adjust the values of the input node to hidden node weights. The testing phase of the GRNN also differs significantly from that of the back-propagation network. In order to describe the GRNN testing phase it is useful to first state what the outputs of the GRNN represent. With the GRNN used as a classifier, the outputs of the GRNN are the probabilities that the test exemplars belong to the categories being discriminated. The GRNN implements a procedure for estimating the probability of a test exemplar vector given a set of training exemplars, based on the principle of Bayesian classification. The GRNN will in fact approach an optimum Bayesian classifier given a large enough number of training exemplars (Wasserman, 1993).

The algorithm used for GRNN testing may be described as follows. In the training phase the entire set of training exemplars is loaded into the network, with the components of each of the exemplars becoming the weights between the input nodes and one hidden node corresponding to that exemplar. The testing phase begins with a testing exemplar being applied

to the input nodes. Each hidden node will thus receive the product, and more precisely the vector dot-product, of the testing exemplar and the training exemplar corresponding to that hidden node. This vector dot-product is a direct measure of the collinearity, or in general terms the similarity, between the test vector and a training vector. Other similarity measures can also be used, such as the sum of squares of the difference between the components of the test and training vectors. This latter measure of collinearity is used in the present analysis.

Each hidden node then performs a non-linear transformation on this dot-product. While in the back-propagation network the transformation generally involves the sigmoidal function, in the case of the GRNN the corresponding transformation involves the exponential function. The meaning of this transformed dot-product is that it represents the probability of obtaining the particular testing exemplar, given a probability density function with a mean equal to the mean of the training exemplar, and standard deviation defined by a parameter referred to as smoothing (generally, smoothing is the only parameter than needs to be selected when using the GRNN). Straightforwardly then, the GRNN computes at each hidden node the probability of the current test exemplar, given the existence of the training exemplar corresponding to that hidden node. The more similar the testing and training exemplar are, or in alternative terms the more nearly collinear they are, the greater will be the resulting probability of that testing exemplar occurring, given the training exemplar.

These individual probabilities next need to be combined in order to generate the desired output of the GRNN. This output is the probability of the current test exemplar given all of the training exemplars. This combining is performed in the hidden to output connections of the GRNN. The transformed output of each hidden node is connected to each output node. As in the back-propagation network, these connections between the hidden and output nodes contain weights. However, and again in contrast with the back-propagation network, these weights in the GRNN are not trained, but rather are assigned values. These values are dummy codes representing the category of each of the hidden nodes. Recall that each hidden node represents one training exemplar, and that this exemplar belongs to one of the categories being discriminated. The dummy codes between a hidden node and all the output nodes are 1 for the output node which represents the same category as the training node, and 0 for all other output nodes. As an example, if there are two categories, A and B, being discriminated, the GRNN will have 2 output nodes, node A and node B. Let us assume that hidden node 1, representing training exemplar 1, belongs to category A. The weight between hidden node 1 and output node A will be 1, and the weight between hidden node 1 and output node B will be 0. The effect of this coding is to connect only hidden and output nodes of the same category, with the result that an output node of a particular category will receive inputs only from hidden nodes of the same category. That output node then simply sums these individual inputs. While each of these

inputs from the hidden nodes represents the probability of the current test exemplar given a particular training exemplar, this sum at an output node represents the probability of the current testing exemplar given all of the training exemplars in one category. Finally, in order to generate an output which represents the actual probability, the value at each output node is normalized by dividing by the sum of all hidden node outputs.

Thus, for this 2 category example, the value generated by the network at output node 1 is the probability that the currently-applied test exemplar belongs to category A. The value at output 2 is the probability that the testing exemplar belongs to category B.

This technique of combining the probability density functions of individual exemplars of a category to approximate the probability density function of the category is due to Parzen (1962). Parzen showed that with a sufficient number of exemplars of a class, the result will approach the true probability density function of the category.

An advantage that the GRNN has over the neural network and the genetic network is the single pass nature of the algorithm. Training and testing can typically be several orders of magnitude faster for the GRNN than for the neural or genetic networks. A potential limitation is that, since all training examples are stored in working memory, the size of the training data set is limited by the amount of available memory.

## Algorithm

The following algorithm describes the testing phase of the GRNN.

### *Symbols:*

- $x_i$       the  $j$ -th test exemplar vector
- $u_j$       the  $i$ -th training exemplar vector
- $h_j$       probability of test exemplar  $x_i$  given the probability of a training exemplar,  $u_j$ .
- $\sigma$       smoothing parameter; defines the standard deviation of the PDF
- $c_k$       output corresponding to category  $k$

For each test exemplar  $x_i$

1 For each training exemplar  $u_j$

1.1 Estimate the probability of  $x_i$  given the probability of  $u_j$ :

$$h_j = \exp[-(x_i - u_j)^T(x_i - u_j) / 2\sigma^2]$$

2 Compute the sum over all probabilities:

$$\text{sum}(1) = \sum h_j$$

3 For each output (category)  $c_k$

3.1 Compute  $\text{sum}(2)$ , the sum over  $h_j$  for all category  $k$  training exemplars:

$$\text{sum}(2) = \sum_{j=k} h_j$$

3.2 Compute the probability of  $x_i$  by dividing  $\text{sum}(2)$  by  $\text{sum}(1)$ :

$$c_k = \left[ \sum_{j=k} h_j \right] / \left[ \sum h_j \right]$$

The value of  $c_k$  now represents an estimate of the probability of test exemplar  $x_i$  given all training exemplars from category  $k$ .

The implementation of the generalized regression neural network function and all supporting operations utilized in the present analysis were performed using the data analysis program Simulnet™ version 2.3.

### 13.3.2 Training and Testing Exemplars

Exemplar vectors used to train and test the network all have the same format. These vectors  $x_i$  consist of two parts, a predictor portion comprised of  $n$  components  $p_j$  and a criterion portion comprised of  $m$  components  $c_k$

$$x_i = \{p_1, \dots, p_n; c_1, \dots, c_m\}$$

In the present application, there is 1 criterion component, a dummy code denoting the category membership of the exemplar. The predictor component of each exemplar is created as follows. From the data for each subject and for each trial, two sets of data are extracted. The first set is a matrix containing the 16 channels of 128 data points each, corresponding to the 1 second interval preceding the blink. The second set is a matrix containing the 16 channels of 128 data points each corresponding to the 1 second interval following the blink. Each of these matrices is then converted into a vector, by concatenating all 16 channels, placing them end-to-end. These vectors thus consist of  $128 \times 16$  or 2048 components. The first 128 components thus correspond to channel 1 (Fp1), and the last 128 components correspond to channel 16 (O2). One such vector is created from each trial for each subject. These vectors form the predictor sections  $\{p_1, \dots, p_n\}$  of each of the exemplars. The criterion dummy code that is added as a final element to each exemplar labels the exemplar as corresponding to the picture or to the control

conditions. There were 54 control condition exemplars and 42 picture condition exemplars, for a total of 96 exemplars.

### **13.3.3 Jackknifed Classification**

In order to make optimal use of the available number of exemplars, a jackknifing classification procedure was chosen. Each one of the 96 exemplars was in turn removed from the total set of exemplars. The remaining 95 exemplars were used to train the neural network, while the single withheld exemplar was then applied to the network for classification. The resulting score assigned to that exemplar by the network, the probability of obtaining that exemplar given the existence of the other 95 exemplars, was recorded. This procedure was repeated for all 96 exemplars. The result was a set of 96 scores, one for each exemplar, denoting the probability that the exemplar belonged to one of the two categories. Next, the effectiveness of the network in carrying out this classification was computed. A t-test of significance was computed on the two groups of scores, the group of scores for control condition exemplars and the group of scores for picture condition exemplars. The resulting numerical value of  $t$  is equal to the difference in the means of the two groups divided by the pooled standard deviation computed over both groups. The value of  $t$  is therefore an index of the relative difference between the two groups as estimated by the network.

## **13.4 Results**

For each of the 8 time-windows, and for each of the 96 exemplars that were tested, (representing the 96 trials, 54 from the picture condition and 42 from the control condition), a score was generated by the network. Each of these scores represents the probability, as computed by the network, of that exemplar occurring, given the probability density functions for each of the two categories (control and picture), constructed using the other 95 exemplars. The network scores for each of the 8 time-windows are shown in Table 13.1.

The results of the t-tests for the difference in network scores for the control and picture conditions are shown in Table 13.2 for the before-blink epoch, and in Table 13.3 for the after-blink epoch. For the before-blink epoch, the values of  $t$  increase monotonically from time-windows 1 to 4. A significant value of  $t$  in time window 4 ( $t = -2.828$ ,  $p = 0.003$ ) indicated that the network was able to significantly well classify exemplars as belonging to either the control or picture conditions on the basis of the EEG recorded during the final 0.25 seconds preceding the blink. For the after blink epoch, non-significant values of  $t$  occurred in time windows 5, 7 and 8, indicating that the network was not able to classify exemplars into the two categories on the basis of EEG recorded during these windows. A significant value of  $t$  did however occur in time-

window 6 ( $t = -1.745$ ,  $p = 0.042$ ), indicating significantly effective classification of exemplars on the basis of EEG from this window. These results are graphed in Figure 13.1.

**Table 13.1 Generalized Regression Neural Network Scores**

Each score represents the probability, computed by the network, of the occurrence of the corresponding exemplar, given the probabilities of the other exemplars, that is, given the probability density functions for the control and for the picture conditions, computed using the other 95 exemplars. The Group label is coded as 0 for the control condition, 1 for the picture condition.

Trial	Group	1	2	3	4	5	6	7	8
1	0	.0419	.588	.883	.36	.122	.0136	.237	.0963
2	0	.388	.842	.0279	.255	.515	.0843	.388	.407
3	0	.0741	.0479	.0307	.529	.165	.658	.354	.184
4	0	.392	.124	.234	.0781	.278	.885	.167	.0305
5	0	.431	.212	.141	.247	.261	.377	.266	.052
6	0	.985	.0076	.851	.254	.238	.394	.825	.544
7	0	.0058	.3	.33	.0845	.0886	.875	.378	.0465
8	0	.331	.882	.386	.542	.531	.539	.19	.314
9	0	.024	.0007	.0445	.853	.608	.0384	.731	.0884
10	0	.0339	.386	.389	.582	.228	.234	.929	.149
11	0	.303	.0879	.301	.233	.116	.516	.646	.82
12	0	.151	.979	.886	.452	.377	.02	.326	.0814
13	0	.531	.124	.182	.157	.122	.996	.489	.0823
14	0	.403	.516	.964	.14	.65	.42	.198	.162
15	0	.013	.556	.575	.0683	.138	.0126	.373	.189
16	0	.15	.957	.205	.967	.394	.0652	.0917	.671
17	0	.0015	.777	.056	.128	.0964	.0682	.375	.866
18	0	.0007	.447	.175	.751	.535	.998	.922	.619
19	0	.22	.0772	.901	.0403	.683	.42	.35	.429
20	0	.813	.0129	.0058	.396	.0139	.519	.277	.372
21	0	.405	.163	.119	.338	.0536	.0624	.0473	.857
22	0	.114	.04	.87	.504	.0925	.0347	.13	.232
23	0	.345	.185	.158	.446	.266	.351	.78	.302
24	0	.975	.657	.341	.161	.105	.104	.126	.2
25	0	.0844	.549	.0282	.172	1	.95	.236	.26
26	0	.0080	.991	.26	.151	.178	.0033	.212	.492
27	0	.0736	.0594	.0783	.286	.566	.341	.399	.226
28	0	.339	.0147	.152	.834	.62	.975	.544	.745
29	0	.175	.0112	.0244	.142	.022	.0447	.202	.233
30	0	.248	.0258	.401	.801	.261	.0663	.0416	.202
31	0	.0411	.0532	.639	.08	.526	.696	.645	.117
32	0	.0007	.197	.37	.143	.0611	.401	.0773	.913
33	0	.0761	.11	.391	.0543	.539	.0712	.0982	.236
34	0	.103	.132	.0229	.637	.204	.832	.254	.277
35	0	.786	.927	.556	.325	.392	.8	.511	.0847
36	0	.161	.0056	.0424	.262	.16	.0775	.299	.19
37	0	.547	.0917	.559	.352	.222	.423	.885	.225
38	0	.066	.108	.961	.836	.842	.265	.0606	.962
39	0	.0285	.0215	.876	.263	.51	.263	.174	.204
40	0	.0079	.0777	.647	.251	.413	.119	.0651	.156
41	0	.059	.304	.15	.0361	.189	.876	.0866	.264
42	0	.132	.281	.306	.585	.406	.61	.286	.065
43	0	.0562	.163	.131	.211	.0118	.866	.586	.272
44	0	.0755	.131	.325	.609	.245	.0988	.0726	.58
45	0	.0039	.0064	.558	.176	.228	.194	.0193	.643
46	0	.0010	.0943	.0202	.293	.498	.0107	.304	.428
47	0	.0037	.019	.144	.127	.297	.0035	.0857	.154

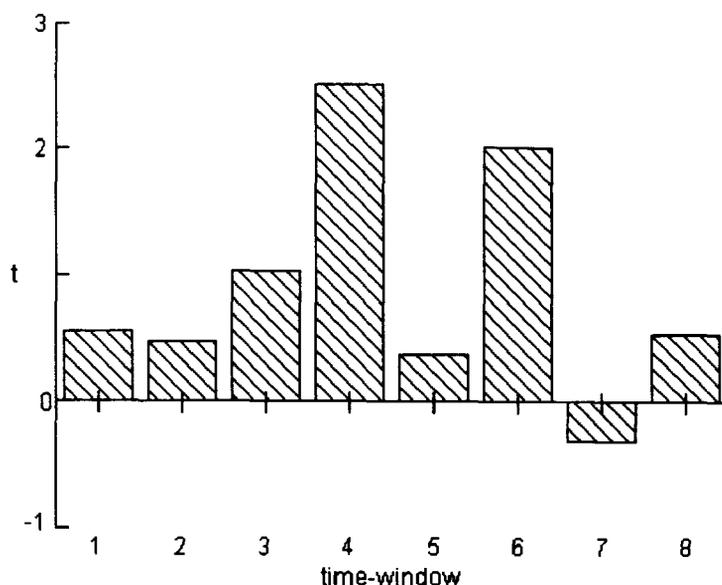
**Table 13.1 Generalized Regression Neural Network Scores**  
(continued)

Trial	Group	1	2	3	4	5	6	7	8
48	0	.0006	.0321	.0221	.139	.748	.0205	.51	.0414
49	0	.46	.22	.0947	.0708	.0747	.143	.596	.404
50	0	.0191	.268	.124	.163	.167	.341	.236	.202
51	0	.099	.0036	.0316	.377	.313	.181	.314	.0873
52	0	.0038	.0048	.0192	.12	.259	.0783	.572	.12
53	0	.242	.296	.0385	.437	.234	.0017	.132	.03
54	0	.0063	.0832	.0727	.0561	.58	.266	.0024	.615
55	1	.0364	.811	.663	.118	.988	.279	.97	.252
56	1	.0069	.0681	.161	.285	.723	.279	.121	.0119
57	1	.0061	.0114	.833	.126	.0333	.897	.279	.626
58	1	.0046	.098	.553	.726	.131	1	.435	.373
59	1	.173	.864	.355	.594	.501	.806	.7	.176
60	1	.282	.0041	.143	.0807	.751	.33	.0113	.792
61	1	.0163	.614	.689	.506	.994	.231	.924	.285
62	1	.134	.172	.247	.115	1	.96	.207	.18
63	1	.613	.011	.0441	.866	.361	.81	.496	.0191
64	1	.549	.244	.0602	.867	1	.977	.581	.0351
65	1	.144	.0172	.0977	.171	.453	.661	.109	.377
66	1	.0852	.0122	.983	.183	.219	.0675	.0845	.452
67	1	.392	.0245	.073	.929	.536	.991	.0539	.321
68	1	.0957	.813	.631	.961	.994	.648	.501	.285
69	1	.0007	.244	.0169	.588	.0307	.022	.309	.0615
70	1	.0244	.52	.0066	.901	.0891	.924	.0479	.72
71	1	.136	.789	.411	.744	.216	.0752	.345	.442
72	1	.0022	.269	.635	.57	.0331	.751	.0338	.853
73	1	.987	.845	.519	.335	.244	.958	.456	.0284
74	1	.0208	.089	.384	.503	.249	.652	.394	.0362
75	1	.267	.0535	.286	.97	.429	1	.737	.993
76	1	.0476	.0612	.953	.397	.169	.0555	.282	.25
77	1	.344	.0069	.12	.367	.13	.221	.925	.962
78	1	.0264	.0145	.0615	.13	.215	.152	.13	.0889
79	1	.528	.241	.751	.82	.199	.0044	.115	.643
80	1	.253	.998	.493	.0107	.286	.934	.014	.64
81	1	.0125	.875	.0233	.171	.178	.243	.0219	.18
82	1	.0313	.025	.516	.58	.427	.792	.269	.16
83	1	.696	.0284	.0080	.991	.315	.0871	.62	.371
84	1	.933	.24	.434	.0938	.0152	.439	.489	.394
85	1	.915	.784	.973	.312	.796	.0013	.142	.0463
86	1	.517	.0008	.199	.0978	.082	.873	.341	.662
87	1	.0082	.915	.182	.538	.0189	.0363	.323	.237
88	1	.0815	.499	.318	.738	.215	.75	.271	.788
89	1	.358	.0284	.506	.349	.326	.491	.0324	.283
90	1	.0018	.0054	.48	.226	.0423	.996	.424	.6
91	1	.537	.047	.667	.579	.117	.108	.446	.116
92	1	.098	.485	.372	.594	.12	.926	.356	.156
93	1	.268	.0026	.36	.422	.362	.0539	.172	.0836
94	1	.0002	.018	.0009	.02	.173	.011	.101	.27
95	1	.0001	.0517	.0113	.487	.0769	.0296	.0141	.0912
96	1	.208	.489	.778	.459	.183	.176	.0917	.293

**Table 13.2 T-test Results for Network Scores**

The values of *t* are computed for the two groups of scores, corresponding to the control and picture conditions, for time-windows 1 to 4 of the before-blink epoch, and 5 to 8 of the after-blink epoch. The control group contains 54 scores, and the picture condition contains 40 scores, for a *df* of 94. In the before-blink epoch, *t* values increase from time-windows 1 to 4. A significant value of *t* in time-window 4 indicates that the network was able to significantly well distinguish between the exemplars of the picture and control conditions, on the basis of EEG recorded during the final 0.25 seconds preceding the blink. In the after-blink epoch, values are not significant for windows 5, 7 and 8, indicating that the network could not find differences between the picture and control conditions in these time-windows. Only window 6 shows a significant value of *t*, indicating that in time-window 6 the network was able to significantly well classify exemplars as to category.

Window	Student's <i>t</i>	<i>p</i> (1-tailed)	Mn (control)	Var (control)	Mn (picture)	Var (picture)
1	.548	.293	.204	.063	.234	.079
2	.474	.318	.264	.091	.295	.115
3	1.04	.151	.317	.091	.381	.088
4	2.517	.0067	.325	.06	.465	.089
5	.371	.356	.323	.053	.343	.096
6	2.021	.023	.346	.107	.493	.147
7	-.317	.376	.335	.063	.318	.07
8	.538	.296	.319	.066	.348	.077



**Figure 13.1 Generalized Regression Neural Network classification results**

The figure shows the values of  $t$  resulting from  $t$ -tests conducted on the GRNN scores for the two sets of exemplars, representing the picture and control groups. Separate  $t$ -tests were conducted for scores for each of the time windows of the before-blink epoch (windows 1 to 4) and the after-blink epoch (windows 5 to 8). The highly significant value of  $t$  in window 4 indicates that the GRNN was able to accurately classify exemplars representing the picture and control groups into their respective categories. In turn, this result indicates that data in time-window 4 provided a clear basis upon which the two groups could be distinguished.

### 13.5 Discussion

Over the time-windows of the before-blink epoch, classification accuracy of the network increased almost monotonically, reaching a highly significant level in time window 4, and a significant level in window 6. These results indicate that the GRNN was able to accurately classify exemplars representing the picture and control groups into their respective categories using data from time-windows 4 and 6. In turn, these results indicates that data in time-window 4, and to a lesser extent time-window 6, provided an effective basis upon which the picture and control groups could be distinguished.

These results, within the before-blink epoch, imply that data from the two conditions contain features that are increasingly dissimilar as the moment of visual discrimination, signaled by the blink, approaches. Applying the results of the previous correlation analyses, it is suggested that these distinguishing features are related to the increasing level of organization of the signals over the time-windows of the before-blink epoch.

These neural network classification results are thus consistent with the results of the cross-correlation, coherence and mutual information analyses, all of which indicated that interregional synchronization increased during the before-blink epoch. In the after-blink epoch, only time-window 6 showed a significant value of  $t$ , which however was smaller than the value of  $t$  found for time-window 4. The network was thus able to discriminate between picture and control conditions on the basis of the EEG recorded during time-window 6. This discrimination was however less effective than that which the network was able to accomplish using time window 4. A possible explanation for the finding of a significant classification in time window 6 is that during this interval subjects were re-focusing their attention the target images in the picture conditions, an event that would therefore be associated with an EEG that was different from that associated with the corresponding time window in the control condition. In the control condition there was no such image on which subjects could re-focus. Thus, during time-window 6 the re-focusing on the target image would have created a short-lived state of increased cortical organization.

As an extension of the present work, it is proposed that an experiment could be designed in which subjects would not blink to indicate discrimination. Instead, subjects would be asked after every trial whether they had been able to recognize the camouflaged target object. Trials would then be separated into two groups, those in which a discrimination event occurred, and those in which discrimination did not take place. A shorter presentation interval for each image would be used. Based on the results of the present study, it is expected that presentation interval could be shortened to about 2 seconds. As in the present study, individual images would be presented more than once. Those images that are not discriminated on a first presentation might be decoded on a subsequent presentation. Network exemplars would be constructed from EEG recordings made during each of these 2 second presentation intervals. Again, as in the present study, the GRNN would be asked to learn to discriminate between exemplars representing intervals during which discrimination occurred from those representing intervals for which discrimination did not occur. If this proposed experiment shows that the network can discriminate between these two cases, a further experiment would be carried out. This further experiment would be similar in intent to the first proposed extension, with the difference that the neural network classification would be performed on-line. That is, the network would be asked to indicate the presence of a recognition event with a short time delay after the event actually occurs. It is estimated, based on the experience of the present study, that a generalized regression network running on a 486DX or better processor should be capable of classifying one or two seconds worth of data within a few milliseconds. This experiment would also be used to determine the minimum number of data points that are required in a real-time situation, for the network to accurately detect the target discrimination event. The present results suggest that the

major features differentiating the picture and control conditions are found within the last 0.25 seconds prior to the eye-blink by which subjects signaled discrimination. If such an experiment has a successful outcome, it will have demonstrated that it would be practical to try to design human-machine interfaces in which an observer would be able to signal detection of a camouflaged target within the time required to collect the minimum required number of data points. Significantly, such a signal would not involve a motor response from the observer, opening the door to potential areas of application in which a motor response is not practical or even available.

## **14 Correlation Dimension Analysis**

### **14.1 Introduction**

The correlation dimension is an estimate of the complexity of a dynamical system. More precisely, the correlation dimension represents a lower bound on the number of variables that are required to adequately model a dynamical system (Farmer, Ott and Yorke, 1983). In relation to the area of study addressed by the present work, it is suggested that a prototypical example of such a system is the organization of intercortical signaling underlying perceptual and cognitive processes. An analysis of the correlation dimension computed for EEG recorded during epochs of dissimilar types and scales of perceptual and cognitive activity therefore is expected to show an effect of these dissimilarities, reflected in differences in the computed estimates of correlation dimension.

In providing a context for this analysis, the following brief review includes studies that represent several areas of application of measures of dynamical complexity (correlation dimension, symbolized as  $d_2$ ) and sensitivity to initial conditions (Lyapunov exponents). These areas include the relationship between these measures and the effects of task and of pathology. Arising from such studies are a number of issues related to methodology, and to the general applicability of these nonlinear analytical measures to analysis of neuroelectric signals. In addressing these general issues of applicability, the following review includes studies involving a variety of independent variables whose scope exceeds the area of visual perception that is the particular focus of the present work.

#### **14.1.1 The Effect of Task**

Correlation dimension has been found to index task complexity, showing a larger value for more complex cognitive tasks relative to less complex tasks. The comparative effect on correlation dimension of mathematical tasks versus a rest condition was examined using five subjects in 3 eyes-closed conditions, resting, serial addition by 2's, and serial subtraction by 7's (Rapp, Bashore, Martinerie, Albano, Zimmerman and Mees, 1989). A resting condition preceded and followed each arithmetic condition. The first finding was that the average value of correlation dimension across all resting conditions was found to be 3.9, while for both of the arithmetic tasks the value was 4.8. The second finding was that the increase in dimensionality from a preceding rest period to an arithmetic task was greater for the subtraction task than for the addition task. This finding suggests, according to the authors, that the subtraction condition involves the relatively more complex task.

Similar results were found by Nan and Jinghua (1988) in a study of relative hemispheric involvement in a mental arithmetic task. Correlation dimension was calculated for recordings made of scalp potentials at several electrode sites before, during and after the arithmetic task. Three subjects participated in this study, one left-handed and two right-handed. Scalp potentials were measured at six electrode sites, FP1, FP2, T3, T4, O1 and O2. For all subjects correlation dimension was significantly affected by task for temporal recordings only. For the left-handed subject the right temporal area showed an increase in dimensionality during the arithmetic task, while the left temporal area showed no change in dimensionality. For both of the right-handed subjects the opposite pattern was observed: During the arithmetic task the left temporal area showed an increase in dimensionality while the right temporal area showed no effect. Following the arithmetic task, correlation dimension for the affected temporal areas returned to pre-task levels. In sum, both prior to and following the mental arithmetic task, dimensionality estimates of the electrical activity at both temporal areas were approximately equal. During the arithmetic task however, dimensionality increases were found in the recordings from temporal areas contralateral to subjects' handedness.

Estimates of correlation dimension for alpha frequency components of the EEG were computed for 6 subjects in an eyes-closed condition (Basar, Basar-Eroglu, Roschke and Schult, 1990). Simultaneous recordings were made at 4 midline sites, vertex, parietal, occipital and frontal. Signals were filtered at 5 to 15 Hz. Overall, correlation dimension at each of the 4 sites varied between 5.5 and 7.5. It has been demonstrated however that frequency filtering can have an effect on the value of correlation dimension. Badii, Broggi, Derighetti, Ravani, Ciliberto, Politi, and Rubio (1988) for example found that correlation dimension increases with filtering. Intuitively, averaging would have an effect on the geometric properties of the phase-space attractor, thus affecting value of the corresponding dimensionality estimate. While the authors concluded that the dimensionality estimates showed a convergence in value across the measured sites, such convergence might reflect the effects of the common filtering process. Basar et al. (1990) suggest nevertheless that for similar initial conditions EEG activity, as reflected in the phase-space attractor and hence correlation dimension, is reproducible.

Using an analysis of the coherence of the EEG, Basar et al. (1990) found support for the view that alpha EEG contains a deterministic, task-related component. While coherence is not in the class of nonlinear techniques discussed in the present Unit, the following description illustrates the value of applying both linear and nonlinear techniques to an analysis application. Recordings were made from 5 subjects at electrode sites Cz, P3, P4, O1 and O2, against an earlobe reference. Subjects were instructed to attend to an 800 ms light-intensity step stimulus presented every 2 seconds, with missing stimuli. Subjects were asked to predict and count the occurrences of the missing stimuli. In the easy condition, every fourth stimulus was missing. In

the difficult condition every 4 to 7 stimuli were missing. The paradigm also included an eyes closed control condition. Recordings were made beginning 1 second before the missing stimulus, and included the evoked response to the missing stimulus. In the easy condition alpha EEG produced in the interval between 300 and 1000 ms prior to the missing stimulus was phase coherent between separate missing stimulus events. These EEG segments were correlated to the extent that the subjects were able to mentally track the missing stimulus. In both the difficult and control conditions there was significantly less phase coherence between alpha responses to the missing stimulus, a result consistent with the observation that in the difficult condition subjects had relatively more difficulty in tracking the missing stimulus events. These findings, the authors suggest, indicate the finding of a coherent brain state during which frequency components in the alpha range were phase-locked to an external signal.

Basar et al. (1990) propose that these alpha coherences support the results of the correlation dimension analysis, that the EEG reflects a deterministic cognitive process, and more particularly in their study, an attentional process. Subjects who could attend to the missing stimuli sufficiently well demonstrated their attentiveness in terms of alpha-range EEG signals that were phase-locked to the stimuli being attended. The authors note as well that correlation dimensions measured across the 4 sites showed very different patterns over the time-span of the recording. They point out that different recording locations may show entirely different patterns of activity, when examined using measures such as correlation dimension and spectral analysis. For this reason, they suggest, such measures are usefully combined with other techniques, such as inter-trial phase-coherence.

Correlation dimension has also found to increase relative to a resting condition during a sequence-learning task (Gregson, Britton, Campbell and Gates, 1991). Correlation dimension was computed for EEG data recorded from 6 subjects in 4 conditions, an eyes-closed resting condition, and three light-stimulus prediction conditions. Recordings were made from electrode pairs O1 to O2, and F3 to F4. Subsequent inspection of the apparent amount of noise in the records led the authors to use only the O1 and O2 recordings for analysis. The three prediction conditions differed in terms of the relative probabilities of the light stimuli, with the third prediction condition being the most difficult. Multiple estimates of the correlation dimension were computed for the recorded EEG. Although there was considerable variability in the dimensionalities both within and across subjects, there was an overall increase in dimensionality from the resting condition ( $d_2 = 8.0$ ) to the prediction conditions ( $d_2 = 9.3, 8.8$  and  $8.1$ ), and with no significant difference among the prediction conditions themselves.

Effects of task complexity were also found in a partial replication of Gregson et al. (1991), but using different electrode positions, Fp1 to P1 and Fp2 to P2 (Gregson, Campbell and Gates, 1992). Each dimensionality value was computed by averaging correlation dimensions for

30 samples for each subject-channel combination. Again differences were found between the eyes-closed condition ( $d_2 = 7.75$ ) and the light-prediction conditions ( $d_2 = 8.3$ ). An effect of task difficulty was also found, with a higher correlation dimension for the most difficult prediction condition ( $d_2 = 8.37$ ) than for the two less difficult prediction conditions ( $d_2 = 8.3$  and  $8.27$ ). The authors note that although there appears to be support for the notion that correlation dimension increases with task complexity, there is at the same time, and as in their previous experiment, a great deal of variability in the dimensionality values both within and across subjects. In sum, estimates of dimensionality using data from occipital recordings showed a difference between task and rest conditions only, while estimates using fronto-parietal recordings showed an additional effect of task complexity. The topography of the values of correlation dimension would appear to contain useful information about the distributional characteristics of neuronal system complexity within and between tasks.

Both Gregson et al. (1991) and Gregson et al. (1992) would appear to be relatively methodologically sounder studies of the correlation dimension, in that multiple values of correlation dimension were computed for a single subject-trial combination. This procedure allows an estimate to be made of the stability across subjects and across trials of the correlation dimension value. These studies found substantial variability in the value of dimensionality, a finding that may help to explain the divergent estimates of dimensionality apparent when comparing results across studies.

In a two-part study, dimensionality estimates were computed in relation to first, a variety of different tasks, and second, degree of hypnotizability (Ray, Wells and Elbert, 1991). In the first part of the study, 12 subjects were engaged in 6 tasks, consisting of visualization tasks, tactile sensory tasks, an observation task, and a verbal alliteration task. Dimensionality was found to be highest for the visualization tasks ( $d_2 = 5.2$  to  $5.4$ ), followed by the tactile and observation tasks ( $d_2 = 4.7$  to  $4.8$ ) and lowest for the verbal task ( $d_2 = 4.4$ ). In the visualization task dimensionality was approximately equal for frontal, parietal and temporal channels. All other tasks showed a lower dimensionality for frontal than for frontal and temporal channels. These spatial differences might be related, the authors suggest, to the dissimilar processes that might occur in the one case during the visualization tasks which involve internally directed attention, and in the second case during the tasks which require some interaction with the external environment.

In the second part of the study, estimates of dimensionality were related to degree of hypnotizability. A pool of 600 subjects was screened to select a group of 60 which was further screened to form a high and a low-hypnotizability group. Subjects underwent a protocol consisting of a baseline period followed by battery of tasks including imagery, spatial manipulation, mental arithmetic, and a Stroop naming task, followed by hypnotic induction, then

a repeat of the tasks, then removal of the trance state, and ending with a second baseline period. No difference in dimensionality was found between the high and low hypnotizables during the baseline periods ( $d_2 = 5.2$ ). Following the induction procedure however there was a significant difference between the two groups. High hypnotizables showed a higher dimensionality ( $d_2 = 5.5$ ) than low hypnotizables ( $d_2 = 5.2$ ). Fourier analysis of theta activity revealed an interaction between state and hypnotizability. Prior to induction, high hypnotizables showed more theta activity than low hypnotizables. This difference disappeared following induction. This double dissociation would appear to indicate that different processes are tapped by the nonlinear correlation dimensionality analysis, and the linear Fourier analysis. The nonlinearity of brain dynamics thus may not be fully describable in terms of a linear model such as Fourier analysis, but instead requires the contribution of nonlinear descriptors such as fractal dimensionality.

Corroborative findings of higher dimensionality estimates during visualization come from 2 studies of the effect of task on correlation dimension, and alpha and beta power (Lutzenberger, Elbert, Birbaumer, Ray and Schupp, 1992). These studies were intended to extend the results of previous studies (e.g., Ray et al. 1991) by involving several modalities, and by using Fourier analysis.

The first study tested the effect of task on dimensionality and both alpha and beta power, by engaging subjects in tasks involving tactile perception (determining the smoothest of a selection of sandpapers), vision (observing a double pendulum swinging), and imagery (imagining a past emotional experience). EEG recordings were made over 16 second intervals producing 2048 data points. A significant effect of task was found on all measures. Correlation dimension and both alpha and beta power increased monotonically from the visual perception task, to the tactile perception task, to the mental visualization task. The alpha power results, higher alpha power in the visualization than in the perception tasks, are consistent with previous findings that alpha power is higher during tasks involving inwardly-directed attention than during tasks in which attention is directed outwardly (Ray and Cole, 1985). The higher dimensionality for the visualization task suggests that this task is associated with a relatively more complex neural dynamical state than the two perception tasks.

In the second study the finding from the first study of dissimilar effects of visual perception and mental visualization was reexamined, but using the same object as the referent for both conditions. In the frontal areas only, a higher dimensionality was found for the visual imagery condition than for the perception condition. Both alpha and beta power showed increases in the visualization condition, mainly in parietal, but also in frontal areas. The finding of a difference in dimensionality between object perception and visualization in frontal areas is consistent with the results of metabolic examinations showing increased frontal metabolism during thinking as compared with perceiving (Roland, 1982). Lutzenberger et al. (1992) suggest

that dimensionality analysis complements traditional techniques of EEG analysis, which in their view are atheoretical and descriptive.

A study of correlation dimension and its relationship to intelligence (Lutzenberger, Birbaumer, Flow, Rockstroh and Elbert, 1992) found evidence of such a relationship, but only during rest and not during task performance. Subjects were tested for intelligence using the Cattell culture fair intelligence test. Subjects were divided on the basis of the test results into a low IQ group (a mean IQ of 84.4), and a high IQ group (a mean IQ of 118.2). Subjects were then engaged in 2 task conditions. In the first condition, EEG's were recorded from subjects during a no-instruction rest period. In the second condition, which followed the rest period, EEGs were recorded while subjects engaged in a mental imagery task. Data samples of 2048 points were recorded, over an interval of 20.48 seconds. Estimates of  $d_2$  were computed for each sample.

Dimensionality, which the authors viewed as indicative of relative complexity of neuronal activity, was found to be higher for the high IQ group than for the low IQ group. This difference was significant during the resting condition in the parietal region, but was not significant during the imaging condition. A significant correlation between IQ and dimensionality was found at electrode Pz, with dimensionality accounting for about a quarter of the variance in IQ. Interestingly, measures of power in the delta, theta, alpha and beta ranges showed no difference between the IQ groups. As an explanation for these findings the authors propose that in the resting condition, higher IQ subjects manifest a greater number of simultaneously active neuronal assemblies, and thus a higher dimensionality, than low IQ subjects. In the imaging condition, task requirements impose equal restrictions on the 2 groups in terms of the number of activated cell assemblies, leading to a smaller difference in dimensionality between the groups.

The effect of processing load was investigated by computing estimates of  $d_2$  and measuring alpha power for EEG's recorded from 12 subjects during no-task eyes open and eyes closed conditions (Pritchard and Duke, 1992). Blocks of eyes open and closed conditions were repeated 4 times. The authors prefer the term dimensional complexity for their measure, in view of what they see as the limitations in the available EEG data with respect to the correlation dimension algorithm. These limitations include first, violations of requirements for an unlimited amount of noise-free data, and second, for a statistically stationary dynamical process. Dimensional complexity was found to be lower in the eyes closed than in the eyes open condition. The increased dimensional complexity in the eyes open condition was found to be well accounted for by the data from the occipital channels. A similar and inverse pattern was found for alpha power which decreased in the eyes open relative to the eyes closed condition. Finally, across the 4 blocks of eyes open and eyes closed tasks, dimensional complexity decreased, while alpha power increased. These results were interpreted as indicating that dimensional estimates may be useful in a relative sense, as an indicator of processing load. The

authors point out however that dimensional estimates may not be valid in an absolute sense. The EEG signal may not fulfill the requirement of representing a stationary process, although there appears to be no generally accepted criterion for determining when a shift in stationarity of the EEG occurs. The non-stationarity of the EEG would appear to be indicated by the finding that dimensionality changed across the 4 identical condition blocks.

#### **14.1.2 The Effect of Pathology**

A number of studies have explored the relationship between nonlinear measures such as correlation dimension and Lyapunov exponents and organic pathological conditions, principally epilepsy and Creutzfeld-Jakob disease.

Comparing states of arousal with epileptic seizure activity, Babloyantz and Destexhe (1986) found correlation dimension values to be lowest for epileptic activity and REM sleep, higher for stage 2 and 4 sleep, and highest for wakefulness. Estimates of the correlation dimension were computed for wakefulness, REM, stage 2 and stage 4 sleep in a normal subject, and an epileptic seizure event in an epileptic subject. For the wakefulness condition, the computation was not able to produce a bound on the correlation dimension, suggesting a high value of dimensionality. A similar result was encountered for the REM sleep condition. For both stage 2 and stage 4 sleep, correlation dimension was computed to have a value of between 4 and 5. For the epileptic seizure events, recordings were made of the differential signal between frontal and parietal regions, and between the vertex and temporal regions. The same results were found for all channels: correlation dimension for epileptic seizure activity was found to be approximately 2. A similar pattern of results was found for the correlation dimension computed using data recorded from the limbic cortex of a rat, during rest, locomotion and an epileptic seizure induced by kindling (Pijn, Van Neerven, Noest and Lopes da Silva, 1991). For the rest and locomotion conditions dimensionality was found to be unbounded and high. For the seizure condition a dimensionality of between 2 and 4 were computed.

These distinctions between wakefulness, sleep and pathology have been explored using a number of different nonlinear measures (Gallez and Babloyantz, 1991). Several nonlinear analyses were applied to EEG recorded during wakefulness with eyes closed, stage 4 sleep, and Creutzfeld-Jakob coma. The first measure involved calculation of Lyapunov exponents, indicators of sensitivity to initial conditions and hence the presence of deterministic chaos. The second method involved computation of Kolmogorov entropy, a measure of the rate at which new information is produced, or the mean time for which a signal can be predicted. The third method used calculation of attractor dimensionality, an estimate of generating system complexity. Attractor dimensionality was estimated using the correlation dimension and two

other related measures of dimensionality based on Lyapunov exponents, the Kaplan-Yorke and the Mori conjectures.

There were three main findings. First, the number of positive Lyapunov exponents was highest for the wakefulness condition, and lowest for the coma and sleep conditions. The greater the number of positive exponents, the greater is the sensitivity of the system to perturbations, and therefore the more chaotic the underlying generating system. The authors interpreted their findings to mean that the higher chaotic level in the wakefulness condition make possible a greater variety of behaviours. Second, metric entropy was found to be higher during wakefulness than during deep sleep. The greater the metric entropy, the greater the rate at which the system is producing information, or in alternate terms, the less predictable it is. The authors relate this finding to the greater rate of information processing during wakefulness relative to deep sleep. The third finding was that bounded, and similar, values of dimensionality were computed by the correlation dimension, and by two estimates of dimensionality based on Lyapunov exponents. The bounded dimensionality values indicate, the authors proposed, the presence of strange attractors during the phases of brain activity which were studied. The convergence in dimensionality values produced by the three methods that were used would indicate that estimates of dimensionality are at least to some extent robust in the face of alternative computational approaches.

The authors note however that typically there is a great deal of variance in the dimensionality values even using the same estimator with different data samples, and suggest that experimental situations should be arranged to provide a clear distinction between the types of tasks that are used, and thus between the associated neural activities. They emphasize that dimensionality estimates are most effective when used to distinguish between the effects of different types of task requirements, rather than when used as indicators of absolute complexity of neural dynamics.

Correlation dimension, Lyapunov exponent, and autocorrelation estimates were computed for EEG recorded during an epileptic seizure in a single subject study (Frank, Lookman, Nerenberg, Essex, Lemieux and Blume, 1990). In contrast with Babloyantz and Destexhe (1986), Frank et al. (1990) were able to obtain a longer-term recording of a seizure, lasting approximately 75 seconds, and including both absence and grand-mal events. The authors estimated the stationarity of the record by looking at the variance of different portions of the entire recording, along principal component axes. They suggest that what they refer to as dynamical nonstationarity - changes in the dynamical properties of the record - would be indicated by different variances along the different dimensions, and between the different portions of the record. No evidence of this was found and the authors conclude that their signal record was statistically stationary. Correlation dimension was found to have a bounded value of

5.6, consistent with the presence of a strange attractor in the dynamics underlying the epileptic activity. The first Lyapunov exponent was 1, consistent with the computed decay rate of the autocorrelation function, and suggesting a chaotic component to the underlying dynamics. Essentially the same results were obtained with the same subject after a 3 month interval.

These findings support the position, the authors conclude, that the neural dynamics during an epileptic seizure are deterministically chaotic, a determination that could not be made without the calculation of Lyapunov exponents. The authors point out that the ubiquitous myoelectric noise contamination of EEG records would have an effect on both the computation both of dimensionality and Lyapunov exponents. They can see no solution to this problem other than intracranial recording.

In an investigation of the association between nonlinear measures and alcohol consumption, Palus, Dvorak and David (1992) found an inverse relationship between blood alcohol level and dimensionality. The authors studied the effect of alcohol intoxication on two measures, the correlation dimension, and a measure which they term linear complexity. Linear complexity they defined as the negative inverse of the sum of the logarithms of the eigenvalues. The authors found that both of these measures were well correlated with level of blood alcohol. Both measures decreased in magnitude with increasing blood alcohol. Intuitively, since linear complexity is a function of the number of significant eigenvalues, it would seem reasonable that this measure show index system complexity in a similar manner to the correlation dimension.

#### **14.1.3 Methodological Issues**

In a study which focused on methodology, Destexhe, Sepulchre and Babloyantz (1988) compared several techniques for computing the correlation dimension of an EEG recording, the standard Grassberger - Procaccia algorithm, a hybrid of this algorithm with singular value decomposition aimed at reducing the noise components in the data, and a multi-channel version of Grassberger - Procaccia method in which data from multiple channels, rather than from only one channel, is analyzed. The authors computed correlation dimension during a number of conditions including Creutzfeld-Jakob seizure ( $d_2 = 3.7$  to  $5.4$ ), alpha ( $d_2 = 6.1$  to  $7.4$ ), deep-sleep ( $d_2 = 4.4$  to  $4.5$ ), and epileptic seizure ( $d_2 = 2.03$  to  $2.05$ ). They concluded that the results of these three methods agree only if the value of correlation dimension is less than 4. It may be observed that in some of the reviewed studies this requirement is not met. Destexhe et al.'s (1988) results suggest that correlation dimension may not be a robust estimator, unless the conditions under which it is calculated are carefully and completely specified.

The study by Dvorak and Siska (1986) work points up some of the difficulties involved in estimating the correlation dimension. Correlation dimension estimates were computed for EEG

recordings were made at sites O1, O2 and C4 using adult male subjects in 3 conditions: vigilant eyes closed, relaxed eyes open, and mental arithmetic (subtracting by 13's down from 1000). The pattern of dimensionality changes was found to be different at the occipital and central sites. At O2, dimensionality was low ( $d_2 = 5.7$ ) for the eyes closed, and equally high ( $d_2 = 6.5$ ) for both relaxed eyes open and arithmetic task conditions. At C4, dimensionality was highest in the eyes closed ( $d_2 = 5.5$ ), medium in the eyes open relaxed ( $d_2 = 5.3$ ), and lowest in the arithmetic task condition ( $d_2 = 4.7$ ). The authors studied the effect of signal stationarity, and of filtering. Using a sample of 1,000 points from a record of 15,000 points, they found that correlation dimension varied with the ordinal position of the sample in the entire record ( $d_2 = 3.8$  to  $5.5$ ). Filtering the entire record with a 30 Hz low-pass filter reduced both the absolute values of dimensionality, and the variability with respect to sample ordinal position ( $d_2 = 3.8$  to  $4$ ).

These results may be summarized in terms of four conclusions. First, the results support, generally, other findings (e.g., Nan and Jinghua, 1988, Rapp et al., 1989) of dimensionality changes with cognitive task. Second, the findings support the observation in other studies that scalp location interacts with the relationship between dimensionality and task condition. Third, they point up the sensitivity of the correlation dimension estimate to signal pre-processing. Fourth, the variability of the dimensionality estimates using the unfiltered data would appear to support the contention that the EEG is a statistically nonstationary signal, reflecting a nonstationarity in the underlying dynamics.

Rapp et al. (1989) also surveyed the use of the correlation dimension in the analysis of EEG recordings by different groups, and present a rationale for using the correlation dimension. The authors suggest that, in comparison with other statistics, the correlation dimension uses more of the information present in a time-series such as the EEG. The correlation dimension, they suggest, is therefore a more robust characterization of the behaviour of such a system. Rapp et al. (1989) also compared the standard computational procedure for the correlation dimension with a hybrid method combining the Grassberger - Procaccia algorithm with singular value decomposition, an idea which was presented by Broomhead and King (1986). The matrix of vectors formed by embedding the time-series in a phase-space is filtered by means of the singular value decomposition. Only the most significant components are then used to generate an estimate of the original matrix, which is then used in the conventional Grassberger - Procaccia computation. The result is a uniform redistribution of noise among the retained components, which form the basis of an embedding phase-space of reduced dimensionality. The latter feature enhances the robustness of the Grassberger - Procaccia computation, while, since only a reduced matrix is analyzed, reducing the computational overhead.

The authors note however the variability in published dimensionality estimates for a variety of experimental conditions. They describe some of the difficulties associated with the

dimensionality computation which may be the cause of the variance in published dimensionality estimates. Besides issues such as selection of algorithm parameter values (e.g., lag, maximum embedding dimension), data collection variables (e.g., sampling rate), they suggest that the EEG is generally not a statistically stationary signal. In other words, statistical properties of the EEG may change significantly over the course of a recording session. The effect of such nonstationary behaviour on the correlation dimension is, they state, not completely understood.

Nonstationarity of the EEG is also addressed by Jansen (1991). He suggests that, while there is evidence for the chaotic nature of the EEG, that, in other words, the EEG is a reflection of an underlying nonlinear dynamical system, the EEG does not meet the requirements that would allow measures such as dimensionality or Lyapunov exponents to be calculated. In particular, he notes, these measures make the assumption that the signal being analyzed is a reflection of a system which has evolved through its asymptotic region and has converged to a pattern of behaviour which is statistically stationary. The question of whether this is the case for the EEG, he suggests, appears to be unanswerable.

It should be noted however that some authors, for example, Rapp et al. (1989) refer to the dimensionality estimate as the correlation index, rather than correlation dimension, in recognition of the EEG's nonstationarity. They suggest, with Gallez and Babloyantz (1991), that the dimensionality estimate is therefore best used in a comparative sense, as an index of the difference in brain dynamical complexity across tasks, rather than as an absolute measure of brain dynamics.

Mayer-Kress and Layne (1987) computed correlation dimension for existing data obtained from a number of sources, and compared their results with those of other studies. The data represented the conditions of resting eyes open and closed, sleep stages II and IV, REM sleep, petit-mal seizures, anesthesia, and verbal memory, visual memory, abstraction and word association tasks. In most cases, the values of correlation dimension were associated with uncertainties in the values that were on the same order of size as the dimensionality estimates themselves. The exception was the awake, eyes-closed, resting condition. The authors concluded that in general correlation dimension cannot be effectively computed due, they suggest, to the nonstationary nature of the EEG, and due to the formal data requirements of the correlation dimension computation procedure. They propose that correlation dimension should be used only in a comparative sense, with subjects acting as their own controls.

A critical analysis of the use of the Lyapunov exponent with EEG data is presented by Principe and Lo (1991). The authors computed Lyapunov exponent for EEG recorded during stage II sleep. They concluded that the range of values which they calculated (2 to 4) represents only an order of magnitude estimate for the following reasons. First, the EEG is statistically nonstationary. Second, the Lyapunov exponent algorithm which they used and which is the only

one so far available (due to Wolf, Swift, Swinney and Vastano 1985) requires knowledge of the generating dynamical system that is unavailable in the case of the EEG. The authors do point out however that the positive sign of their computed values is consistent with the view that the EEG reflects a deterministically chaotic rather than a stochastic dynamical process.

In summary, both correlation dimension and Lyapunov exponents have been used to analyze EEG recordings made in a number of different sets of conditions, involving both clinical and non-clinical subject groups, and a number of different task conditions. The consensus would appear to be that, because of data limitations such as the statistically non-stationary nature of the EEG, correlation dimension can best be viewed as an index of relative, rather than absolute, system complexity. Correlation dimension appears to increase with task complexity, and in tasks such as mental visualization relative to resting conditions. The positive values calculated for Lyapunov exponents suggest that the EEG reflects a chaotic generating system. Again however, EEG nonstationarity on the one hand and unavailability of easily-applied algorithms on the other hand, suggest that caution should be applied in interpreting the findings as indicating the presence of chaos in the EEG.

## **14.2 The Hypothesis**

The present experiment is aimed at finding out whether changes in cognitive and perceptual processes can be found to be reflected in corresponding changes in the correlation dimension computed for time series recordings of EEG. Estimates of the correlation dimension will be computed for EEGs recorded during two conditions. In the first (picture) condition, subjects inspecting an image containing a camouflaged target object are able to discriminate the target from the background, signaling the discrimination event with an eye-blink. In the second (control) condition subjects blink at will while fixating on a neutral screen. It is expected that the neuronal events over a short interval preceding discrimination will be reflected in a higher value of correlation dimension than the neuronal events that are associated with spontaneous eye-blinks while fixating on the neutral screen. This hypothesis is made on the basis of the results of the cross-correlation analysis, which demonstrated that over a 1 second epoch preceding discrimination there was evidence of significant changes in the configuration of interregional association. It is this change in the pattern of interregional communication that is proposed to be responsible for a higher value of correlation dimension, relative to the control condition. The results of the cross-correlation analysis showed that over the 1 second epoch preceding the blink in the control condition there was a relatively less change in the pattern of interregional association.

## **14.3 Method**

### **14.3.1 Preprocessing**

The following operations were carried out on the recordings from every subject and every trial, and for both control and picture conditions. First, the 1 second (128 points) intervals representing the before-blink epoch and the after-blink epoch were isolated from the matrix of data for each trial. The result was, for each trial, a pair of matrices, each 128 time-points by 16 channels. No other transformations were performed on the data. Next, each of these matrices was used to construct a phase-space attractor in a 16-dimensional space. The dimensionality of each of these attractors was then computed, as the estimates of correlation dimension. The construction of the phase-space attractor and the correlation dimension computation are described in the following section.

### **14.3.2 Correlation Dimension**

Correlation dimension represents an estimate of the lower bound on the number of variables that are involved in the dynamical behaviour, or evolution over time, of a multivariate system. The computation of correlation dimension involves first the construction of a phase space attractor. This attractor is essentially a graph plotted on a multi-dimensional set of axes. The axes represent the various dimensions of the multivariate system. Each point on the attractor represents the value, at a particular time, of each of these dimensions, and thus represents the state of the multivariate system at that time. Once the attractor has been constructed, its dimensionality can be computed. This dimension will in general have a non-integer value. Attractors with such non-integral values of dimensionality are referred to as strange attractors. Algorithms have been developed that allow attractor dimensionality to be computed. One of the easiest to apply is the Grassberger-Procaccia algorithm (Grassberger and Procaccia, 1983).

The present analysis makes use of a multiple time-series version (e.g., Destexhe et al., 1988) of the Grassberger-Procaccia algorithm. The original single time-series version of this algorithm was intended to provide a means by which a dimensionality estimate could be computed when data from only a single variable of the multivariate system under study was available. This algorithm was based on a result demonstrated by Takens (1980) showing that a number of dynamical properties of the underlying system are preserved when the time series is used to reconstruct a multi-dimensional phase-space attractor. In particular the dimensionality of the multivariate process is represented as the dimensionality of the attractor, when a number of conditions are met. These conditions raise a number of problematical issues when the method is

applied to real data. In the standard algorithm a single time series is used to generate a phase space attractor for which dimensionality can then be computed. This process involves embedding the time series in a phase space of suitable dimensionality. Two key parameters are involved in this embedding. The first is the choice of embedding dimension. Schaffer, Truty and Fulmer (1988) suggest a value of at least  $2m + 1$ , where  $m$  is the hypothesized value of dimensionality of the attractor. Since  $m$  is not known beforehand, prudence dictates a large value of embedding dimension. The larger this value however the more demanding are the data requirements. As embedding dimension increases, the number of data points that are required to construct the attractor correspondingly increase. The second key parameter involved in embedding the phase space attractor is lag. The lag parameter is used in the selection of sets of points from the time series to serve as the coordinates of individual points on the attractor. While with an unlimited quantity of noise-free data the choice of lag is not critical, with a limited number of noisy data points the value of lag can be critical. In particular dimensionality in such a real case will vary with the value of lag.

These limitations of the Grassberger-Procaccia algorithm are circumvented in the present approach of using the records of multiple variables, that is EEG channels, from the neuronal generating system. This approach is of course only possible when multiple time series are available. Using this method, the embedding process is eliminated. The starting point for this approach is the matrix of data points, organized as  $m$  data points by  $p$  channels. The phase space attractor is created by simply taking the  $p$  data points corresponding to a single time point, and then using these data points as the coordinates of a single point on the attractor in a  $p$ -dimensional phase space. The process is repeated for all  $m$  time-points in the original data matrix. Once the attractor has been created, the correlation integral  $C(r)$  is computed. For each of a series of values of a parameter  $r$ , the correlation integral is the probability of finding an attractor point within a distance  $r$  of a given reference point also on the attractor. This probability is averaged over a number of points on the attractor,  $n$ . In the present case,  $n$  is equal to the total number of points on the attractor,  $m$ . Thus,

$$C(r) = \lim_{n \rightarrow \infty} \frac{1}{n^2} \sum_{i \neq j=1}^n H\{r - |x_i - x_j|\} \quad \dots \text{eqn 14.1}$$

where  $x_i$  is a vector representing a reference point and  $x_j$  is a vector representing any other point on the attractor,  $n$  is the number of points on the attractor that are averaged over,  $r$  is the distance from a reference point,  $|\cdot|$  denotes the modulus, and  $H$  is the Heaviside function defined as

$$H(x) = 0; \quad x \leq 0 \quad \dots \text{eqn 14.2}$$

$$H(x) = 1; \quad x > 0$$

The Heaviside function thus simply counts the number of pairs of points that are separated by a distance less than  $r$ . Once the correlation integral has been evaluated for a range of values of  $r$ , the correlation dimension,  $d_2$ , is computed as

$$d_2 = \ln C(r) / \ln r \quad \dots \text{eqn 14.3}$$

The correlation dimension computation was carried out using the data analysis program Simulnet™ version 2.3.

#### 14.4 Results

The estimates of correlation dimension were found to be significantly larger for the 54 trials from the picture condition than for the 42 trials from the control condition, but only for the before-blink epoch and not for the after-blink epoch. Correlation dimension estimates are shown for all trials in Table 14.1, and in Figure 14.1 for the before-blink epoch, and Figure 14.2 for the after-blink epoch.

For the before-blink epoch, the mean values of  $d_2$  were highly significantly different between the picture and control conditions. The mean correlation dimensions were 5.44 for the control condition and 5.96 for the picture condition. The resulting value of  $t$  was 2.24 ( $p = 0.014$ ). For the after-blink epoch, the mean values of  $d_2$  were only marginally significantly different between the control and picture conditions. The mean correlation dimensions were 5.86 for the control condition and 6.32 for the picture condition. The resulting value of  $t$  was 1.55 ( $p = 0.063$ ). The means are shown in Figure 14.3, and the results of the  $t$ -tests are shown in Table 14.2.

**Table 14.1 Correlation Dimension Estimates**

Each correlation dimension value represents an estimate of the relative complexity of the dynamical system underlying the EEG measurements during that trial. The Group label is coded as 0 for the control condition, 1 for the picture condition. BBE indicates the before-blink epoch. ABE indicates the after-blink epoch.

Trial	Group	BBE	ABE
1	0	6.03	5.5
2	0	4.64	4.32
3	0	4.38	4.82
4	0	4.55	5.62
5	0	4.69	4.98
6	0	4.24	5.91
7	0	4.93	4.29
8	0	5.2	5.41
9	0	5.83	5.7
10	0	5.08	5.71
11	0	4.69	6.26
12	0	4.63	7.29
13	0	4.41	4.22
14	0	4.44	4.13
15	0	5.04	4.86
16	0	5.49	4.68
17	0	4.43	4.29
18	0	4.75	4.07
19	0	5.82	4.75
20	0	4.36	4.4
21	0	4.88	5.69
22	0	6.07	7.99
23	0	6.15	4.45
24	0	5.4	5.19
25	0	4.97	6.52
26	0	5.62	4.37
27	0	4.15	5.73
28	0	6.44	4.48
29	0	6.13	5.5
30	0	5.01	6.87
31	0	7.09	4.91
32	0	5.27	4.92
33	0	4.62	7.83
34	0	4.45	5.17
35	0	4.62	8.02
36	0	5.01	4.91
37	0	7.39	7.47
38	0	4.58	8.24
39	0	4.96	7.84
40	0	4.77	5.17
41	0	4.91	5.14
42	0	5.92	7.18
43	0	7.92	7.71
44	0	5.93	8.61
45	0	7.87	5.55
46	0	7.66	6.32
47	0	6.23	6.84
48	0	5.1	6.7

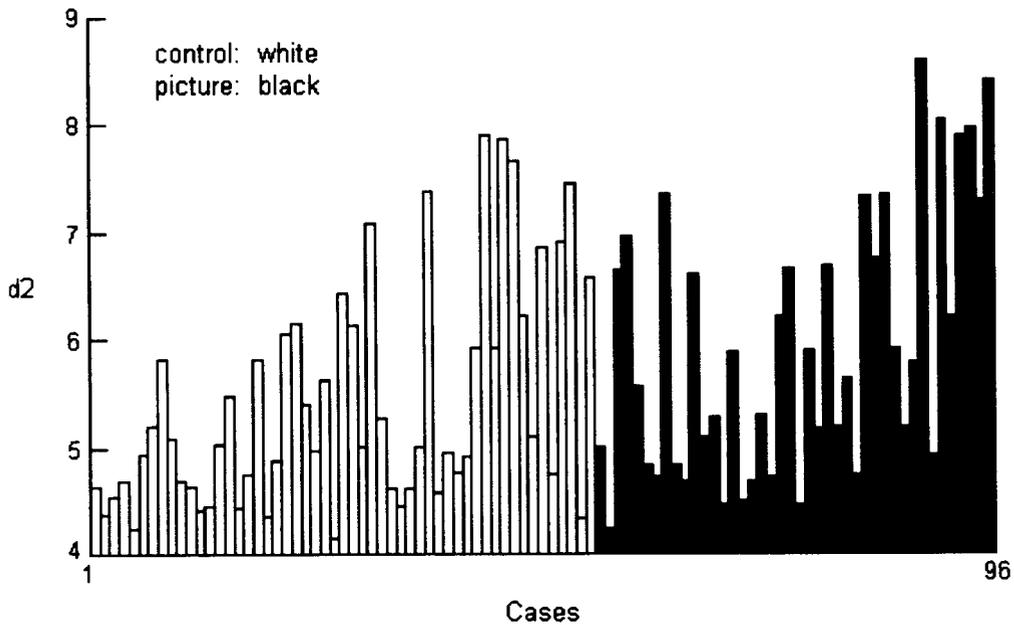
**Table 14.1 Correlation Dimension Estimates**  
(continued)

Trial	Group	BBE	ABE
49	0	6.86	8.12
50	0	4.75	8.08
51	0	6.93	7.47
12	0	7.47	6.27
53	0	4.34	5.46
54	0	6.58	4.65
55	1	5.02	5.12
56	1	4.24	4.34
57	1	6.67	8.1
58	1	6.98	4.14
59	1	5.58	4.24
60	1	4.84	4.23
61	1	4.73	4.4
62	1	7.36	4.12
63	1	4.84	6.42
64	1	4.7	4.44
65	1	6.61	7.68
66	1	5.1	4.69
67	1	5.28	6.07
68	1	4.46	6.58
69	1	5.89	5.09
70	1	4.51	4.18
71	1	4.69	4.56
72	1	5.31	7.3
73	1	4.72	5.54
74	1	6.23	4.85
75	1	6.67	6.63
76	1	4.46	6.42
77	1	5.9	7.42
78	1	5.17	7.92
79	1	6.7	5.72
80	1	5.2	6.56
81	1	5.65	4.78
82	1	4.74	5.08
83	1	7.34	6.88
84	1	6.76	8.36
85	1	7.37	8.29
86	1	5.93	6.29
87	1	5.21	9.17
88	1	5.8	6.6
89	1	8.63	8.49
90	1	4.94	6.66
91	1	8.06	7.19
92	1	6.23	8.78
93	1	7.91	8.32
94	1	8	8.14
95	1	7.31	9.03
96	1	8.44	6.68

**Table 14.2 T-test Results for Correlation Dimensions**

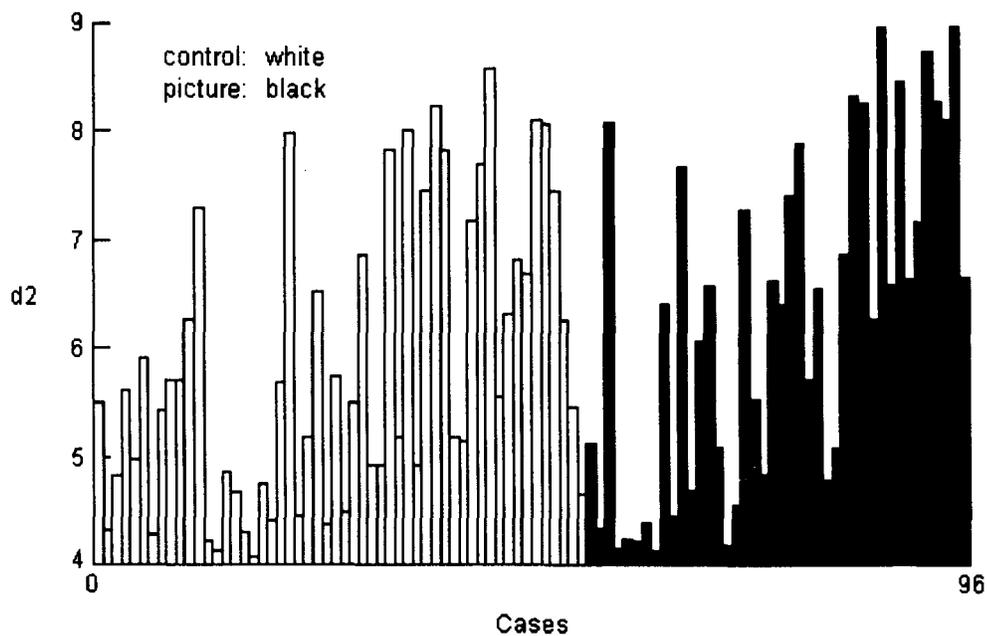
The values of *t* are computed for the two groups of scores, corresponding to the control and picture conditions, for the before-blink epoch and the after-blink epoch. The control group contains 54 scores, and the picture condition contains 42 scores, for a *df* of 94. In the before blink epoch, mean correlation dimension for the picture condition was highly significantly larger than for the control condition. In the after-blink epoch, mean correlation dimension for the picture condition was only marginally significantly larger than for the control condition.

Epoch	Student's <i>t</i>	probability	Mn (control)	Var (control)	Mn (picture)	Var (picture)
before blink	2.24	0.014	5.44	1.07	5.96	1.52
after blink	1.55	0.063	5.86	1.76	6.32	2.51



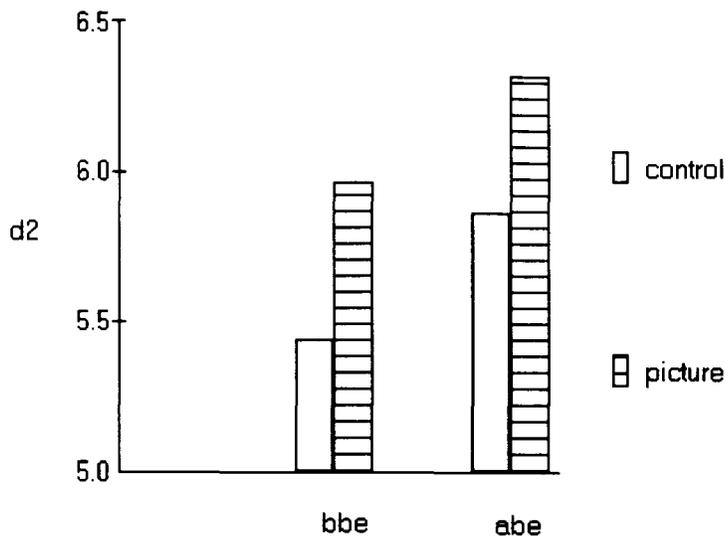
**Figure 14.1 Correlation dimension estimates for the before-blink epoch**

The figure shows the correlation dimension estimates for each of the 96 cases. Cases 1 to 54, shown in white, represent the control condition. Cases 55 to 96, shown in black, represent the picture condition. The mean correlation dimension for the picture condition is significantly higher than for the control condition.



**Figure 14.2 Correlation dimension estimates for the after-blink epoch**

The figure shows the correlation dimension estimates for each of the 96 cases. Cases 1 to 54, shown in white, represent the control condition. Cases 55 to 96, shown in black, represent the picture condition. The mean correlation dimension for the picture condition is marginally significantly higher than for the control condition.



**Figure 14.3 Mean correlation dimensions**

The figure shows the mean of the correlation dimension estimates for the control and picture conditions, for the 1 second epoch preceding the blink signaling discrimination (the before-blink epoch, BBE), and for the 1 second epoch starting after the cessation of blink artifacts (the after-blink epoch, ABE). There is a highly significant difference in  $d_2$  between the control and picture conditions in the BBE, and a marginally significant difference in  $d_2$  between conditions in the ABE.

## 14.5 Discussion

There was a significant difference in  $d_2$  between conditions in the 1 second interval preceding the eye-blink by which subjects signaled discrimination, with a larger value of dimensionality in the picture than in the control conditions. The larger value of  $d_2$  in the picture condition, preceding discrimination, supports the hypothesis made at the outset, that the neural processes occurring over a short interval preceding discrimination would be reflected in a higher value of correlation dimension than the corresponding processes associated with gaze-fixation on a neutral screen. This proposal was made on the basis of the cross-correlation results which showed that over the 1 second epoch preceding discrimination there were profound changes in the configuration of interregional signaling, changes which it was proposed would result in a higher relative value of  $d_2$  in the picture condition. In the 1 second interval following the eye-blink, there was a corresponding marginally significant difference, with again a larger value of dimensionality in the picture than in the control conditions. There were no predictions made at the outset in regards to this difference. Such a finding might be explainable however as indicating that following the eye-blink in the picture condition the re-fixation on the discovered

target image are associated with a relatively more complex dynamical behaviour in terms of the configuration of interregional signaling. Thus, following the blink in the picture conditions, subjects would re-fixate on the presented image and the discovered target object, and continue with some degree of visual analysis of the image. In contrast, following the blink in the control condition the re-fixation on the fixation point within a neutral screen would be associated with relatively less complex dynamical behaviour.

Another finding, for which no hypothesis had been initially advanced, was the result that values of  $d_2$  were higher, although not significantly, in the 1 second epoch following the blink than in the 1 second epoch preceding the blink. It could be conjectured that this finding is a reflection of a progression of more varied, less cognitively focused, visual analyses of the presented image following the blink than before the blink. Thus, it is suggested that over the 1 second interval before the blink subjects were engaged in a relatively cognitively focused search for the camouflaged target object. Over the course of the 1 second interval following the blink however, it is suggested that subjects would engage in a more varied sequence of cognitive tasks, starting with, for example, visually refocussing on the discovered target image, and proceeding to retrieving memory templates representing memory associations with the target object. As an alternative explanation, it might be possible that the after-blink epoch included neuronal processes associated with the termination of the eye-blink. These neural processes, which would occur only at the start of the epoch, would be followed by processes associated with more perceptually-oriented processing of the visual image, would then result in a relatively more complex pattern of dynamical behaviour in the after-blink epoch.

The finding of a difference in the estimate of correlation dimension between conditions and between epochs suggests that the estimate of correlation dimension as computed in the present analysis is a sensitive index of the quality of the neuronal activity underlying perceptual and cognitive events. One issue that must be addressed is the implications of the finding of a fractional value of dimension for the question of whether or not the neuronal system in question is behaving chaotically.

There are two points to be raised. First of all, computation of correlation dimension by itself does not speak to this question. The determination of the presence of chaotic dynamics would depend of the computation of a measure such as Lyapunov exponents, a measure that indexes the degree of sensitivity displayed by the dynamical system to changes in initial conditions, and to perturbations generally. In the present analysis no such index was computed, and therefore no statement can be made regarding the possible presence of a chaotic dynamic component to the behaviour of the relevant neuronal systems. The second point concerns the meaning of the computed values of what has been referred to as the correlation dimension. Recalling the discussion presented earlier regarding the data requirements of a correlation

dimension computation, it is probably more appropriate in the present case to refer to the computed values as estimates of a correlation index. Correlation index is, then, appropriately treated as a measure of the relative complexity of the corresponding dynamical systems, rather than as measures of absolute dimensionality (Rapp et al., 1989; Gallez and Babloyantz, 1991; Mayer-Kress and Layne, 1987).

Thus, the present findings of estimates in the region of from 5.44 to 6.32 may only with caution be interpreted as indicating that the corresponding neuronal dynamics can be adequately described in terms of between 5 and 6 variables. A statement regarding the absolute level of dimensional complexity would probably require a greater number of data points for the computation than are available in the present analysis. Smith (1992) for example suggests that for an estimated dimension of 5, an RMS error of 0.1 would require on the order of 5000 data points. What would appear to be safely concluded from the present results is that the cortical dynamics underlying a short interval preceding the moment of discrimination are, on average, significantly more complex than the dynamics underlying the corresponding interval in the control condition.

## **IV Discussion**

## 15 Summary of Findings

The paradigm used in the present study has addressed the question of the neuronal basis of visual discrimination, and through this the issue of the neural events associated with visual feature binding. This paradigm has been designed to present subjects with images constructed in such a way that, initially, the central representations of image features are visually unorganized and an intended target object is unperceived. Preceding successful discrimination, these representations, it is suggested, undergo a process of transformation that involves binding or integration of the representations into feature ensembles, as a result of which a unified percept is constructed through association of these feature ensembles with information contained within existing memory templates. Discrimination and, it is suggested, this process of feature binding and association, is observed to be associated with an increasing level of synchronization between increasingly spatially-extended cortical regions. More particularly, the findings of the present study, using three different measures of association, support the conclusion that, as the moment of visual discrimination approaches, the process of discrimination involves the coordinated activity of both more, and more widely separated cortical regions, involving the left hemisphere preferentially, and including occipital, temporal, central and frontal cortical regions.

The findings of the cross-correlation analysis were that in the picture condition, correlations increased over a short time interval preceding the moment of discrimination, while in the control condition, correlations remained relatively constant over a short interval preceding the voluntary eye-blink. In both picture and control conditions, and in the time intervals both before and after the blink, correlations decreased with increasing distance between electrodes. Finally, in the picture condition, correlations increased by a greater amount between a subset of widely separated electrode positions than between more closely spaced positions. This subset of electrode positions included scalp areas over occipital, temporal and frontal regions. Significant effects on lag were found only as a function of distance, with an increase in lag with inter-electrode distance prior to discrimination in both the control and picture conditions, and a decrease in lag with increasing distance prior to the blink in the control condition.

Coherence analysis showed that for a subset of electrode pairs coherence was higher in the picture condition than in the control condition. Consistent with the results of the cross-correlation analysis, this subset of electrodes included areas over occipital, temporal and frontal regions, with a greater involvement of the left hemisphere. In both conditions, the value of coherence was an approximately inverse quadratic function of inter-electrode distance, supporting the view (Thatcher et al., 1986) that EEG coherences are the result of axonal, rather than volume conduction.

The pattern of changes of mutual information was similar to that for cross-correlation. Mutual information decreased with increasing inter-electrode distance. Mutual information increased with time up to the moment of discrimination, and increased by a greater amount for more distantly spaced electrode pairs than for more closely spaced pairs. These findings suggested that the rate of information exchange between cortical systems increased with time up to the moment of discrimination. One conclusion that can be drawn from the similarity in the results of the cross-correlation and the mutual information analyses is that the relationship between the activity at different cortical regions appears to be reasonably linear.

The findings of these linear analyses were supported by the results of the nonlinear analyses. Correlation dimension was found to be higher for the picture condition than for the control condition. This finding reflects, it is suggested, the relatively more complex time-evolution of the interregional signaling configuration in the picture condition. That is, it was observed that there was a greater change in the pattern of interregional signaling, as indicated by interregional associations, over the 1 second interval preceding the eye blink, in the picture than in the control condition.

Based on these findings of a robust difference between picture and control conditions on a variety of linear and nonlinear measures, it was expected that a neural network classifier would be able to distinguish between the EEG recordings of the 1 second before-blink epoch for the picture and control conditions. It was in fact found that a generalized regression neural network could significantly well perform this discrimination.

The predictions made at the outset on the basis of the cortical self-organization model of visual discrimination were that (1) interregional associations should increase, over a short interval preceding discrimination, involving wide extents of cortex, rather than only particular regions, (2) the strength of association should vary inversely with interregional distance, (3) the strength of association should increase preferentially between widely separated cortical regions, and (4) these associations should be based on oscillatory signal components. The present findings would appear to be in general accord with these predictions.

## 16 A Neural Basis for Object Discrimination

The correlational activity studied in the present work was analyzed using oscillatory features within the recorded electrical activity. These features, which may include both periodic and aperiodic components, reflect corresponding oscillatory characteristics in the activity of the neural processes underlying visual discrimination. These oscillatory characteristics, it is proposed, in turn, reflect reciprocal information transfer or signaling between multiple cortical regions. This proposal is made on the basis of the following argument. Two ways in which oscillatory activity can arise is first, relaxation phenomena, and second, reciprocal signaling between subsystems with a positive value of gain. An example of a relaxation phenomenon leading to oscillatory activity is the sum-and-fire characteristic of a neuron. Reciprocal signaling in the simplest case can consist of feedforward and feedback paths connecting multiple subsystems. Such connections, together with the time delays inherent in any dynamical system, and the forward gain in the connected subsystems form a network that has the potential for exhibiting a rhythmic pattern of activity involving the connected systems. These three elements, gain, reciprocal connections, and delay, are all represented in neuronal systems. Signaling between cortical regions occurs along the profuse cortico-cortical and cortico-thalamic connections that have been described. Such information interchange, for example along arcuate fibers between cortical regions, along projection fibers between cortical regions and subcortical centers such as the thalamus, and along commissures between the left and right hemispheres, is the first requirement for the generation of rhythmic activity. The second requirement is similarly met in that time-delays are inherent in neural function, and are the result of synaptic and axonal transmission delays. Finally the gain function is likewise an inherent property of neurons. On the level of individual neurons, gain or amplification is represented in terms of the signal regeneration properties of axons, and the sum-and-fire behaviour typical of neurons.

A potential result of such reciprocal signaling, along association fibers with inherent delays, between the neuronal groups comprising disparate areas of the brain, can thus reasonably be expected to be an oscillatory pattern of activity. A direct reflection of this activity is the observed oscillatory nature of EEG signals found over all cortical areas. This oscillatory pattern may contain both periodic and aperiodic components. Stated more generally, the frequency of this oscillatory activity may change over time in a complex way. A neuronal basis for such a changing pattern of frequencies could be that the frequency that is observed at any time over any cortical region is the result of signaling within a network comprised of some particular number of components such as cortical microcolumns. As the size and configuration of this network change over time, so would the associated time-delays, and thus the frequency associated with this dynamic configuration would also change from moment to moment.

It might be suggested that the band-pass filtering used to preprocess the data prior to analysis would inevitably emphasize such periodic activity. While band-pass filtering can make such periodic activity more easily observable, filtering alters the characteristics of the activity relatively minimally when the passband of the filter is sufficiently wide, relative to the center frequency of the passband. On the other hand, methods such as analysis of EEG components that rely on averaging over many stimulus or response registered recordings, do impose a transformation on the original signals. Notably, the results of the averaging process present a distorted picture of both the magnitude and the temporal structure of the averaged signals. The point here is that in the present study the oscillatory activity that forms the basis of the subsequent correlational analysis was present in the original recorded EEG, and was not created by the techniques used to preprocess the signals. In contrast with features such as EEG components, the observed periodic activity should be a direct reflection of the behaviour of underlying neuronal systems, and as such should be a reasonable basis upon which to construct a description of the neurophysiological events associated with visual discrimination.

The discrimination of representations of complex real-world objects from a visually complex background is seen in the present study to involve the correlated activity between most, rather than between only a few, cortical regions. This correlated activity, it is suggested, is the result of interregional signaling involving bilateral occipital, temporal, central, parietal and frontal areas. The neuronal activity associated with object discrimination is found to be an oscillatory pattern, with frequencies predominantly in the theta band, and with a correlation between the different brain areas that changes in a distinct way over the 1 second interval preceding the moment of discrimination. First, mean correlation, averaged over all electrode pairs, increases. Second, spatial variability in the correlation, computed over all electrode pairs, decreases. The implications of these observations together is first, that over the course of the before-blink epoch, there is an increased level of coupling between neuronal regions, and second, that the magnitude of this increase in coupling is approximately proportional to the distance separating the neuronal regions. In sum, visual discrimination is characterized by coupling increases between most neuronal regions, but with a greater coupling increase between distantly spaced regions.

The neurophysiological processes associated with visual discrimination are thus associated with relatively little change in coupling between closely spaced regions. The level of coupling between such regions, indicated by a correlation of 0.75, is already substantial. At the start of the discrimination process, these regions, according to CSO model, carry out relatively independent and low-level analyses of stimulus features. According to the model, successful visual discrimination requires that in addition the coupling between distantly spaced sites also increases, as increasingly complex feature transformations are generated through associations

formed between correspondingly larger extents of cortex. Over the interval preceding discrimination, cross-correlations between such distantly spaced sites are seen to increase from approximately 0.2 to 0.4. If the magnitude of the correlation coefficient is assumed to reflect the degree of coupling between cortical regions, then mean coupling strength between these regions increases by on the order of 100%, between the beginning and end of the 1 second interval preceding discrimination.

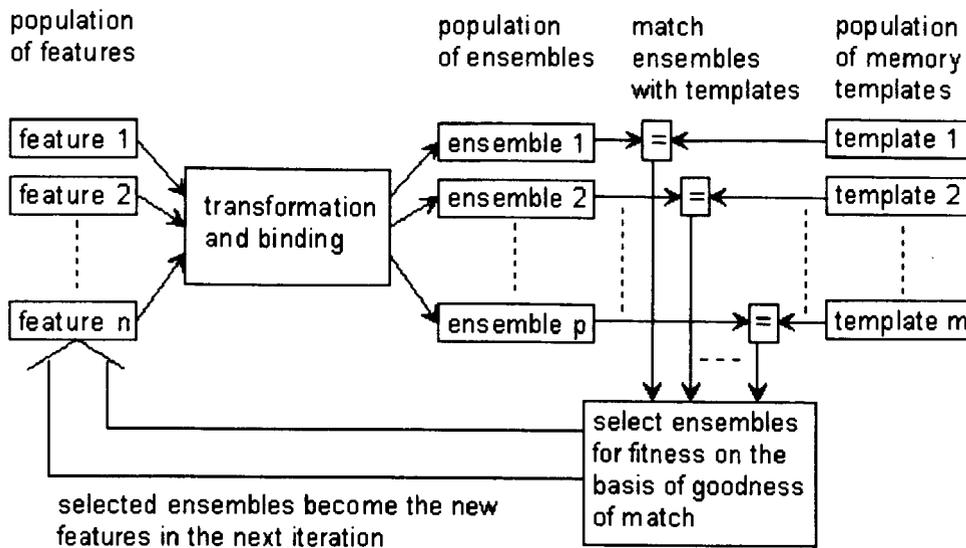
These observations can be interpreted as supporting the CSO model of object discrimination. That is, it is predominantly the increase in coupling strengths between distantly spaced cortical regions that mediates the process of visual discrimination. Correlations between closely spaced regions are indicators of the ongoing relatively local processing that transforms primary stimulus features and lower level ensembles into more complex feature ensembles. Such ensembles can be conceptualized as being in effect particular configurations of interregional signaling by means of which the central representation of the features of the visual image are transformed and bound. Visual discrimination is only able to occur when image features have been bound or transformed into a sufficiently complex, multidimensional transformation of the original retinal representation of the image, that then matches a pre-existing memory template. Such feature transformation or binding is in turn indicated by the increase in the magnitude of correlations, as discrimination approaches, between relatively widely separated cortical regions.

According to the CSO model then, a unitary percept, which in the present study is a discriminated target object, is the emergent product of a process of self-organization of the signaling configuration between a set of neuronal collectives. This self-organization, it is proposed, involves signaling, in successive iterations, between multiple and wide-spread cortical regions, that carries out a recursive sequence of transformations of the original retinal representation of the stimulus. Through these transformations the elementary features of this retinal representation are successively bound by being transformed into successively more complex feature ensembles, which as part of this transformation process are compared with successively more complex memory templates. The transformation process is in this way guided by the influence of existing memory templates. In essence, therefore, the processes of transformation and template matching are not separable.

To this description an additional feature is added, one that while not addressed by the present findings nevertheless allows the task of visual discrimination to be carried out with a computationally simple approach. Computational simplicity, it is proposed, translates into an economy of description given an inherent attribute of neuronal signaling pathways. This attribute is the high degree of interconnectivity both within and between neuronal regions, an

interconnectivity which in turn implies the possibility of a correspondingly high degree of functional parallelism in terms of intra and inter-cortical signaling.

In any one iteration this functional parallelism inherent in cortical signaling, it is suggested, can support a population of simultaneous feature ensemble-creating transformations. This in turn would result in the simultaneous creation of a corresponding population of feature ensembles, that is a population of coexisting, simultaneously active configurations of interregional signaling. Out of this population of available ensembles that may be generated in any given iteration, the ensembles that survive to become the features entering into the next round of transformation and memory matching are those that result in a sufficiently accurate match with existing templates. This description of these successive iterations might thus be conceptualized as representing an analog of Darwinian evolution, in which not organisms, but interregional signaling configurations are evolved. This evolution takes place within an environment in which fitness corresponds to a successful match between a feature ensemble and an existing memory template. An evolutionary model of this sort is an instance of the genetic algorithmic approach to the task of searching for global minima within complex problem spaces. Genetic algorithms have shown a competitive advantage in situations in which such exploration of a problem space is made difficult by the presence of multiple distracting local minima, and in which the possibility exists for multiple operations to occur in parallel (Goldberg, 1989; Holland, 1975). The CSO model is schematically diagrammed in Figure 16.1.



**Figure 16.1 The Cortical Self-Organization Model**

According to the model, it is proposed that in an iterated sequence, the brain chronically and automatically carries out a process of transformation and binding of the population of  $n$  features, by means of which a population of  $p$  feature ensembles are created. Feature ensembles are in effect specific configurations of interregional signaling through which sets of features are transformed and bound. By matching each of the  $p$  ensemble against the population of  $m$  existing memory templates, the  $p$  ensembles are rated for fitness. On the basis of this fitness rating a sub-population of the  $p$  ensembles is selected to act as the population of initial features in a subsequent iteration. This sub-population consists of ensembles that most closely matched existing memory templates. This process continues iteratively until a sufficiently accurate match occurs between a feature ensemble and a memory template. The winning ensemble represents the particular configuration of interregional signaling by which the central representation of visual image features is successfully matched with a memory template. The driving force for this process is proposed to be a mechanism involving energy relaxation. The more closely an ensemble matches a memory template the less energy is required to maintain the configuration of interregional signaling corresponding to that ensemble. The ultimately winning ensemble thus represents the particular inter-regional signaling configuration associated with a minimum energy state of the associated neuronal system.

While the term computation has been used for expository convenience in the preceding discussion it must be emphasized that the CSO model does not propose that the various transformational processes occurring during visual discrimination involve some sort of manipulation of symbols according to computational rules, a perspective associated with the Artificial Intelligence approach to the modeling of cognitive behaviour (e.g., Minsky, 1968). Rather, the CSO makes the assumption that all such transformations take place in an asymbolic, fully distributed fashion (McClelland and Rumelhart, 1988), making use of the substrate of dense anatomical pathways that exists, subserving cortical communication.

At the start of this transformation and matching process, the corresponding memory templates would conceivably consist of elements such as edges, orientations and colors. At some lowest level such elements may be hard-wired into the visual system (e.g., Hubel and

Wiesel, 1962). At some higher level, these templates would arise as a result of interactions between the individual and the environment in early development. As this process continues, such memory templates could represent more complex feature constellations including, for example, complex shapes such as the sensitivities to pattern partials hypothesized to exist in the inferotemporal cortex (Fujita et al., 1992; Perrett and Oram, 1993). This sequence of iterations would only terminate when a match occurs between the results of a stage of feature transformation and an existing memory template. At a particular level of iteration, that would be determined by high-level considerations such as intention, the purpose for which the particular discrimination is being conducted. Thus, if the intention is to discriminate a particular line orientation or elementary shape, the iterations might terminate after relatively few iterations. If, however, the purpose of the discrimination task is, as in the present study, to detect a complex depiction of a real-world object such as an animal or a bird, then a correspondingly greater number of iterations might be required. At the termination of the iterations, the corresponding network of interregional signaling may be thought of as having self-organized into a configuration that allows a successful match to occur between the highest level feature ensembles that correspond to the goal of the discrimination task, and existing memory templates. At this point, behaviourally, the target object has been successfully discriminated from its visual context.

The phenomenon of visual discrimination can be considered to be an instance of the more general process of induction, or theory formation. A rule or unified model is induced from initially apparently isolated data fragments. Such model construction underlies enterprises such as scientific theorizing; it underlies as well the ubiquitous phenomenon of camouflaged target discrimination. Both of these examples may, on different occasions, involve both conscious and non-conscious components. Thus, Kekule is reported to have discovered the structure of the benzene molecule in a flash of insight while boarding a bus. In a fundamentally similar way, a subject in the present study discovers the identity of a camouflaged target when the target appears to suddenly pop out after a period of visual inspection. While a pattern of evolution of interregional association through cortical self-organization is proposed to underlie the phenomenon of visual discrimination, such cortical self-organization should in fact serve as the ground of theory formation in general.

## 17 Cortical Self-Organization

### 17.1 Interregional Signaling Topologies

In this section the present findings will be discussed in relation to possible configurations of intercortical signaling underlying object discrimination, and how these configurations may evolve during perceptual integration. A primary finding of the present study is that immediately before the point of discrimination, multiple cortical regions, including bilateral occipital, temporal, central, frontal and prefrontal, showed evidence of mutually synchronous activity. At least two general configurations could serve as a basis for such associations. First, these associations might be attributable to the influences of a common source driving each of the several cortical regions. Second, the interregional associations might be interpreted as an increase in the level of mutual coupling among these regions.

The first possibility involves driving of the oscillatory activity within multiple cortical regions by a common source. This mechanism would require extensive connections between all cortical regions and a common neuronal hub, or pacemaker. Such connections might be provided by, for example, the projection fibers of the corona radiata that connect thalamic efferents with the cortex. Arguments have been made both in support of and in opposition to this view. On the one hand, as Bressler (1995) has suggested, a neuronal substrate for a pacemaker might include the large-scale pathways that have been found to project from non-specific nuclei within the thalamus to multiple and widespread cortical target areas (Goldman-Rakic, 1988, 1992). On the other hand, it has been pointed out that beyond the optic chiasm, there is a segregation of optic afferents to the two hemispheres, and that therefore there is no common input to the two hemispheres which could act as a driving source (Engel, Konig, Kreiter, Schillen and Singer, 1992; Llinas, Grace, and Yarom, 1991; Steriade, Curro Dossi, Pare and Oakson, 1991). It has also been proposed that multiple sources of activity synchronized with small relative phases must be driven by a common source, because of the effects of transmission delays (Ts'o, Gilbert, and Wiesel, 1986). However, the findings of Engel et al. (1991), in a study of inter-columnar coherence in the visual system of the cat, suggest otherwise. These findings showed that reciprocal signaling between cortical regions can result in synchronized activity among these regions with a zero relative phase delay, in the presence of transmission delays of up to several milliseconds. In support of these findings, simulations carried out by Konig and Schillen (1991) have shown that synchronized activity with a zero phase delay can be achieved if conduction delays are not greater than one-third of the period of the oscillatory activity. Results such as these, suggesting that a small relative phase angle does not necessarily imply a pacemaker driving source, are consistent with the second possible mechanism that could

account for the increased level of coupling among multiple cortical regions, a non-hierarchical network of mutual interconnections.

This second possibility involves reciprocal cortico-cortical (Engel, Konig, Kreiter and Singer, 1991) and cortico-thalamic (Edelman, 1989) signaling. The present findings are consistent with the view that at the start of the 1 second epoch preceding discrimination the frequencies of the activity at the various cortical regions may be relatively dissimilar, as each region operates in relative autonomy. The result of such autonomy would be the observed low values of interregional association. As the moment of discrimination approaches, an increasing level of interregional signaling would result in multiple cortical regions becoming increasingly coupled. This, in turn, would result in mutual entrainment of the activity at these regions, which would then be reflected in the observed increasing level of interregional association.

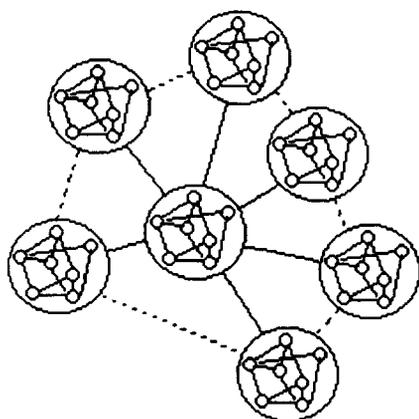
These two contrasting positions can be very approximately modeled in terms of multiple functional units interconnected by the two topologies. In the first case, this topology consists of relatively weak coupling strengths between all units, with the exception of relatively strong level of coupling between these units and a central or hub unit acting as a common driving source. This configuration will be referred to as the hub topology, and is diagrammed in Figure 17.1a. In the second case, the topology consists of inter-unit couplings with relatively equal strengths, and with no preferentially strong couplings to a common unit that could function as a pacemaker. This arrangement will be referred to as the distributed topology, diagrammed in Figure 17.1b. In both cases units have the capability to oscillate in any one of a number of distinct modes, with the frequency and phase associated with each mode being a function of the signals received from other units through the inter-unit couplings.

In both of these prototypical topologies, each of these units can be conceptualized as itself consisting of a network of processing units. The result is a recursive structure, with the network configuration repeating at different spatial scales. In particular, each unit might be considered to consist of a recurrently connected network, a configuration that constitutes a dynamical system with a potential for a rich set of behaviours (e.g., Ermentrout, 1994). These behaviours, and their corresponding encapsulation in terms of phase space attractors, can range, depending on inter-unit coupling parameters, from static states, corresponding to point attractors, through oscillatory behaviour with various combinations of frequencies, corresponding to limit cycles, to chaotic behaviour describable by a phase space attractor with a fractal dimension and with a positive Lyapunov exponent, commonly termed a strange attractor (e.g., Moon, 1987, p. 23). These potential classes of behaviour, characteristic of each of the coupled units, can in turn give rise to a corresponding range of dynamical behaviours in the network comprised of the interconnected individual units. Experimental observations of simultaneous coherences at multiple frequencies during cognitive tasks is suggested by Bressler et al. (1993) as providing

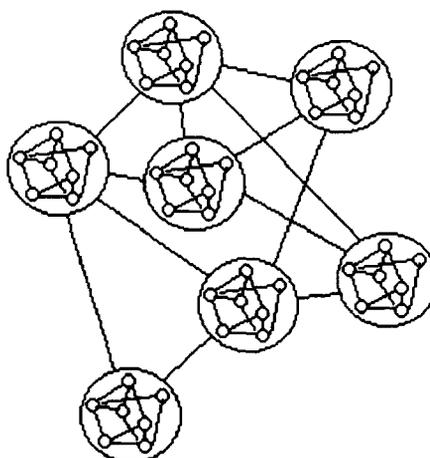
cortical information processing with a degree of flexibility in perceptual tasks not available from oscillatory behaviour over a more limited range of time-scales. On purely theoretical grounds, Churchland (1995) has proposed that such recurrent neural networks, with the ability to operate on information in a recursive manner, inherently possess the behavioural flexibility to account for such phenomena as figure-ground discrimination, and ambiguous-figure resolution.

In the proposed model topologies, each of the individual units has, in addition to the inputs from other units or from the pacemaker, a perturbing noise input. This noise input would have the capability to modulate the frequency of the oscillatory activity of individual units. Figure 17.2 shows a simplified schematic diagram of some of the connections that comprise the visual system. The double-headed arrows in the figure indicate reciprocal connections. A more complete functional connection diagram of the visual system is presented by Van Essen and DeYoe (1995). In their more complete diagram of functional interconnections within the visual system, and as is suggested by Figure 17.2, the topology of the visual system would appear to be well modeled as a distributed rather than as a hub configuration. On the other hand, some combination of these two topologies is also a possibility. A diagram such as that presented by Van Essen and DeYoe (1995) is intended to model a particular subset of the functions in visual perception, those involved in the processing of visual features. A wider perspective on the issue of visual perception might identify as well functions not directly concerned with such processing, but concerned with global modulation of neural states, for example by processes associated with attention and arousal. Such global functions might be well modeled by a configuration resembling a hub rather than a distributed topology. In any event, each of these two possible topologies can be associated with particular predictions. These predictions will be discussed in the following section.

(a) hub topology

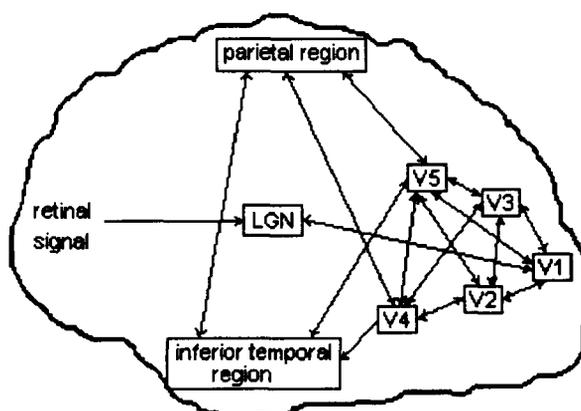


(b) distributed topology



**Figure 17.1 Prototypical inter-regional signaling topologies**

These network configurations consist of discrete nodes interconnected by strong (solid lines) or weak (dashed lines) links. Each node itself consists of a recurrent network structure. The hub topology is characterized by relatively strong connections between all processing units and a single central unit, with relatively weaker connections between the processing units themselves. The distributed topology is characterized by relatively equal strength connections between all processing units.



**Figure 17.2 Reciprocal interconnections in the visual system**

This schematic diagram shows a few of the major connections that have been identified between cortical and subcortical regions involved in visual perception. Double-headed arrows indicate reciprocal connections. (Adapted from Edelman, 1989).

## 17.2 Predictions From the Two Topologies

The first prediction concerns the spatial variance, that is the instantaneous variance computed across multiple network nodes, in the relative phases of the oscillatory activity at the processing nodes. In a distributed topology perturbed by noise, one mode of operation should be

synchronous activity with some non-zero value of spatial phase variance. In contrast, a hub topology similarly perturbed by noise should show a smaller relative value of spatial phase variance. This suggestion is made on the basis of the argument that in the distributed topology, inter-node coupling would result in interactions at each node between signals arriving from other nodes and the perturbing noise inputs. These interactions could then propagate throughout the network along the inter-node links, resulting in higher level interactions. The result would be a relatively large value of spatial phase variance. In the case of the hub topology, noise perturbations would again interact with the signal from the hub at individual nodes. However, since with this topology inter-node coupling strength is low, these interactions would not propagate as readily, resulting in less higher-level interaction. The outcome would be a relatively smaller spatial phase variance. Thus, the higher level interactions within the distributed topology would result in a relatively more complex and nonlinear system behaviour, with a correspondingly greater spatial and temporal phase variance, in contrast with the hub topology. With the hub topology, the lower level of interaction would result in the relatively less complex and relatively linear behaviour, with a correspondingly smaller phase variance. An analysis of spatial phase variance would thus address the issue of the possible configuration of interregional signaling. A low value would be more consistent with the view that intercortical signaling can be modeled by a hub topology, with multiple cortical sites driven by a common pacemaker. A large value would be more consistent with a distributed topology, with synchronization of multiple sites mediated by their mutual interconnections. In order for it to be useful in discriminating between the two topologies however, the value of spatial phase variance would need to be calibrated using measurements on known patterns of intercortical connections. Such measurements would also probably be best carried out using single cell recordings from multiple sites, rather than using the more spatially extended and diffuse measurements from the scalp. For these reasons, this prediction regarding phase variance can not be applied to the present results.

The second prediction relates to the relative number of oscillatory modes associated with the two topologies. The hub topology would appear to have a single dominant mode of oscillation, in which all of the driven units oscillate in synchrony. In contrast, the distributed topology should be capable of exhibiting a wide range of oscillatory modes, with the particular mode being determined by the instantaneous distribution of inter-unit coupling strengths. A network of  $N$  recurrent networks, each with the potential for  $m$  dynamic modes of behaviour would itself have the potential, in the absence of additional constraints related to for example degree of connectivity and coupling strengths, for exhibiting  $N^m$  states or modes of behaviour (MacGregor, 1993; p. 303). Of these, only a subset of modes would be characterized by more-or-less synchronous activity among all units, similar to the dominant mode of the hub topology.

Many other possible modes however would be characterized by relatively more complex behaviours, and corresponding more complex patterns of signaling between the coupled nodes. A robot simulation of insect gait patterns has shown that such a system of coupled processing units will self-organize (e.g., Brooks, 1991; Chiel and Beer, 1993; Fetz, 1993). Self organization can be defined in the present context as follows. The system of coupled processing units will settle into stable and organized patterns of behaviour that are determined by interactions with the environment, that is, external inputs to the network, as well as by constraints directing mutual interactions between the coupled units, but importantly, without the necessity for a central controlling signal to direct this organization. Simulations of gait patterns such as those cited above find that with such a distributed signaling topology multiple operating modes are possible, with a particular mode selected by the nature of the interaction between the network and the environment. In summary, a hub topology should be characterized by a single dominant mode with all coupled units operating in synchrony, while a distributed topology should be characterized by multiple possible operating modes, only a subset of which would correspond to synchronous activity among the coupled units. This admittedly simplistic line of reasoning nevertheless suggests that the observation of synchronous activity among a network of processing units, which in the present study corresponds to the observed high degree of correlation among multiple cortical regions prior to discrimination, implies a greater likelihood that the processing units are interconnected by a unimodal hub topology rather than a multimodal distributed topology. This conclusion may be drawn only in the absence of additional constraints.

As an example of such a constraint, an additional factor can be brought into this analysis of interregional signaling configurations, the notion of the energy within a coupled system. For the distributed topology, the various operating modes can each be associated with a particular level of energy. In the absence of additional requirements, this energy level should be a minimum for the totally synchronous mode. In such a case, this synchronous mode, with all coupled units operating in synchrony, would be the preferred state of the system. The system would tend to relax into this minimum energy configuration. An analogy that can be drawn from physics is the case of Rayleigh-Benard convection within a body of fluid receiving energy at the bottom and dissipating it at the top (e.g., Baker and Gollub, 1990, p. 133). A simplified model of such convection dynamics was studied by Lorenz (1963) who developed a system of three coupled ordinary differential equations:

$$x' = \sigma (y - x)$$

$$y' = Rx - y - xz$$

$$z' = xy - Bz$$

This highly simplified approximation has nevertheless been found to be capable of displaying a wide range of behaviours, from fixed points to limit cycles to deterministic chaos, depending on the values assigned to the system parameters  $\sigma$ ,  $R$  and  $B$ . In these equations the prime denotes a differential with respect to time,  $x$  corresponds to the intensity of the convective movement and  $y$  corresponds to the temperature difference between top and bottom of the mass of fluid. At low levels of energy input, heat transfer within the body of fluid occurs primarily by conduction. As the rate of energy input increases, a transition in the nature of the energy transfer mechanisms within the fluid occurs. Energy transfer eventually comes to involve, in addition to conduction, convection in the form of multiple local circulating systems of fluid, that is, convection columns. This transition, from an essentially stochastic process involving the relatively random, unorganized, motion of individual molecules, to an organized process involving a system of convection columns each consisting of large numbers of molecules moving in relative synchrony, is an example of the process of self-organization. At increasing levels of energy input, further transitions take place within this system. In these further transitions, not only does the number of convection columns increase, but increasing numbers of columns of different sizes occur simultaneously. At a sufficiently high level of energy input, the motion of the fluid becomes turbulent, or in alternate terms, chaotic. Even at high levels of energy input however, the phenomenon of intermittency can interrupt the chaotic regime. Limit cycles, that is intervals of predictable oscillatory behaviour, can occur, punctuated by erratic chaotic bursts (e.g., Peitgen, Jurgens and Saupe, 1992; p. 253). The driving force for the initial self-organization of the fluid into columns, and for the subsequent transitions, is minimization of the quantity of energy stored within the fluid. The amount of stored energy can be related to the temperature difference between the top and bottom of the body of fluid. By self-organizing, the dynamics of the body of fluid are able to effect the transfer of a relatively larger amount of energy from bottom to top, and thus minimize the amount of energy stored within the fluid.

It is not suggested that cortical processes resemble fluid convection dynamics on any but the most general level. On this general level however, it is proposed that the multiple possible interregional signaling modes of a network of cortical systems coupled by a distributed signaling topology may, in analogy with a dynamical system such as that which describes fluid convection, similarly evolve into a minimum configuration along a dimension equivalent to energy. With the constraint that the level of input to the neuronal system of this equivalent variable is not unlimited, the minimal system configuration should correspond to the fully synchronous mode. The principal characteristic of such a mode is synchronized activity among the coupled cortical regions. Furthermore, and again in analogy with convection dynamics, at increasing levels of equivalent input, modes more complex than the fully synchronous mode might occur.

Some care must be taken in making the analogy between fluid convection dynamics and neuronal interregional signaling dynamics. The concept of energy minimization in the fluid convection case translates, it is suggested, into an equivalent variable in the neuronal signaling case, a variable that involves information. Thus, while the fluid is seen to self-organize in order to minimize the amount of stored energy, the neuronal signaling configuration can be conceptualized as self-organizing in order to minimize the amount of stored information. That is, the reason as it were for the interregional signaling configuration to self-organize would be to minimize the amount of information that the associated neural system would need to deal with, essentially the amount of information that it would need to contain, by means of an appropriate organization of that information.

A final consideration in comparing the two possible interregional signaling configurations is that, in the absence of additional constraints, of the two topologies the fully distributed configuration would appear to be the more economical description. In the fully distributed topology all nodes and links between nodes are hierarchically equal. In contrast, the hub topology presumes a two-level hierarchy, on the one hand the hub and its connections to all other nodes, and on the other hand the nodes and their interconnecting links. Thus, while the present findings appear to be generally consistent with both topologies, parsimony of description would suggest choosing the fully distributed model as the neuronal signaling configuration responsible for the present results. Constraints or requirements beyond those of absolute simplicity of description reasonably do exist: The complex of perceptual and cognitive functions involved in propelling an organism over the course of its span of existence is most reasonably supported by whatever neural signaling structures represent the most effective trade-off between reliability and efficiency. Such structures may involve the characteristics of a distributed topology for functions such as memory that are efficiently and robustly implemented by means of the sharing of information among multiple processing systems. In complement, such structures may involve the characteristics of a hub topology for functions related to global modulatory functions such as those related to attention and arousal. Both of these classes of functions should be involved in perceptual-cognitive operations in general, and in the visual discrimination task of the present study in particular.

Summarizing this discussion of how the present findings relate to the topography of the underlying intercortical signaling, the observation of synchronous activity among widespread cortical regions at recognition is consistent with both a hub and a distributed topology. The hub topology is preferred on the basis of the greater likelihood of fully synchronous operation with this configuration. This advantage is mitigated however by consideration of a relaxation mechanism within a distributed topology. With the constraint that the information input to the neuronal systems is not unlimited, this relaxation mechanism should favor the fully synchronous mode for

the distributed topology. The distributed topology would also appear to be preferred on the grounds that it represents a simpler description of the underlying configuration of interregional signaling. Finally, a more realistic assessment of the diversity of functions required by tasks such as discrimination of camouflaged targets suggests that both topologies are reasonably involved.

### **17.3 Perception as a Relaxation Phenomenon**

The idea that object perception can be conceptualized as a relaxation process, proposed by Tom Richardson (in conversation, 1995), suggests that the underlying neuronal processes can be considered as a dynamical system evolving towards a state corresponding to a minimum along some dimension. Dynamical systems can typically be considered as relaxing or evolving into a state of minimum energy. A non-trivial example of a dynamical system with a point attractor is a soap film. In conforming to the constraints imposed by a supporting wire-frame structure, the soap film adopts the surface configuration which corresponds to a minimum in the level of energy stored within the film, in terms of the forces associated primarily with surface tension. In an analogous way, the neuronal systems involved in object perception, constrained on the one hand by the information within the central representation of the visual stimulus, and on the other hand by the information stored within existing memory templates, might in a similar way be considered to relax or evolve into a state that corresponds to a minimum along some dimension. As stated earlier, this dimension need not be energy in the case of a neuronal system, but rather could be an equivalent variable such as information. Borrowing from statistical physics, the process of evolving towards an information minimum is equivalent to evolving towards a corresponding entropy minimum. By a definition of entropy, a decreasing value of entropy for a system implies that a decreasing amount of information is needed to describe the state of the system. This should in fact be the case in the present example of object discrimination. By way of a simple analogy, a network of oscillators can be described in terms of fewer bits of information when the coupling between the oscillators is strong enough to entrain the frequencies of the individual oscillators, than when the coupling strength is low and the oscillators operate with relatively independent frequencies and phases. In a similar way, the state of a system comprised of multiple neuronal regions operating in synchrony should in principle be describable in terms of less information, that is fewer frequencies and phases, than the state of a system consisting of multiple neuronal populations operating relatively autonomously. A description of the latter system would involve a relatively greater number of frequencies and phases.

Features of two of the views of visual perception that were discussed earlier, the RCI model and the convergence zone framework, can be encompassed by the proposal that perception involves a relaxation process, in which multiple cortical systems self-organize by relaxing into a state corresponding to a minimum along a dimension involving information. According to the RCI model, the process of perceptual categorization involves reentrant signaling between neuronal populations within a neural system that creates sensory-cortical, limbocortical and corticocortical mappings at multiple scales. Such mappings may be viewed as neuronal configurations that implement transformational functions between sensory, cortical, and subcortical systems. The creation and evolution of these maps might be restated as the organization of discrete neuronal populations into functional networks. Importantly, this organization is directed, not by some controlling device external to the neural system, but through interaction between the system and the environment, and by the degree to which the configuration of these mappings is successful in adapting the individual to the environment. This is a description of a process of self-organization in which, as larger networks are created out of more local structures, the information within the overall system might be seen to tend towards a minimum. In other words, the more highly organized the overall system comprised of these neuronal groups becomes, the fewer bits of information are needed to describe the state of the system.

In a generally similar way, the convergence zone framework suggests that discrete and widely separated neuronal populations organize through the mechanism of feedforward and feedback connections that link such populations with local control centers, the convergence zones. Again, this organization occurs, not as the result of an imposed directive from a higher level source, but as the result of a property inherent in the neuronal system itself: During perception, convergence zones encode a pattern of interconnections among multiple cortical regions, a pattern that captures the configuration associated with the perception. These zones then, in a sense, play back that code to reestablish the interconnection pattern during memory recall and recognition. The information needed to specify the state of this self-organized system is less, it is suggested, than the information needed to describe the system before it has self-organized.

In general, statements about the neuronal basis of perception, such as the RCI model and the convergence zone framework, can be conceptualized, it is proposed, as being different perspectives on a common model, exemplified by the presently proposed cortical self-organization model. According to the CSO model, a coherent perception is the emergent result of a process of self-organization of a system comprised of multiple discrete neuronal populations, driven by a natural tendency, a relaxation process, which involves the minimization of a dimension such as information.

To summarize the discussion thus far, it is suggested that object perception can be conceptualized as a relaxation process in which the pattern of interregional signaling between the components of the associated neuronal system relaxes towards a state of minimum information, and correspondingly, minimum entropy. The decreasing entropy level of this neuronal system corresponds to an increase in the level of organization of the associated components. While the analogy of fluid convection dynamics involves self-organization driven by energy relaxation, the case of neuronal interregional signaling dynamics involves a corresponding self-organization through information relaxation, a relaxation into an information minimum.

It may be, however, that in the case of neuronal signaling, the two variables, energy and information, can be considered to be related. Thus, the increasing degree of organization of the pattern of interregional signaling may correspond to a decreasing level of energy contained within the associated neuronal system. A possible mechanism for this correspondence can be suggested. Since a finite amount of energy is required to carry out an element of interregional signaling, a more organized and hence generally simpler signaling configuration should require less energy to sustain than a less organized and more complex pattern of signaling.

Object discrimination might therefore be viewed as a tendency towards self-organization of multiple cortical systems, driven by the requirement of reducing the total energy within the system. Through this process of self-organization, the relatively high information content associated with the complex original retinal representation of the visual image is transformed into an integrated, discriminated, percept associated with corresponding information and energy minima.

These statements form part of the CSO model of object discrimination. According to the model, object discrimination is subserved by a sequence of operations that occurs chronically and automatically within the neuronal systems associated with perceptual and cognitive processing, and which operate on primary sensory input initially, and in a recursive manner on the products of these operations themselves subsequently. Thus, it is proposed, visual discrimination entails an iterated process in which a sequence of transformations of the central representation of a visual image are used to successively approximate of the results of prior learning. This iterated process, creating a series of increasingly complex transformations, or feature ensembles, of the image elements, continues until a match occurs between the feature ensembles and information contained within existing memory templates. The sequence of these successive iterations can be considered as a process of self-organization occurring among multiple neuronal populations, driven by the requirement that the total amount of energy and information contained within this system is to be minimized. The result of this self-organization is an emergent unified perception, which in the present case is represented by the target object,

successfully discriminated from its background. In equivalent information-theoretic terms, the associated neuronal system can be conceptualized as relaxing towards a state of minimum entropy.

The view presented here of the network of neuronal functions associated with object discrimination is that of a dynamical system within which the pattern of interregional communication, driven by a mechanism of energy relaxation, is able to self-organize in order to coordinate the processing resources within multiple and widespread cortical regions. A consequence of this view of the process of cortical integration in terms of energy minimization is that it endows the process with a teleological component. A central coordinating device or structure is thus not required in order to direct the flow of events during the process of cortical integration in the direction of generating a unified percept. The principle of energy relaxation provides a natural driving force, and thus a direction, for the sequence of processes involved in generating such an integrated percept.

## **18 Relating the Present Findings to Alternative Models**

In this section, a key feature of the results of the present study, the spatial extent of synchronization, will now be discussed in relation to two models of visual perception. The first model is the reentrant cortical integration (RCI) model (Finkel and Edelman, 1989). The RCI model is an example of a model involving relatively distributed pattern of communication. The second model is Damasio and Damasio's (1993) convergence zone framework. This framework, in contrast with the RCI, is an example of a model that proposes a hierarchical system of interregional signaling.

The present study found that immediately prior to the moment of discrimination, multiple cortical areas including bilateral occipital, temporal, central and frontal regions showed evidence of mutually coherent activity. As discussed earlier, this finding would appear to be more appropriately modeled by a distributed topology, and is thus more consistent with the reentrant cortical integration model than with the convergence zone framework.

The convergence zone framework posits the existence of controlling neuronal collectives, the convergence zone, that function to organize multiple and widespread cortical regions into a network of nodes capable of operating in mutual synchrony. Such convergence zones in turn require the existence of feedforward and feedback paths between any one zone and multiple other cortical areas. Pathways that have been shown to exist, connecting thalamic (Goldman-Rakic, 1988, 1992) and non-thalamic (Rolls, 1989) sources with multiple cortical areas, are generally limited in scope. Such pathways do not generally project from a single source to the wide range of cortical regions that are observed in the present study to be involved in coherent activity. Wider-scale projection systems have been found however. These include the identification of corticothalamic ascending and descending pathways between the intralaminar nucleus of the thalamus and all areas of the cortex (Llinas and Ribary, 1993). The convergence zone framework suggests that a large number of controlling regions may exist on many scales, each of which coordinates the activity of a collection of subordinate cortical areas. The activity of such individual controlling regions can, in turn, become synchronized by means of still higher order convergence zones. Such higher level zones would thus coordinate signaling among lower level zones.

In contrast with the convergence zone framework, the reentrant cortical integration model does not require central controlling structures to direct the organization of multiple cortical regions into a topology capable of mutually coherent activity. According to the RCI model, coherent activity, and the consequent binding of stimulus features, is a result of the complex reciprocal signaling among these regions, termed reentry. In contrast with the convergence zone framework, the RCI model suggests that these multiple regions are interconnected in a non-

hierarchical configuration. The RCI model, with an underlying distributed topology would, at least in this limited respect, appear to be a more economical explanation consistent with the present results.

In summary, it is suggested that the present findings, primarily in terms of the observed wide extent of coherently-coupled cortical regions, are to some extent more economically explained in terms of a description such as the reentry-based RCI model, than in terms of a view such as the convergence zone framework. It is suggested further that the interregional coherences observed in the present study are more consistent with a model involving interregional information interchange, similar to the reentry mechanism proposed by Edelman (1989), rather than with a model involving a common-source driving by central controlling structures. It must be acknowledged however that the present results do not appear to be useful in making a discrimination between these two positions on grounds other than economy of description. Thus the present findings, particularly in terms of the topographic distribution of interregional associations, do not provide the level of spatial resolution that would be required to address the question of the existence of convergence zones. This question might however be addressable using techniques such as MEG recording with dense sensor arrays.

These conclusions, regarding the possible signaling topologies consistent with the present results, must be qualified by an important limitation that is to some degree inherent in the measurement paradigm used in the present study. By making measurements of scalp potentials there is a tendency to associate such potentials with generating structures located within immediately adjacent cortical regions. The validity of such an assumption rests in part on the physical proximity of such cortical sources to the scalp electrodes themselves. This proximity argues for the view that the effect of cortical sources will be represented in the scalp electrical activity preferentially with respect to the effects of subcortical sources. Reasonably, the effects of such subcortical sources must nevertheless represent some component of the potentials measured at the scalp, either relatively directly through volume conduction from subcortical regions to the scalp, or indirectly as a modulatory influence on cortical activity. For example, the observation in the present study that associations were maximal between oscillatory components in the theta frequency range hints at the possibility that interactions between cortical regions and the hippocampus may be involved (Miller, 1991; Basar, Schurmann, Basar-Eroglu and Demiralp, 1994). Miller (1991) for example suggests that theta activity may be the result of a corticohippocampal resonance, pointing out that the total transmission delay within a loop involving cortex and hippocampus is of the right order to result in a theta band periodicity. The function of this theta activity, he proposes, is to modulate the level of activation of networks of cortical cells in aid of feature binding and memory retrieval processes. In support of this notion, local negative potential excursions in the upper cortical

layers have been associated with a lowering of cortical activation thresholds (Birbaumer, Lutzenberger, Elbert and Trevorrow, 1994), while EEG positivities have been suggested to reflect increases in activation thresholds (Mitzdorf, 1985).

In order to try to estimate the effect of such non-cortical generating structures on the basis of scalp potentials, essentially inferential techniques have been developed that generally make use of iterative optimization algorithms. Such algorithms attempt to infer the location, orientation and strength of one or more subcortical generators on the basis of an observed pattern of scalp potentials (e.g., Kertesz, 1994). These valuable techniques nevertheless suffer from the effects of the inverse problem: there is no unique solution in terms of the locations of subcortical generators corresponding to any observed topography of scalp potentials. Typically, therefore, source localization procedures make use of biologically-driven constraints, in order to try to develop solutions which correspond to subcortical structures that have been hypothesized to exist on the basis of independent theoretical or empirical work. In any event, in the present study such source localization procedures have not been employed. For this reason any statements that are made here regarding inter-cortical connection topologies consistent with the present findings must acknowledge the fact that the effects of subcortical sources has not been estimated. The use of subcortical source localization procedures represents a possible direction for future extensions to the present work.

## 19 Perception and Awareness

In general terms, the present findings show that visual object discrimination is associated with a transient wave of synchronization that sweeps out from the primary visual areas, and which eventually includes virtually all cortical regions, occipital, temporal, central and frontal. It is this synchronization transient, representing a momentary increase in the degree of coupling of these diverse cortical regions, that defines the discrimination event, as reflected in the subjective state of conscious awareness of the target object as an entity distinct from its background.

Thus, on the basis of the present findings it might be suggested that it is only when sufficiently large and numerous areas of cortex are participating in synchronous activity that the phenomenon of conscious - that is, reportable - awareness of the discrimination of an object occurs. According to the CSO model, however, preceding such discrimination of a complex object there should be a sequence of precursor events involving over time a progressively more complex and multidimensional bundle of information about the central representation of the stimulus. Although not specifically tested in this study, introspection suggests that, beyond the very general awareness of elementary visual forms such as lines, colors and orientations, the neural events occurring during the discrimination process do not generally give rise to any conscious or reportable awareness of such intermediate feature ensembles. The effect does, however, vary with the stimulus. For some images, such a pop-out effect occurs, while for others, the discrimination process is somewhat more continuous.

One interpretation of this observation is that, for the stimuli for which the pop-out effect occurs, intermediate processing products do not exist, but rather than the discrimination process involves initial elementary visual feature analysis followed by a massively-parallel recursive memory-matching or search process. In such a case, there would be no feature-transformation products of intermediate complexity of which the subject could be aware. Rather, only when the memory search had succeeded in generating a successful match would conscious, reportable, awareness of the target occur. This suggestion is supported by the everyday observation that how, or even if, an object is perceived is largely determined by what one is prepared to see.

These statements represent a possible modification of the CSO model. According to the model, object discrimination involves a relatively continuous process of successive iterations of a process involving feature analysis, transformation, and memory matching, creating successively more complex feature ensembles. According to the present discussion, object discrimination, in at least some cases, involves a somewhat discontinuous process of successive iterations of initial feature identification and relatively low-level binding of these features into a population of feature ensembles, followed by a massively-parallel memory search. An unsuccessful search would be followed by a rebinding of primary visual features into a new

population of feature ensemble along new feature dimensions, with a subsequent memory search for a match to these new ensemble. In the present results there appears to be no evidence upon which to discriminate between one or the other of these hypotheses, other than the general observation that, for at least a subset of the stimuli, reportable awareness appears to be associated only with successful discrimination.

## 20 Extensions

A number of extensions of the present study suggest themselves. One extension that involves a modification of the present methodology would be to compute an on-line measure of association. This measure would indicate the relative degree of intercortical coupling. In the present paradigm, associations were measured both before and after the eye-blink. In a follow-up study, aimed for example at refining the discrimination signal itself, a measure of association could be computed for signals recorded only prior to discrimination. In this approach, an increase in association, detected using an appropriate algorithm, would then be used to turn the computer display off. The subject would immediately be asked whether discrimination had taken place, and the relative timing of the discrimination event and the offset of the display. This approach would have the advantage of not requiring some action on the part of the subject, such as the eye-blink, to signal discrimination. This approach would have the associated benefit of eliminating the effect of the observed timing uncertainty between the increase in association and the eye-blink.

A second extension would be to replace the on-line computation of a measure of association with a neural network associator. The neural network could be trained to associate characteristics of the EEG signals with the object discrimination event. This approach would have the advantage that a neural network can operate as a universal function approximator, making optimal use of EEG signal features. A neural network would not be biased by any particular model of what the nature of the relationship between EEG signal characteristics and the object discrimination event should be.

A third extension, discussed earlier, would involve the use of source localization techniques in order to try to estimate the effects of subcortical, such as for example, thalamic, sources. Ideally, such localization techniques would use information present over a window of time, such as the 0.25 second time-windows defined in the present work, rather than information from a single time-point. An alternative to localization techniques based on scalp measurements would be functional imaging techniques, such as for example magnetic resonance imaging or regional cerebral blood flow measurements within the present target discrimination paradigm.

A fourth extension would be to repeat the present study using measurements of magnetic rather than electric fields over the scalp. Such MEG measurements have as one advantage an increased level of spatial resolution: In comparison with electric fields, magnetic fields interact minimally with the tissues that intervene between cortical and subcortical generators, and the scalp. The result is a smaller degree of blurring using magnetic rather than electric fields, in that such fields more accurately represent the topography of the underlying generating structures, than do scalp potentials.

## 21 Applications

A number of areas of application of the results of this study suggest themselves. One area in which these results could be applied is in the design of human-machine interfaces. This area includes any application that requires a means of communication between a human being and a computerized system. Such a system might, for example, perform environmental monitoring, or facilitate interpersonal communication. The data analysis methodology used in the present study involved aligning the EEG time-segments on the blink by which subjects signaled the object discrimination event. A more general, and perhaps more powerful procedure would be to continuously, that is, on an on-line basis, process the EEG signals to compute some measure of association such as intercorrelation, and then to look for increases in association between a subset of the 1020 electrode ensemble. According to the present results, such increases, between bilateral occipital, and fronto-temporal sites, should effectively index the occurrence of an object discrimination event. An alternative, as discussed earlier, would be to use a neural network associator to indicate to moment of discrimination. The overall result would be that a person could signal discrimination of a target object from a camouflaging or distracting background in a non-verbal, non-motoric way, through the sharp increase in magnitude of intercorrelation. Applications for this effect would include tasks in which individuals are involved in scanning visually-complex or dynamic scenes, such as computer-generated displays, in search of particular objects or groupings of objects while ignoring a non-essential background. It should be possible to design such human-machine interfaces in such a way that an individual could react to an object or constellation of object features on the display without producing an actual motor response. In this way it would be possible to eliminate the reaction times associated with generating such a response.

Another general areas of applications might involve providing a channel of communication for persons with motor disabilities that make it difficult for them to generate responses to events in their environment based on muscle action. With appropriate training, it may be possible for persons to generate a discrimination response to a target object while ignoring potentially distracting non-target objects. In this respect an extension of the present paradigm would be to look for differences in the evolution of the pattern of interregional associations in relation to differences in the visual features of the discriminated target objects. The question to be answered would therefore be, are there differences in the pattern of interregional associations, or in the time-evolution of this pattern, as a function of target features such as shape or color? An experiment designed to answer this question would represent yet one more possible extension to the present study.

The present findings also suggest that the paradigm used in the present study may have applications within neuro-cognitive studies generally, as a means of generating a response corresponding to item discrimination. In the present study subjects were instructed to blink following onset of the conscious awareness of the identity of the target objects. It may be possible however that a similar large-scale coherence response might be found to be associated with only implicit, rather than explicit tests of recognition of a stimulus (for a review see Schacter, Chiu and Ochsner, 1993).

## **Appendices**

## **Appendix 1 Effect of Noise on Correlation**

### **A1.1 Introduction and Method**

The issue has been raised in the literature that the observed magnitude of correlations are affected by the signal to noise ratio of the data. For example, Pijn, Vijn, da Silva, Van Emde Boas, and Blanes (1989) point out that the strength of a correlation between 2 sets of data is artificially decreased by the presence of random noise. To estimate the effect of random noise on the strength of correlation a numerical experiment was carried out using artificial data, to find the relationship between correlation and signal-to-noise ratio. Next, in order to measure the amount of noise in the experimental data that was due to EEG amplifier and to environmental sources, a dummy-input circuit to the EEG amplifiers used in the present study was constructed. This artificial input was intended to approximately model the subject for the purpose of assessing the amount of noise that was recorded as part of the experimental data. Using this dummy input, the output at all 16 channels was recorded. This recording was done immediately after data from subject 8 had been recorded. The RMS amplitude of the actual experimental data (excluding eye-blink waveforms) was measured, and the corresponding values of signal to noise ratio were computed for each session.

The artificial data for the numerical experiment consisted of 10 pairs of vectors with each pair of vectors having a different signal-to-noise ratio. Each vector consisted of 32 data points, and consisted of one cycle of a sine wave with an RMS amplitude that was constant across all vectors, together with additive gaussian noise with an RMS amplitude that was adjusted for each pair of vectors. In this way a different signal-to-noise level was created for each pair. This one cycle sine-wave in the 32 data points is equivalent, in the object discrimination data, to a frequency of 4 Hz. This frequency is the geometric mean of the frequency band used in the analysis of the object discrimination data, 2 to 8 Hz. Signal-to-noise ratios ranged between 0.25 and 10. Correlations were then computed between each pair of vectors.

In order to measure the noise level of the EEG amplifiers used in the present study, the dummy input was connected to the EEG amplifier inputs, in place of the scalp electrodes. The dummy input was a set of 19, 4.7 kilo-ohm resistors. One resistor, modeling the electrode to scalp resistance, was connected to each input of the head-box that was used in the experiment: 16 resistors to the 16 channel inputs Fp1 through O2, 2 resistors to the reference inputs, and 1 resistor to the ground input. The other ends of all 19 resistors were connected together, and connected to a 5 foot length of insulated wire. The purpose of this wire was to approximately model the antenna effect of the body, in order to assess the common mode rejection capability of the amplifiers. This dummy input was constructed, connected to the head-box of the EEG

machine, and a recording was made of the resulting outputs of the amplifiers. All machine settings were identical to those used in the study. In particular, filters were set to 3 Hz high-pass, and 70 Hz low-pass, and a total of 1024 data points was recorded at a sampling rate of 128 points per second. The 3 midline channels, Fz, Cz and Pz were removed. The resulting matrix of 16 channels of 1024 data points was band-pass filtered to extract the 2 to 8 Hz frequency band. Standard deviation was then computed over the entire matrix. The result of this calculation is the RMS noise amplitude of the recording equipment.

The noise level in the actual data recorded during the study was measured using the following procedure. The data was first band-pass filtered to select the 2 to 8 Hz frequency band. The standard deviation was next computed over the matrix of 1024 points by 16 channels for a single trial, and averaged over 5 randomly selected trials from each session. The result is the RMS signal amplitude. Portions of the data containing eye-blink waveforms were excluded from the computation.

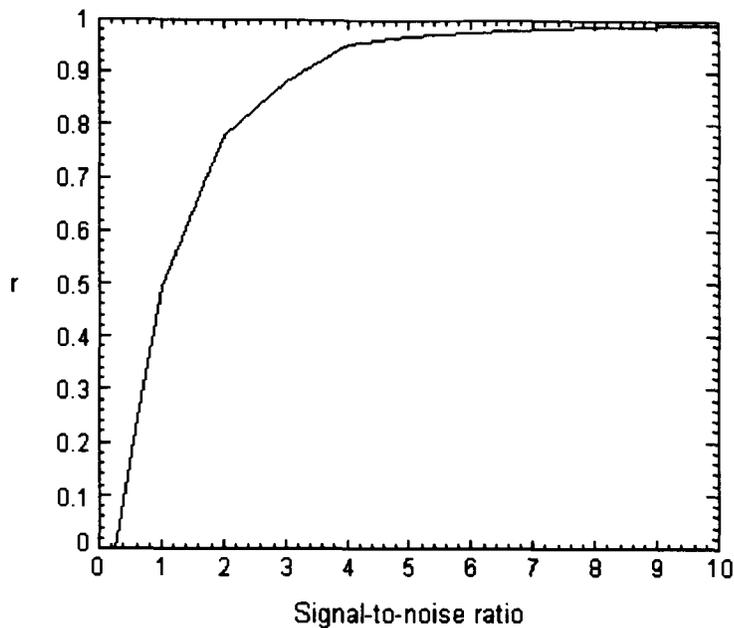
## A1.2 Results and Discussion

The results of the numerical experiment showed that the magnitude of correlation was relatively unaffected when signal-to-noise ratio was greater than approximately 4. At a signal-to-noise ratio of 1, correlation decreased to approximately 0.5. The results are shown in Figure A1.1.

The RMS signal level in the experimental data was found to range between 1.1 micro volts (subject 6) and 2.1 micro volts (subject 7). RMS noise amplitude measured using the dummy load was 0.106 micro volts. This noise is some combination of intrinsic amplifier noise, together with environmental noise. The signal-to-noise ratio for subject 8 is accurate, since the machine noise measurements were made immediately after data from subject 8 had been recorded. The signal to noise ratios for sessions 1 through 7 must be considered to be approximate. The results for all sessions are shown in Table A1.1. Signal to noise ratios range from 10.4 for subject 6 to 19.3 for subject 7.

**Table A1.1 Signal-To-Noise Ratios with Theta Band Filtering**

Subject	Signal (micro volts)	Noise (micro volts)	Signal-to-noise ratio
4	1.75	0.106	16.5
5	2.0	0.106	18.9
6	1.1	0.106	10.4
7	2.05	0.106	19.3
8	1.68	0.106	15.8
9	1.55	0.106	14.6



**Figure A1.1 Pearson product-moment correlation as a function of signal-to-noise ratio.** Correlation was computed between two time series. In each of these, the signal component was a 1 cycle sine-wave. An independent noise component was added to each of these signals. This noise component had a gaussian distribution. The length of the time-series used in the analysis was 32 data points.

The results of the numerical experiment indicate that signal-to-noise ratio does affect magnitude of correlation, and that the effect of noise becomes noticeable at signal-to-noise ratios of less than about 4. The computed estimates of signal-to-noise ratios for the recorded data suggest that external noise is not a problem in this study for data filtered at 2 to 8 Hz.

In order to get an estimate of the signal to noise ratio for other frequency bands, the frequency distribution of the recorded data was examined by performing a Fourier analysis on the data. It was found that the resulting amplitude spectrum was well modeled by a  $1/f$  frequency distribution, over the frequency range of 2 to 64 Hz. That is, the amplitude of the frequency components within the data was found to be inversely proportional to frequency. Using this model, the RMS signal level within bands other than the 2 to 8 Hz band can be estimated. The RMS signal level in frequency bands other than the theta band should be roughly inversely proportional to the ratio between the mean geometric mean frequency of the theta band, 4 Hz, and the geometric mean of the other frequency bands. Using this approach, it is estimated that the resulting signal to noise ratios in the alpha band will be approximately 1/3 of the values for the theta band. For the beta band signal to noise ratio will be 1/6 of the theta band value, and for the gamma band signal to noise ratio will be 1/10 of the theta band value. The

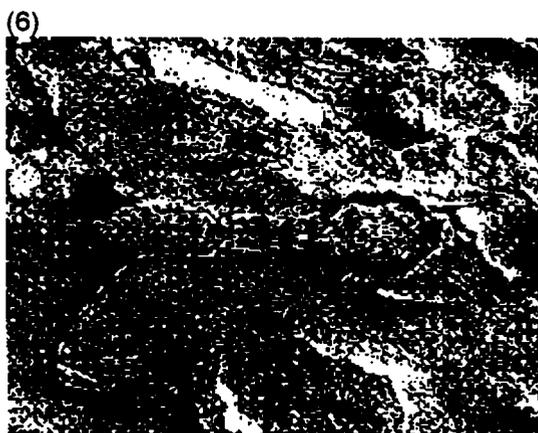
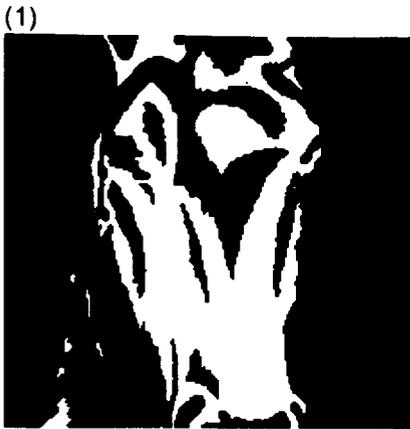
significance of these values is that for the alpha band the signal to noise ratio is marginally high enough to allow its effect on correlation to be ignored. For the beta and gamma bands however there can be expected to be a significant effect of signal to noise ratio on correlation. These effects of signal to noise ratio on correlation would of course be discountable if the level of noise did not vary significantly over the course of a recording. In that case, noise would have the effect of depressing the value of correlation, but this effect would be uniform over time and would thus not impair comparisons made between conditions, and between time-intervals. If on the other hand the level of noise were to change with time then such comparisons could no longer be carried out. It seems reasonable that the component of noise due to the amplifiers themselves should be relatively constant over the time-spans of the data recordings. The component of the noise that is due to external, environmental sources on the other hand can not reasonably be expected to remain constant over the time scale of the recordings. An overall, and conservative conclusion would appear to be that in the present study, using correlational analyses, theta, and to a somewhat lesser extent alpha, band filtering allows for minimal interaction between system noise and level of correlation. On the other hand filtering the data to attempt to extract beta and gamma band components is probably not appropriate when the resulting filtered data is subjected to correlational analysis.

## Appendix 2 - Stimulus Pictures

### Picture Credits

All stimulus pictures used in this study were adapted from the original sources using the methodology outlined in the methods section in Unit 2. None of the items was used directly. Stimuli were adapted from the following sources:

Picture	Adapted From:
1, 8	Mooney and Ferguson (1951)
3	Porter (1954)
5	James (1989)
2, 4, 6, 9, 10, 29	Fogden (1974)
7, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 22, 28	MacKay (1990)
21, 22, 23, 27	Hosking and MacDonnell (1979)
24, 26	Reedy (1973)
25, 30	Frisch (1973)



**Figure A3.1 Stimulus pictures**

(1) horse's head; (2) pheasant; (3) face (4) plover chick (5) Dalmatian dog (6) moth

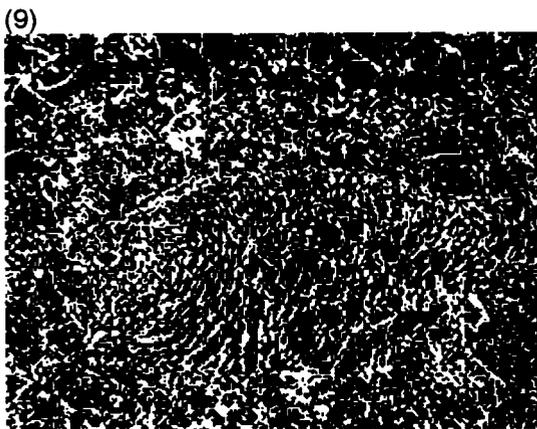


Figure A3.1 (continued) Stimulus pictures  
(7) horses; (8) facial profile; (9) ptarmigan; (10) frog; (11) bear; (12) deer

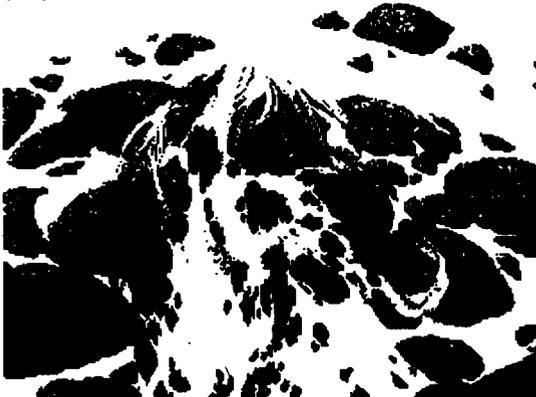
(13)



(14)



(15)



(16)



(17)



(18)



Figure A3.1 (continued) Stimulus pictures  
(13) eagle; (14) pack horses; (15) horse's head; (16) rabbit; (17) horse through trees; (18) horse and rider

(19)



(20)



(21)



(22)



(23)



(24)



Figure A3.1 (continued) Stimulus pictures  
(19) face in rocks; (20) face in rocks; (21) courser; (22) eagle; (23) owl; (24) fist

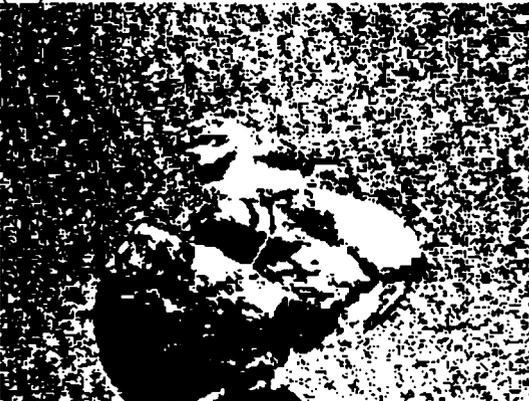
(25)



(26)



(27)



(28)



(29)



(30)



Figure A3.1 (continued) Stimulus pictures  
(25) lamb; (26) rose in hand; (27) plover; (28) leopard; (29) heron; (30) tenrec

## References

- Andersen, R. A. and Mountcastle, V. B. (1983) The influence of the angle of gaze upon the excitability of the light-sensitive neurons of the posterior parietal cortex. Journal of Neuroscience, 3: 532-548.
- Asanuma, C., Andersen, R. A. and Cowan, W. M. (1985) The thalamic relations of the caudal inferior parietal lobule and the lateral prefrontal cortex in monkeys: Divergent cortical projections from cell clusters in the medial pulvinar nucleus. Journal of Comparative Neurology, 241: 357-381.
- Babloyantz, A. and Destexhe, A. (1986) Low-dimensional chaos in an instance of epilepsy. Proceedings of the National Academy of Science, 83: 3513-3517.
- Badii, R., Broggi, G., Derighetti, B., Ravani, M., Ciliberto, S., Politi, A., and Rubio, M. A. (1988). Dimension increase in filtered chaotic signals. Physical Review Letters, 60 (11): 979-982.
- Baker, G. L. and Gollub, J. P. (1990) Chaotic Dynamics, an introduction. Cambridge: Cambridge University Press.
- Barr, M. L. and Kiemann, J. A. (1988) The Human Nervous System: An Anatomical Viewpoint (5th ed.). Philadelphia: Lippincott.
- Basar, E., Basar-Eroglu, J., Roschke, J., and Schult, J. (1990). Strange attractor EEG as sign of cognitive function. In E. Roy John (Ed.), Machinery of the Mind, New York: Birkhauser Press.
- Basar, E., Schürmann, M., Basar-Eroglu, C. and Demiralp, T. (1994) Theta and delta responses in cognitive event-related potential paradigms and their possible psychophysiological correlates. In: H. J. Heinze, T. F. Munte and G. R. Mangun (Eds.) Cognitive Electrophysiology. Boston: Birkhauser, pp. 334-367.
- Birbaumer, N., Lutzenberger, W., Elbert, T. and Trevorrow, T. (1994) Threshold variations in cortical cell assemblies and behaviour. In: H. J. Heinze, T. F. Munte and G. R. Mangun (Eds.) Cognitive Electrophysiology. Boston: Birkhauser, pp. 248-264.
- Braitenberg, V. (1978) Cortical architectonics: General and areal. In: M. A. B. Brazier and H. Petsche (Eds.), Architectonics of the Cerebral Cortex. New York: Raven Press.
- Brennan, J. F. (1991) History and Systems of Psychology. Englewood Cliffs, N. J.: Prentice-Hall.
- Bressler, S. L. (1995) Large-scale cortical networks and cognition. Brain Research Reviews, 20: 288-304.
- Bressler, S. L. (1990) The gamma wave: A cortical information carrier? Trends in Neural Science, 13(5): 161.
- Bressler, S. L., Coppola, R. and Nakamura, R. (1993) Episodic multiregional cortical coherence at multiple frequencies during visual task performance. Nature 366: 153-156.
- Brooks, R. A. (1991) New Approaches to Robotics. Science, 253: 1227-1232.

- Broomhead, D. S. and King, G. P. (1986) Extracting qualitative dynamics from experimental data. Physica 20D: 217-236.
- Bust, J. and Galbraith, G. C. (1975) EEG correlates of visual motor practice in man. Electroencephalography and Clinical Neurophysiology 38: 415-422.
- Chiel, H. J. and Beer, R. D. (1993) Neural and peripheral dynamics as determinants of patterned motor behaviour. In: D. Gardner (Ed.) The Neurobiology of Neural Networks. Cambridge Mass.: MIT Press, pp. 137-164.
- Churchland, P. (1995) The Engine of Reason, the Seat of the Soul. Cambridge, Mass.: MIT Press.
- Cohen, J. (1977) Statistical Power Analysis for the Behavioral Sciences (rev. ed.). New York: Academic Press.
- Damasio, A. R. and Damasio, H. (1994) Cortical systems for retrieval of concrete knowledge: The convergence zone framework. In: C. Koch and J. L. Davis (Eds.) Large-Scale Neuronal Theories of the Brain, Cambridge, Mass.: MIT Press.
- Damasio, A. R. and Damasio, H. (1993) Cortical systems underlying knowledge retrieval: Evidence from human lesion studies. In: T. A. Poggio and D. A. Glaser (Eds.) Exploring Brain Functions: Models in Neuroscience, Chichester: Wiley.
- Damasio, A. R., Damasio, H. and Tranel, H. [1990] Impairment of visual recognition as clues to the processing of memory. In: G. M. Edelman, W. E. Galland W. M. Cowan [Eds.], Signal and Sense: Local and Global Order in Perceptual Maps. New York: Wiley.
- Damasio, A. R., Tranel, H. and Damasio, H. (1989) Amnesia caused by herpes simplex encephalitis, infarctions in basal forebrain, Alzheimer's disease, and anoxia. In F. Boller and J. Grafman (Eds.), Handbook of Neuropsychology, Vol. 3. Amsterdam: Elsevier, pp. 149-166.
- Damasio, A. R., Yamada, T., Damasio, H., Corbett, J. and McKee, J. (1980) Central achromatopsia: Behavioral, anatomic and physiologic aspects. Neurology, 30: 1064-1071.
- Desimone, R., Albright, T. D., Gross, C. G., and Bruce, C. [1984] Stimulus-selective properties of inferior temporal neurons in the macaque. Journal of Neuroscience 4: 2051-2062.
- Desimone, R., Miller, E. K., Chellazi, L. and Lueschow, A. (1995) Multiple memory systems in the visual cortex. In: M. Gazzaniga (Ed.) The Cognitive Neurosciences, Cambridge, Mass.: MIT Press, pp. 475-486.
- Destexhe, A., Sepulchre, J. A., and Babloyantz, A. (1988) A comparative study of the experimental quantification of deterministic chaos. Physics Letters A, 132: 101-106.
- DeYoe, E. A. and Van Essen, D. C. [1988] Concurrent processing streams in monkey visual cortex. Trends in Neuroscience 11: 219-226.
- Diamond, M. C., Scheibel, A. B. and Elson, L. M. (1985) The Human Brain Coloring Book. New York: HarperPerennial.
- Dvorak, I. and Siska, J. (1986) On some problems encountered in the estimation of the correlation dimension of the EEG. Physics Letters A, 118 (2): 63-66.

- Eccles, J. C. (1981) the modular operation of the cerebral neocortex considered as the material basis of mental events. Neuroscience 6: 1839-1856.
- 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.
- Edelman, G. M. (1992) Bright Air, Brilliant Fire. New York: Basic Books.
- Engel, A. K., Konig, P., Gray, C. M. and Singer, W. (1990) Stimulus-dependent neuronal oscillations in cat visual cortex: Inter-columnar interaction as determined by cross-correlation analysis. European Journal of Neuroscience 2: 588-606.
- Engel, A. K., Konig, P., Kreiter, A. K. and Singer, W. (1991) Interhemispheric synchronization of oscillatory neuronal responses in cat visual cortex. Science 252: 1177-1178.
- Engel, A. K., Konig, P., Kreiter, A. K., Schillen, T. B. and Singer, W. (1992) Temporal coding in the visual cortex: New vistas on integration in the nervous system. Trends in Neuroscience 15(6): 218-226.
- Ermentrout, B. (1994) An introduction to neural oscillators. In F. Ventriglia (Ed.) Neural Modeling and Neural Networks. Oxford: Pergamon Press, pp. 79-110.
- Farmer, J. D., Ott, E. and Yorke, J. A. (1983) The dimension of chaotic attractors. Physica 7D: 153-180.
- Fausett, D. W. (1990) Strictly local backpropagation. International Joint Conference on Neural Networks, Vol. 3, San Diego, Calif., pp. 125-130.
- Felleman, D. J. and Van Essen, D. C. [1991] Distributed hierarchical processing in the primate cerebral cortex. Cerebral Cortex 1: 1-47.
- Fetz, E. E. (1993) Dynamic neural network models of sensorimotor behaviour. In D. Gardner (Ed.) The Neurobiology of Neural Networks. Cambridge Mass.: MIT Press, pp. 165-190.
- Finkel, L. H. and Edelman, G. M. (1989) Integration of distributed cortical systems by reentry: A computer simulation of interactive functionally segregated visual areas. The Journal of Neuroscience, 9(9): 3188-3208.
- Flotzinger, D. Kalcher, J. and Pfurtscheller, G. (1993) Suitability of learning vector quantization for on-line learning: A case study of EEG classification. Proceedings of the World Conference on Neural Networks (WCCN - 93), Vol. 1. New Jersey: Lawrence Erlbaum Associates.
- Fogden, M. and Fogden, P. (1974) Animals and their Colors. New York: Crown.
- Frank, G. W., Lookman, T., Nerenberg, M. A. H., Essex, C., Lemieux, J., and Blume, W. (1990) Chaotic time-series analyses of epileptic seizures. Physica 46D: 427-438.
- Fraser, A. M. and Swinney, H. L. (1986) Independent coordinates for strange attractors from mutual information. Physical Review A 33(2): 1130-1140.

- Frisch, O. (1973) Animal Camouflage. London: Collins.
- Fujita, I., Tanaka, K., Ito, M. and Cheng, K. (1992) Columns for visual features of objects in monkey inferotemporal cortex. Nature 360: 343-346.
- Gabor, Andrew J. and Seyal, Masud (1992) Automated interictal EEG spike detection using artificial neural networks. Electroencephalography & Clinical Neurophysiology, 83(5): 271-280.
- Gallant, S. I. (1993) Neural Network Learning and Expert Systems. Cambridge, Mass.: MIT Press.
- Gallez, D. and Babloyantz, A. (1991) Predictability of human EEG: a dynamical approach. Biological Cybernetics, 4 (5): 381-391.
- Gevins, A. S., Morgan, N. H. and Bressler, S. L. (1987) Human neuroelectric patterns predict performance accuracy. Science 235: 580-585.
- Gilbert, C. D. (1992) Horizontal integration and cortical dynamics. Neuron 9: 1-13.
- Gilbert, C. D. (1995) Dynamic properties of adult visual cortex. In: M. Gazzaniga (Ed.) The Cognitive Neurosciences, Cambridge, Mass.: MIT Press, pp. 73-90.
- Gilbert, C. D. and Wiesel, T. N. (1989) Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. Journal of Neuroscience 9: 2432-2442.
- Goldberg, D. E. (1989) Genetic Algorithms in Search, Optimization, and Machine Learning. Reading Mass.: Addison-Wesley.
- Goldman-Rakic, P. (1988) Topography of cognition: Parallel distributed networks in primate association cortex. Annual Review of Neuroscience, 11: 137-156.
- Goldman-Rakic, P. (1992) Working memory and the mind. Scientific American, 267: 111-117.
- Goodale, M. A. and Milner, A. D. (1992) Separate visual pathways for perception and action. Trends in Neuroscience 15: 20-55.
- Grady, C. L. and Haxby, J. V. (1992) Dissociation of object and spatial vision in human extrastriate cortex: Age-related changes in activation of regional cerebral blood flow measured with <sup>15</sup>O water and positron emission tomography. Journal of Cognitive Neuroscience 4(1): 23-34.
- Grassberger, P. and Procaccia, I. (1983) Characterization of strange attractors. Physical Review Letters, 50 (5): 346-349.
- Gray, C. M., Konig, P., Engel, A. K. and Singer, W. (1989) Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. Nature 338: 224-337.
- Gray, C. M. and Singer, W. (1989) Stimulus specific neuronal oscillations in orientation columns of cat visual cortex. Proceedings of the National Academy of Science 86: 1698-1702.
- Gray, R. M. (1990) Entropy and Information Theory. New York: Springer-Verlag.

- Gregson, R. A. M., Briton, L. A., Campbell, E. A. and Gates, R. (1991) Comparisons of the nonlinear dynamics of electroencephalograms under various task loading conditions: A preliminary report. Biological Psychology, 31: 173-191.
- Gregson, R. A. M., Campbell, E. A. and Gates, R. (1992) Cognitive load as a determinant of the dimensionality of the electroencephalogram: A replication study. Biological Psychology, 35: 165-178.
- Gross, C. G., Rocha-Miranda E. C. and Bender, D. B. (1972) Visual properties of neurons in inferotemporal cortex of the macaque. Journal of Neurophysiology 35: 96-111.
- Grozinger, M., Kloppel, b. and Roschke, J. (1993) Recognition of rapid-eye movement (REM) sleep by artificial neural networks. Proceedings of the World Conference on Neural Networks (WCCN - 93), Vol. 1. New Jersey: Lawrence Erlbaum Associates.
- Haxby, J. V., Grady, C. L., Horwitz, B., Ungerleider, L. G., Mishkin, M., Carson, R. E., Herscovitch, P., Schapiro, M. B. and Rapoport, S. I. (1991) Dissociation of object and spatial visual processing pathways in human extrastriate cortex. Proceedings of the National Academy of Science 88: 1621-1625.
- Haxby, J. V. and Horwitz, B. (1994) The functional organization of human extrastriate cortex: A PET rCBF study of selective attention to faces and locations. Journal of Neuroscience 14(11, Pt 1): 6336-6353.
- Hebb, D. O. (1949) The Organization of Behavior. New York: Wiley
- Hermann, B. P. and Seidenberg, M. (1993) Dissociation of object recognition and spatial localization abilities following temporal lobe lesions in humans. Neuropsychology 7(3): 343-350.
- Holland, J. (1975) Adaptation in Natural and Artificial Systems. Ann Arbor: The University of Michigan Press.
- Horwitz, B., Grady, C. L., Haxby, J. V., Schapiro, M. B., Rapoport, S. I., Ungerleider, L. G. and Mishkin, M. (1992) Functional associations among human posterior extrastriate brain regions during object and spatial vision. Journal of Cognitive Neuroscience 4(4): 311-322.
- Hosking, E. J. and MacDonnell, K. (1979) A passion for birds : Fifty years of photographing wildlife. New York: Coward, McCann & Geoghegan.
- Hubel, D. H. and Wiesel, T. N. (1962) Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. Journal of Physiology (London) 160: 106-154.
- Hubel, D. H. and Wiesel, T. N. (1968) Receptive fields and functional architecture of monkey striate cortex. Journal of Physiology (London) 195: 215-243.
- James, R. C. (1989) Photograph reproduced in E. B. Goldstein, Sensation and Perception. Belmont Calif.: Wadsworth.
- Jando, G., Siegel, R. M., Horvath, Z and Buzaki, G. (1993) Pattern recognition of the electroencephalogram by artificial neural networks. Electroencephalography and Clinical Neurophysiology, 86: 100-109.

- Jansen, B. H. (1991) "Is it?" and "So what?" - A critical view of EEG chaos. In D. W. Duke and W. S. Pritchard (Eds.) Proceedings of the Conference on Measuring Chaos in the Human Brain. Singapore: World Scientific.
- Kandel, E. and Schwartz, J. H. (1985) Principles of Neural Science. San Francisco: Elsevier North Holland.
- Kanizsa, G. (1979) Organization in Vision: Essays in Gestalt Perception. New York: Praeger.
- Keppel, G. (1991) Design and Analysis: A Researcher's Handbook. New Jersey: Prentice-Hall.
- Kertesz, A. (1994) Localization and Neuroimaging in Neuropsychology. San Diego: Academic Press.
- Kloppel, B. (1994a) Neural networks as a new method for EEG analysis. Pharmacoelectroencephalography, 29: 33-38.
- Kloppel, B. (1994b) Application of neural networks for EEG analysis. Pharmacoelectroencephalography, 29: 39-46.
- Konig, P. and Schillen, T. B. (1991) Stimulus dependent assembly formation of oscillatory responses: I. Synchronization. Neural Computation 3(2): 155-166.
- Leahey, T. H. (1987) A History of Psychology: Main Current in Psychological Thought, Second Edition. Englewood Cliffs, N. J.: Prentice-Hall.
- Leibniz, Gottfried Wilhelm (1965). The monadology and other philosophical writings, translated by Robert Latta. London: Oxford University Press.
- Llinas, R., Grace, A. A. and Yarom Y. (1991) In vitro neurons in mammalian cortical layer 4 exhibit intrinsic oscillatory activity in the 10 to 50 Hz frequency range. Proceedings of the National Academy of Sciences 88: 897-901.
- Llinas, R. R. and Ribary, U. (1993) Coherent 40-Hz oscillation characterizes dream state in humans. Proceedings of the National Academy of sciences, 90: 2078-2081.
- Lorenz, E. N. (1963) Deterministic non-periodic flow. Journal of Atmospheric Science, 20: 130-141.
- Lueschow, A., Miller, E. K. and Desimone, R. (1993) Effect of stimulus transformations on short-term memory mechanisms in inferior temporal cortex. Society of Neuroscience Abstracts 23: 975.
- Lutzenberger, W., Birbaumer, N., Flor, H., Rockstroh, B. and Elbert, T. (1992) Dimensional analysis of the human EEG and intelligence. Neuroscience Letters, 143: 10-14.
- Lutzenbeger, W., Elbert, T., Birbaumer, N., Ray, W. J. and Schupp, H. (1992) The scalp distribution of the fractal dimension of the EEG and its variation with mental tasks. Brain Topography, 5(1): 27-34.
- MacGregor, R. J. (1993) Theoretical Mechanics of Biological Neural Networks. Boston: Academic Press.
- MacKay, E. (1990) The Art of Bev Doolittle. New York: Bantam Books.

- Maunsell, F. H. R. and Van Essen, D. C. (1983) Functional properties of neurons in middle temporal visual area (MT) of macaque monkey. I. Selectivity for stimulus direction, velocity and orientation. Journal of Neurophysiology 49: 1127-1147.
- Maunsell, F. H. R. and Ferrera, V. P. (1995) Attentional mechanisms in visual cortex. In: M. Gazzaniga (Ed.) The Cognitive Neurosciences, Cambridge, Mass.: MIT Press, pp. 451-461.
- Maunsell, F. H. R., Sclar, G., Nealey, T. A. and DePriest, D. D. (1991) Extraretinal representations in area V4 in the macaque monkey. Visual Neuroscience 7: 561-573.
- May, R. M. (1976) Simple mathematical models with very complicated dynamics. Nature 261: 459-467.
- Mayer-Kress, G. and Layne, S. P. (1987) Dimensionality of the human electroencephalogram. In S. H. Koslow (Ed.), Perspectives in Biological Dynamics and Theoretical Medicine. Annals of the New York Academy of Sciences, 54: 62-87.
- McClelland, J. L. and Rumelhart, D. E. (1988) Explorations in Parallel Distributed Processing. Cambridge, Mass.: MIT Press.
- Meunier, M., Bachevalier, J., Mishkin, M. and Murray, E. A. (1993) Effects on visual recognition of combined and separate ablations of the interrhinal and perirhinal cortex in rhesus monkeys. Journal of Neuroscience 13: 5418-5432.
- Miller, E. K. and Desimone, R. (1994) Parallel neuronal mechanisms for short term memory. Science 263: 520-522.
- Miller, E. K., Li, L. and Desimone, R. (1993) Activity of neurons in anterior inferior temporal cortex during a short-term memory task. Journal of Neuroscience 13: 1460-1478.
- Miller, R. (1991) Cortico-Hippocampal Interplay and the Representation of Contexts in the Brain. Berlin: Springer.
- Minsky, M. L. (1968) Matter, mind, and models. In M. L. Minsky (Ed.) Semantic Information Processing. Cambridge, Mass.: MIT Press.
- Mishkin, M. (1993) Cerebral memory circuits. In: T. A. Poggio and D. A. Glaser (Eds.) Exploring Brain Functions: Models in Neuroscience. Chichester: John Wiley & Sons.
- Mishkin, M. and Appenzeller, T. (1987) The anatomy of memory. Scientific American 256: 62-72.
- Mishkin, M. and Manning, F. J. (1978) Non-spatial memory after selective prefrontal lesions in monkeys. Brain Research 143: 313-323.
- Mishkin, M., Ungerleider, L. G. and Macko, K. (1983) Object vision and spatial vision: Two cortical pathways. Trends in Neural Science, 6: 414-417.
- Mitzdorf, U. (1985) Current source-density method and application in cat cerebral cortex: Investigation of evoked potentials and EEG phenomena. Physiological Review, 65: 37-99.
- Miyashita, Y. (1988) Neural correlate of visual associative long-term memory in the primate visual cortex. Nature, 335: 817-820.

- Moon, F. C. (1987) Chaotic Vibrations: An Introduction for Applied Scientists and Engineers. New York: Wiley.
- Mooney, C. M. and Ferguson, G. A. (1951) A new test of closure. Canadian Journal of Psychology, 5: 129-133.
- Moran, J. and Desimone, R. (1985) Selective attention gates visual processing in the extrastriate cortex. Science 229: 782-784.
- Nagel, T. (1974) What is it like to be a bat? Philosophical review, 83(4): 435-450.
- Nan, X. and Jinghua, X. (1988) The fractal dimension of EEG as a physical measure of conscious human brain activities. Bulletin of Mathematical Biology, 50: 559-565.
- Newcombe, F. and Ratcliffe, G. (1987) Dissociable visual and spatial impairments following right posterior cerebral lesions: Clinical, neuropsychological and anatomical evidence. Neuropsychologia 25(1-B): 149-161.
- Nieuwenhuys, R., Voogd, J. and van Huijzen, C. (1981) The Human Central Nervous System: A Synopsis and Atlas (2nd revised ed.). Berlin: Springer-Verlag.
- Otnes, R. K. and Enochson, L. D. (1972) Digital Time Series Analysis. New York: John Wiley & Sons.
- Otnes, R. K. and Enochson, L. D. (1978) Applied Time Series Analysis. New York: John Wiley & Sons.
- Palus, M., Dvorak, I., and David, I. (1992) Spatio-temporal dynamics of human EEG. Physica A, 185 (1/4): 433.
- Parzen, E. (1962) On estimation of a probability density function and mode. Annals of Mathematical Statistics, 33: 1065-1076.
- Peitgen, H., Jurgens, H. and Saupe, D. (1992) Fractals for the Classroom, Part Two. New York: Springer-Verlag.
- Perrett, D. I. and Oram, M. W. (1993) Neurophysiology of shape processing. Image Visual Computation, 11: 317-333.
- Petsche, H., Lacroix, D., Lindner, K., Rappelsberger, P. and Schmidt-Henrich, E. (1992) Thinking with images or thinking with language: A pilot EEG probability mapping study. International Journal of Psychophysiology 12: 31-39.
- Pfurtscheller, G., Flotzinger, D., Mohl, W. and Peltoranta, M. (1992) Prediction of the side of hand movements from single-trial multi channel EEG data using neural networks. Electroencephalography & Clinical Neurophysiology 82(4): 313-315.
- Pijn, J. P. M., Van Neerven, J., Noest, A. and da Silva, F. H. (1991) Chaos or noise in EEG signals: Dependence on state and brain site. Electroencephalography and clinical neurophysiology, 79(5): 371-381.
- Pijn, J. P. M., Vijn, P. C. M., Lopes da Silva, F. H., Van Emde Boas, W. and Blanes, W. (1989) The use of signal-analysis for the localization of an epileptogenic focus: A new approach. Advances in Epileptology 17: 272-276.

- Porter, P. B. (1954) Another picture puzzle. American Journal of Psychology, 67: 550-551.
- Principe, J. C. and Lo, P. (1991) Towards the determination of the largest Lyapunov exponent of EEG segments. In D. W. Duke and W. S. Pritchard (Eds.) Proceedings of the Conference on Measuring Chaos in the Human Brain. Singapore: World Scientific.
- Pritchard, W. S. and Duke, D. W. (1992). Dimensional analysis of no-task human EEG using the Grassberger-Procaccia method. Psychophysiology, 29 (2): 182-192.
- Rapp, P. E., Bashore, T. R., Martinerie, J. M., Albano, A. M., Zimmerman, I. D. and Mees, A. I. (1989) Dynamics of Brain Electrical Activity. Brain Topography, 2(1-2): 99-118.
- Rappelsberger, P. and Petsche, H. (1988) Probability mapping: Power and coherence analyses of cognitive processes. Brain Topography 1(1): 46-54.
- Ray, W. J., Well, R. and Elbert, T. (1991) EEG and chaos: Dimensional estimation of sensory and hypnotic processes. In D. W. Duke and W. S. Pritchard (Eds.), Proceedings of the Conference on Measuring Chaos in the Human Brain. Singapore: World Scientific.
- Ray, W. J. and Cole, H. (1985) EEG alpha reflects attentional demands, beta reflects emotional and cognitive processes. Science, 228: 750-752.
- Reber, A. S. (1985) Dictionary of Psychology. London: Penguin.
- Reedy, W. H. (1973) Impact: Photography for Advertising. New York: Eastman Kodak.
- Regan, D. (1989) Human Brain electrophysiology - Evoked Potentials and Electromagnetic Fields in Science and Medicine. Toronto: Elsevier.
- Richardson, T. (1995, personal communication) Professor, Department of Kinesiology, School of Applied Science, Simon Fraser University.
- Roland, P. E. (1982) Cortical regulation of selective attention in man. Journal of Neurophysiology, 48: 1059-1078.
- Rolls, E. T. (1989) Parallel distributed processing in the brain: implications of the functional architecture of neuronal networks in the hippocampus. In R. G. M. Morris (Ed.), Parallel Distributed Processing: Implications for Psychology and Neurobiology. Oxford: Oxford University Press, pp. 286-308.
- Rosenweig, M. R. and Leiman, A. L. (1989) Physiological Psychology. New York: Random House.
- Schaffer, W. M., Truty, G. L. and Fulmer, S. L. (1988) Dynamical Software: User's manual and introduction to chaotic systems. Tucson: Dynamical Systems Inc.
- Searle, J. (1987) Minds and brains without programs. In C. Blakemore and S. Greenfield (Eds.), Mindwaves. Oxford: Basil Blackwell.
- Shacter, D. L., Chiu, C. Y. and Ochsner, K. N. (1993) Implicit memory: A selective review. Annual Review of Neuroscience, 16: 159-182.
- Schiller, P. H. and Lee, K. (1991) The role of the primate extrastriate area V4 in vision. Science 251: 1251-1253.

- Shannon, C. E. (1948) A mathematical theory of communication. Bell System Technical Journal 27: 379-423.
- Slater, J. D., Wu, F. Y., Honig, L. S., Ramsay, R. E. and Morgan, R. (1994) Neural network analysis of the P300 event-related potential in multiple sclerosis. Electroencephalography and clinical neurophysiology, 90: 114-122.
- Smith, L. A. (1992) Comment: Relation between statistics and chaos. Statistical Science, 7(1): 109-113.
- Smith, Murray (1993) Neural networks for statistical modeling. New York: Van Nostrand Reinhold.
- Smolensky, P. (1988) On the proper treatment of connectionism. Behavioral and Brain Sciences 11: 1-74.
- Specht, D. F. (1990) Probabilistic Neural Networks. Neural Networks, 3: 109-118.
- Specht, D. F. (1991) A general regression neural network. IEEE Transaction on Neural Networks, 2(6): 568-576.
- Steriade, M., Curro Dossi, R., Pare, D. and Oakson, G. (1991) Fast oscillations (20 - 40 Hz) in thalamocortical systems and their potentiation by mesopontine cholinergic nuclei in the cat. Proceedings of the National Academy of Sciences 88(10): 4396-4400.
- Takens, F. (1980) Detecting strange attractors in turbulence. Dynamical Systems and Turbulence. Lecture Notes in Mathematics, 898: 366-381.
- Tanaka, K., Saito, H., Fukada, Y. and Moriya, M. (1991) Coding visual images of objects in the inferotemporal cortex of the macaque monkey. Journal of Neurophysiology 66: 170-189.
- Thatcher, R. W., Krause, P. J. and Hrybyk, M. (1986) Cortico-cortical associations and EEG coherence: A two-compartmental model. Electroencephalography and clinical neurophysiology, 64: 123-143.
- Tresch, M. C. and Sinnamon, H. M. (1993) Double dissociation of spatial and object visual memory: Evidence from selective interference in intact human subjects. Neuropsychologia, 31(3): 211-219.
- T'so, D. Y., Gilbert, C. D. and Wiesel, T. N. (1986) Relationships between horizontal interactions and functional architecture in cat striate cortex as revealed by cross-correlation analysis. Journal of Neuroscience 6: 1160-1170.
- Turing, A. M. (1937) On computable numbers, with an application to the Entscheidungsproblem. Proceedings of the London Mathematical Society, 42: 230-265.
- Ullman, S. and Shashua, A. (1988) Structural saliency: The detection of globally salient structures using a locally connected network. Cambridge, Mass.: Artificial Intelligence Laboratory, MIT, A.I. Memo No. 1061.
- Ungerleider, L. G. and Mishkin, M. (1982) Two cortical visual systems. In: D. J. Ingle, M. A. Goodale and R. J. W. Mansfield, (Eds.) Analysis of Visual Behaviour. Cambridge, Mass.: MIT Press.