

REFERENCE ONLY



UNIVERSITY OF LONDON THESIS

Degree phd Year 2007 Name of Author OLIVER
HULME

COPYRIGHT

This is a thesis accepted for a Higher Degree of the University of London. It is an unpublished typescript and the copyright is held by the author. All persons consulting the thesis must read and abide by the Copyright Declaration below.

COPYRIGHT DECLARATION

I recognise that the copyright of the above-described thesis rests with the author and that no quotation from it or information derived from it may be published without the prior written consent of the author.

LOAN

Theses may not be lent to individuals, but the University Library may lend a copy to approved libraries within the United Kingdom, for consultation solely on the premises of those libraries. Application should be made to: The Theses Section, University of London Library, Senate House, Malet Street, London WC1E 7HU.

REPRODUCTION

University of London theses may not be reproduced without explicit written permission from the University of London Library. Enquiries should be addressed to the Theses Section of the Library. Regulations concerning reproduction vary according to the date of acceptance of the thesis and are listed below as guidelines.

- A. Before 1962. Permission granted only upon the prior written consent of the author. (The University Library will provide addresses where possible).
- B. 1962 - 1974. In many cases the author has agreed to permit copying upon completion of a Copyright Declaration.
- C. 1975 - 1988. Most theses may be copied upon completion of a Copyright Declaration.
- D. 1989 onwards. Most theses may be copied.

This thesis comes within category D.

- This copy has been deposited in the Library of UCL
- This copy has been deposited in the University of London Library, Senate House, Malet Street, London WC1E 7HU.

PhD Thesis:

Visibility, Invisibility and Reportability

Oliver James Hulme

University of London

October 2002- October 2005

Anatomy Department, Wellcome Department of Imaging Neuroscience.

University College London.

Gower Street, London. WC1E6BT.

Email:

UMI Number: U592915

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U592915

Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Contents:

Abstract and Author Information

A. Abstract.....6
B. Acknowledgements8
C. Publications.....10

Glossary and Overview

D. Glossary11
E. General Overview.....15
F. Overview Experiments 1.....16
G. Overview Experiment 2.....19

1. Historical Survey

1.1 Introduction to Historical Survey.....23
1.2 Functional Specialisation.....23
1.3 Faces and Objects.....31
1.4 Places and Houses.....36
1.5 Visual perception and awareness.....37

2. Theoretical Introduction

2.1 Theory of Microconsciousness.....41
2.2 Reportability and Consciousness.....42
2.3 Philosophical Enquiries.....44
2.4 Scientific Enquiries.....49
2.4.1 Lamme.....51
2.4.2 Global workspace Models.....57
2.4.3 Baars.....57
2.4.4 Global Neuronal Workspace.....58

3. Signal Detection Theory

3.1 Introduction to Signal Detection Theory.....61
3.2 Neuroimaging and Signal Detection Theory.....68

4. Methodological Introduction

4.1 Introduction to fMRI.....	70
4.2 Analysis of fMRI data.....	72
4.2.1 SPM.....	74

5. Reiteration of Aims

5.1 Reiteration of Aims.....	75
-------------------------------------	-----------

6. Methods and Results

6.1 Methods Experiment 1.....	76
6.2 Results Experiment 1.....	81
6.3 Methods Experiment 2.....	90
6.3.1 Psychophysics.....	90
6.3.2 Scanning.....	97
6.4 Results Experiment 2.....	100
6.4.1 Position General effects.....	108
6.4.1.1 Position general stimulus effects.....	108
6.4.1.2 Position general decision effects.....	109
6.4.1.3 Position general report effects.....	116
6.4.2 Position specific effects	120
6.4.2.1 Stimulus effects.....	120
6.4.2.2 Decision and report effects.....	126
6.4.2.3 Attentional effects.....	126
6.5 Anatomical analysis of attentional effects in the thalamus.....	129

7. Discussion

7.1 Discussion Experiment 1.....	136
7.2 Discussion Experiment 2.....	141
7.2.1 Psychophysics.....	141
7.2.2 Scanning	147
7.2.2.1 Stimulus specific responses.....	147
7.2.2.2 Decision specific responses.....	160
7.2.2.3 Attention specific responses.....	161

7.2.2.4 Report specific responses.....	163
7.2.3 Interpretational problems.....	164
7.3 A new perspective	170
8. Theory	
8.1 Experiment 1.....	172
8.2 Experiment 2.....	177
9. General theory	
9.1 Phenomenal consciousness.....	180
9.2 Access consciousness.....	187
10. Conclusions	
10.0 Conclusion.....	189
11 Appendices	
Appendix 1.....	191
Appendix 2.....	192
Appendix 3.....	193
<i>12 References</i>	
<i>13 Quotes</i>	

Declaration:

I Oliver James Hulme, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

List of tables

Table 1 shows performance values, visibility ratings and stabilisation times for the different delay conditions in the contrast calibration phase.

Table 2 shows visibility ratings and performance values for different delay values

Table 3 shows performance and visibility ratings for the different delay conditions of in-scanner psychophysics.

Table 4 contains the anatomical labels for each significant cluster identified with two or more spatial selectivities.

Table 5 shows a taxonomy of reportability.

A. Abstract

This thesis attempts to investigate the relationship between neural activity and visual consciousness. It attacks two problems central to this endeavour.

The first problem is that of visibility and invisibility. This is addressed by looking at the neural determinants of whether something can be seen (or not). The relationship between neural activity and object perception has received considerable attention using sophisticated stimulus manipulations such as masking and dichoptic presentation. Here we investigated this relationship simply by occluding objects with an opaque screen that dissociated the direct perception of the object from the awareness of its presence. We used fMRI to measure brain activity when subjects viewed objects (faces and houses) that were occluded and found that the response of object selective areas is invariant to whether the object is visible or occluded. This suggests that when objects are directly viewed, the response of object-selective regions is due to the awareness of the object's presence, not just the direct perception of it.

The second problem is that of reportability. This is addressed by looking at the neural determinants of how visual stimuli are reported. Understanding the neural underpinnings of reportability is critical since experiments on consciousness invariably employ reportability as an 'index' of consciousness. This involves operationally defining whether the subject was conscious of something by simply asking 'did you see it?'. It can be argued that this operationalisation conflates two processes; perception and its reportability whose neural underpinnings are potentially dissociable. In experiment 2 I used a 'partial report methodology' to dissociate the neural correlates of these two processes in an fMRI study of object detection. The task involved presenting a visual array of objects, and then presenting a spatial cue after a variable delay. Using a signal detection theoretic approach we were able to decouple 4 stages of stimulus reportability, stimulus processing, attentional selection, perceptual decision and motoric report. We show that retinotopically specific responses in early visual cortex reflect stimulus processing, a network of parietal and temporal regions reflects perceptual decisions, fronto-parietal-pulvinar networks

reflect attentional selection and motor regions reflect report. These findings provide a tentative neural framework with which to evaluate the relationship between 'consciousness' and reportability.

By consideration of both findings I attempt to synthesise a speculative theoretical perspective that unites the neural correlates of visibility, invisibility and reportability into a consistent framework.

B. Acknowledgements

I begin by expressing my sincere gratitude to Prof. Zeki for his support along the rocky road of my PhD. His dedication and no-nonsense approach to science, academia and life in general has been influential. Particular to his approach is the strive for the simple, the elegant and the bold. This attitude has undoubtedly influenced my approach to science and the undertaking of this thesis. The attitude in the lab that no topic is beyond the reach of science has also led me to receive a broader scientific education than I anticipated. Working in the Zeki lab is a pretty unique experience, with numerous parties, games of darts, bull sessions (informal debates to those outside of the lab), long evenings at SOAS, and Californian beach parties.

There have been many people over the years that have helped me from the lab. I would particularly like to thank Matt Self who gave me technical training, instilled the basics of statistics and experimental design, and taught (often forced) me to maintain a life-work balance. His attitude towards helping others was bordering on the truly altruistic, and this has had a profound and long-lasting influence on the sense of teamwork in the lab. I would like to thank Richard Lewis for the energy, exuberance and verbosity that he brings to the lab and everything else he embraces. I would like to thank Barrie Roulston for his companionship, sense of humour (particularly his flamboyant dancing and his impersonations of Christoph Koch) and his helpful input with the PhD. I would like to thank everyone in the lab, past and present, for the nights out where discussions about science continued for hours after work. (This is admittedly sad but to be expected). I would also like to thank Stuart for his no-nonsense scepticism, depth of knowledge, and in general his input on my ideas (particularly the 2nd experiment reported here). For technical and mathematical matters I have to thank Karl Friston whose methodological input was priceless. Then there is John and Thomas who have helped me with technical matters on almost a daily basis without complaint.

I would like to express my gratitude to my girlfriend Louise Whiteley for her help with all aspects of the thesis from advice on statistics, grammar and basic punctuation to most importantly moral and academic support and belief. Last of all I would like to thank my parents without whose love and support I would probably not have been

able to do a PhD. I hope for them that the end of this PhD spells an end to my fiscal dependence on them (at the age of 26!).

C. Publications

Abstracts

O. Hulme, S. Zeki. VISUAL PERCEPTION AND VISUAL KNOWLEDGE ACTIVATE COMMON AREAS Program No. 176.2. 2004 Abstract Viewer/Itinerary Planner. Washington, DC: Society for Neuroscience, 2004. Online.

Hulme, O.J, Chumbley, J., Eickhoff, S.B., Prince, S., Shipp, S.(2006). Functional Topology of Attention in the Human Pulvinar. CoSyne 2006

Papers

Hulme, O. and Zeki, S. (2006). The Sightless View: Correlates of Occluded Objects. Cerebral Cortex July 14

Hulme, O. Friston, K and Zeki, S. (2006). Neural Correlates of Stimulus Reportability. In Preparation

Hulme, O.J., Shipp, S., Functional Topology of Attention in the Human Pulvinar (2006) In preparation

Talks

Visibility, Invisibility and Reportability, Anatomy Dept. UCL, England, 2005

D. Glossary

When discussing conceptual (some would say ‘woolly’) issues such as perception and consciousness, it is easy for people to talk at cross-purposes unless one is explicit about the definition and use of key terms. Unfortunately there are no canonical definitions for words such as ‘consciousness’ and ‘perception’ in a neuroscientific context, partly because their migration from philosophy and common parlance has been haphazard. In the same way the Inuits are purported to have an arsenal of different words for snow, neuroscientists/philosophers/psychologists use many words for consciousness (as well as many definitions of each one). I will therefore begin by defining the key terms I will use in this thesis.

Visual perception here shall be used as a synonym for visual experience; ‘*seeing things*’. In this use of the term it is contradictory to use the phrase ‘*unconscious perception*’ and tautological to use the phrase ‘*conscious perception*’.

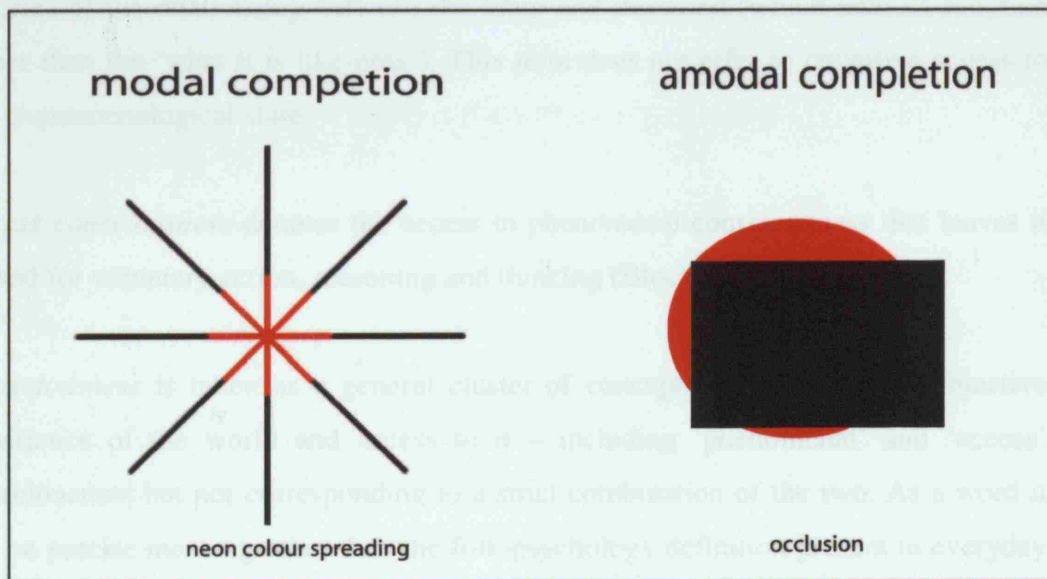


Figure 1. Demonstrates the difference between modal and amodal perception.

Modal perception is the direct perception of something (Helmholtz 1876, Michotte 1950, Gibson 1950); a perception with ‘substance’ (see figure 1). In the example redness spills out over the gaps between the spindles. One experiences faint redness where there is, veridically speaking, no red. This is a modal perception because there

is substance to the perception in that it looks like the property of a surface. Modal perception applies to all real surface perception not just illusions. For example in the occlusion example of figure 1 the square should be modally perceived, as should the unoccluded portions of the circle.

Amodal perception (Helmholtz, 1876, Michotte 1950, Gibson 1950) is indirect perception in the sense that there is no actual sensory experience (of the object) *per se*, but rather an implicit sense that something is there (see the circle in figure 1 occlusion). In the example depicted you are aware that the circle's surface continues behind the square however for the portion of the circle that is occluded you have no perceptual experience of the red surface. In this case you have an amodal perception of the occluded parts of the circle.

Phenomenal consciousness denotes the phenomenological aspect of consciousness – *actually experiencing something as having a particular quality* (e.g. Block 1990). This is the aspect that many philosophers have defended, seeing it unexplained by theories of the relationship between the brain and the mind (which address function rather than the 'what it is like-ness'). This term does not refer to cognitive access to the phenomenological state

Access consciousness denotes the access to phenomenal consciousness that leaves it poised for voluntary action, reasoning and thinking (Block 1990,1992, 1995).

Consciousness is taken as a general cluster of concepts concerning our subjective experience of the world and access to it – including 'phenomenal' and 'access' consciousness but not corresponding to a strict combination of the two. As a word it has no precise meaning other than the folk-psychology definition present in everyday language.

Reportability is the ability to generate reports based on ones perception. This should strictly be called perceptual reportability, since one could have emotional reportability for instance, which would constitute reporting ones emotions (for some, this has primacy over all other types of reportability). For the purposes of this thesis reportability generally refers to perceptual reportability. The concept of a report can

be decomposed into two components. Consider the following example: A subject is instructed to wave their right hand when they see a face on a screen.

- a) *Meaning of the report.* The subject waving their hand *means* they report seeing the face. Waving their hand has the same meaning as 'I saw a face'.
- b) *Motor act of reporting.* The subject waving their hand is achieved through motor processes of the brain leading to the actual behaviour of the moving hand

Report. For the purposes of this thesis, when talking about brain processes subserving the report I specifically mean the motor act of reporting ((b) above).

Perceptual decision. This is the term used to refer to the meaning of the report ((a) above). The subject forms a decision based upon their perception and this is obviously intrinsically related to the *meaning of the report* if they are reporting the decision. One can change the *motor act of reporting* without changing the actual *meaning of the report*. For example one could instruct the subject to wave their foot when they see the face and it would still have the same meaning as 'I saw a face'.

Object is a category of visual entities not just limited to inanimate objects in the world but in this context can mean semantic entities such as faces and houses (such as in experiment 1) or more abstract entities such as luminance defined discs (in experiment 2)

Perceptual-reporting system is the neural system supporting the perceptual processing and all cognitive-intentional-motoric operations underlying the behavioural report of a percept. It therefore subserves all of the defined concepts above from perception, perceptual decision and the motor act or reporting

The essential dynamic is the dynamic of neural activity that is posited to be most necessary and minimally sufficient (for clear discussion see Chalmers 1998, 2004) in determining or causing consciousness. Examples include:

Recurrent interactions (Lamme 2003)
Thalamo-cortical re-entry (Edelman 2001)
Synchrony (Engel and Singer 2001)
Information (Chalmers 1996)
Ignition (Deheane and Naccache 2001)

Recurrent processing/interactions generally is used to refer to processing where input is fed forward from one area to another that is then fed back onto the same area.

Signal detection profile is a profile of brain responses corresponding to each of the 4 different signal detection categories (hit, miss, false alarm and correct rejection). Larger profiles can be constructed by considering more than one condition (and are still referred to as signal detection profiles).

E. General Overview

This thesis is concerned with the neural correlates of visual consciousness. In plain English this is about understanding what it is our brain is doing that allows us to see things. Inevitably, this is a difficult problem that spans many different disciplines, scales, and methodologies. The approach that this PhD project has taken is to study the neural basis of visual consciousness using functional Magnetic Resonance Imaging (fMRI), a technique that can measure the brain activity of human subjects who are viewing and responding to controlled visual events (for example on a computer screen). This technique allows one to measure a blood oxygenation signal coupled to neural activity across the whole brain at a resolution of seconds and millimetres. Even at this rather coarse scale, there are many different problems that need addressing before one could claim even a limited understanding of visual consciousness. The following thesis constitutes an attack on just two such problems, with one experiment devoted to each. The thesis therefore consists of two parts that could be taken separately, although both address the different aspects of the same fundamental question and fit into a coherent theoretical framework.

F. Overview of experiment 1

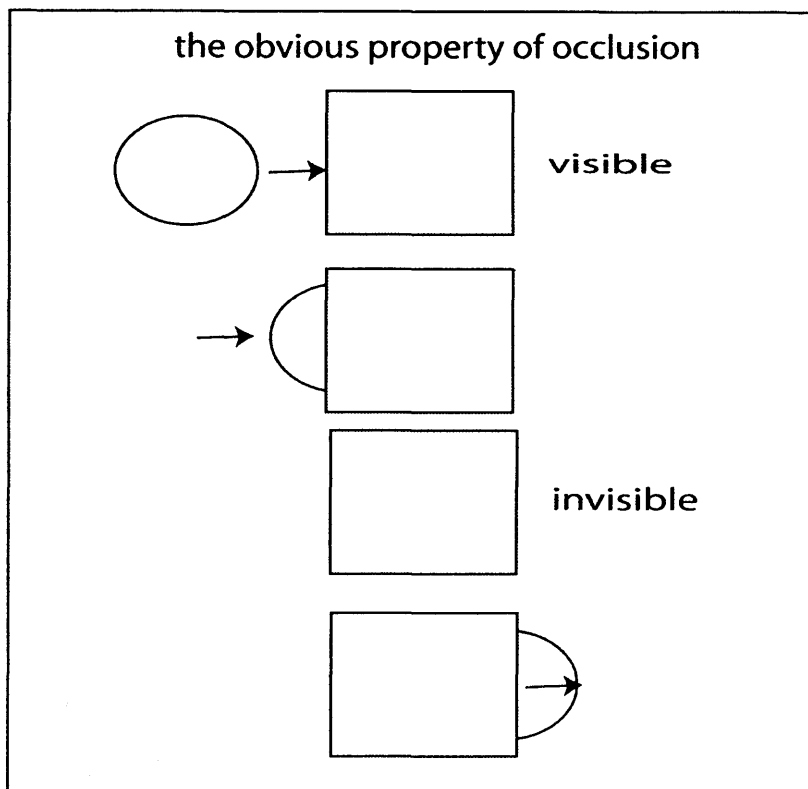


Figure 2. The cartoon depicts the passing of a circle behind an opaque square rendering it invisible. The past history of having seen the circle prior to its occlusion normally results in object permanence, whereby an awareness of the objects presence persists in the absence of immediate sensory input (from the circle).

In a world of perpetual flux, objects and surfaces continually pass in front of one another. In doing so they pass in and out of sight. The reason is as simple as it is obvious. If light is reflected from the object to the eye then it is visible, but if light is obstructed by a surface in-between the object and the eye then the object is invisible. This fundamental property of the visual world, however simple and obvious, has the potential to help further our understanding of the neural basis of object perception.

For species dependent on vision, occlusion is problematic because it means that the availability of visual information is at the mercy of erratic and unpredictable surfaces that are constantly moving with respect to the viewer. This is of particular concern in the natural world because the actual presence of objects such as predators, prey and

conspicuous is independent of whether they are occluded or not. Any species that can process the presence of these objects in a way that is robust against the effects of occlusion is at an obvious advantage to those that cannot. Indeed there is considerable evidence that humans (Sekuler *et al* 1994; Behrmann *et al* 1998; Kellman *et al* 1998; Rauschenberger and Yantis, 2001; Michotte 1950, Scholl and Pylyshyn 1999, Yi *et al* 2003 Joseph and Nakayama K 1999) including infants (Johnson *et al* 2004, Mareschal 2000, Baillargeon 1985, 1986) and many animal species (Churchland *et al* 2003) have evolved brains capable of solving the problems posed by visual occlusion.

The capacity of humans to process occluded objects is of particular relevance to investigations into visual perception. If you look at an object that has become occluded, you perceive only the surface of the occluder and yet you are also aware of the occluded object's location, its spatial extent and its identity. Original accounts of occluded objects described this phenomenon as being 'visually amodal' (Michotte 1950). This intuitively contradictory description has a rather elegant logic. On the one hand the occluded object is 'amodal' in that there is no local sensory experience of its attributes, and on the other hand it is visual in modality, in the sense that it is determined by visual information (Palmer 1999).

This property of being aware of an object's presence despite having no direct perception produces a dissociation between perception and 'awareness of presence' which is very useful for investigating the neural basis of perception. Previous studies have shown that areas that are selectively activated by the visual presentation of faces (Kanwisher *et al* 1997, McCarthy *et al* 1997) or houses (Epstein and Kanwisher 1998 Aguirre *et al* 1998) are the same ones that correlate with the perception of faces and houses (Moutoussis and Zeki 2002, Tong *et al* 2002, Grill-Spector *et al* 2004). One possible interpretation is that activity in these areas must pass a certain threshold for perception of faces and houses to occur. This hypothesis is supported for motion processing in patient GY (who has damage to V1 resulting in a blind hemifield). When GY is presented with moving stimuli in his blind field, his perception of motion correlates best with activity in V5 such that the activation of V5 activity is higher in when motion is perceived than when it is not (Zeki and ffytche 1998). A general term for this principle is the "activation strength hypothesis" (Kanwisher 2001).

The problem for experimental investigations of the ‘activation strength hypothesis’ is that whenever one perceives an object (face or house), one is obviously also aware that it is *present*. Thus in normal conditions perception and awareness of presence are inextricably coupled. The dissociation provided by occlusion circumvents this problem and allows us to ask ‘Do object selective areas correlate with the perception of an object *per se* or with the awareness of an objects’ *presence*?’

We used occlusion in this way, enabling us to directly address the ‘activation strength hypothesis’ in this context of face and house perception. We presented subjects with two types of object (images of faces or houses), which were positioned either in front of (‘visible’) or behind (‘hidden’) an upwardly moving opaque screen. Functional Magnetic Resonance Imaging was used to measure the brain activity under these two conditions, allowing us to distinguish activity reflecting visible objects from that reflecting hidden objects alone. If the response of the object selective areas is due to the direct perception of the object, then its response to a hidden object will be attenuated. Alternatively, if the object selective response is due to the awareness of an object’s presence independently of the perception of its surface, then the object-selective area will respond with equal strength whether the object is visible or hidden.

G. Overview of experiment 2: Reportability

The second problem I sought to address is that of reportability. ‘Reportability’ (a word which doesn’t actually appear in any dictionary) is defined in this context as the ability to generate a report about a particular perception, through any volitional communicative action. The report therefore doesn’t inherently depend on the modality of report for instance a perception could be reported in any arbitrary way for example waving ones foot or hand. Understanding reportability is critical, because visual consciousness - like consciousness of any modality - is fundamentally private in nature in the sense that one person cannot directly access another person’s experience. This poses a fundamental problem for those wishing to scientifically study its physical basis. How does an experimenter know what brain state correlates with which percept? The strategy that most investigators use is to make assumptions (otherwise known as bridging principles (Chalmers 2004) about what objectively observable indices are associated with the existence of a particular perceptual state. The most common assumption is that if a subject can generate a voluntary report about percept ‘X’ (either verbally or by manual button press) then the subject is conscious of ‘X’. By measuring the correlation between *X* and its pre- or post- ceding brain states one has a handle on the neural correlates of percept *X*. Thus the bridging principle utilised in such experiments (most often implicitly), is that of reportability, whereby the ability or inability of the subject to voluntarily report their perception of a stimulus, operationally classifies those stimuli as being conscious or unconsciously processed. This approach has been widespread, with most experiments employing a strategy best characterised by the maxim ‘*keep the stimulus constant change the percept*’ (for a canonical example see Logothetis et al 1989). This paradigm, however simple, has the elegant property of dissociating the neural correlates of stimulus processing from those of perceptual processing. Examples include binocular rivalry (Logothetis et al 1989, Lumer and Rees 1998, Tong et al 1998), bistable figures (Kleinschmidt et al 1998), thresholds tasks (Rees and Heeger 2003), masking (Haynes and Rees 2005) dichoptic colour fusion (Moutoussis and Zeki 2002) and the attentional blink (Sergent et al 2005), whereby the same stimulus can result in two or more conscious states.

These studies often show correlates in Fronto-Parietal and posterior visual cortices and there is considerable disagreement about which of these correlates with conscious visual perception.

Of central concern, however, is the interpretational vulnerability due to the bridging principle that these experiments adopt. As has been pointed out previously (Chalmers 2004), if scientists choose a bridging principle to define consciousness then they will only ever discover the neural correlates of the index that the bridging principle adopts. Thus the operational definition of consciousness is critical in interpreting experimental results and yet this definition is rarely discussed in scientific papers on consciousness let alone explicitly studied. The difficulty inherent in all of these experiments stems from their (often implicit) adherence to the assumptions of reportability. The major problem is that whilst experimenters go to great lengths to dissociate stimulus from percept they rarely attempt to dissociate the operations underlying reportability from the perceptions themselves. The problem then faced is how is it possible to dissociate the correlates of perception from the correlates of the reportability that provides the evidence for the existence of the percept? The problem faced is that many experimenters conduct experiments which make assumptions about reportability in the absence of any framework for understanding its neural basis. This therefore motivates an investigation into reportability itself without any prior assumptions about its relationship to consciousness. Once an understanding of reportability itself is possible then it is appropriate to re-examine the relationship between reportability and consciousness.

To probe the neural correlates of stimulus reportability, we used a paradigm inspired by the classical partial report experiments of Sperling (1960) (see figure 3). In the original experiment subjects were presented with a 3x4 array of letters after which they were cued to report the identity of the letters. If cued to report all letters (full report), only 4 on average could be reported correctly. However if cued 400ms after the offset of the letters to report a randomly chosen row of letters (partial report) subjects could accurately report all 4 letters. Cueing later than 400ms resulted in significantly reduced performance. This has been taken as evidence for subjects having access to an iconic memory trace (Neisser 1967) for a limited period, which

can be accessed after offset and reported only if cued. In this sense, the letters are reportable, but only reported if cued.

We adapted this partial report paradigm replacing letters with a simple luminance-based object detection task which allows one (by virtue of retinotopic topography), to generate a clear hypothesis for the stimulus driven areas. With this paradigm we can address two issues. The first issue is that a partial report necessarily involves some stimulus locations not being reported. The fact that if stimuli are cued at the right time they are reported with high accuracy means that they are still potentially reportable even if they are not reported. With this paradigm we are able to ask what is the difference in neural activity between unreported (but reportable) and reported stimuli? The second issue concerns a full delineation of the neural stages between stimulus onset, and the motor output of the report. A parsimonious hypothesis for the object detection partial report task involves four stages, stimulus processing, attentional selection, perceptual decision and motoric report. Experimental manipulations allow us to decouple these stages and find the neural correlates of each. Firstly, by varying the timing of the cue it is possible to manipulate the performance such that at the short delays, report is coupled to stimulus presence, but at longer delays the report is decoupled from stimulus presence. This allows us to decouple the neural correlates of stimulus presence from those of report. Secondly by manipulating the location of attentional selection, it is possible to measure the neural correlates of attentional selection mechanisms, selecting stimuli to be decided upon and reported. Thirdly by manipulating the mapping of the perceptual decision into the motor act of reporting, it is possible to decouple the neural correlates of the decision from those of the report.

With this design it is therefore possible to decouple 4 neural stages of stimulus reportability, stimulus processing, attentional selection, perceptual decision and report. A stimulus processing area will correlate with the presence of the stimulus at a specific location independently of the perceptual decision ('present vs. absent') or the motor response ('left hand vs. right'). An attentional selection area will correlate with the location of attention independently of stimulus presence, decision and report. A perceptual decision area will correlate with the 'present vs. absent' judgment independently of the stimulus presence, attentional selection or motor report. A motor area will correlate with the motor report independently of all other factors. Once the

G

neural underpinnings of stimulus reportability are outlined, it is then appropriate to discuss their relevance to theories of consciousness and previous experiments.

Consciousness

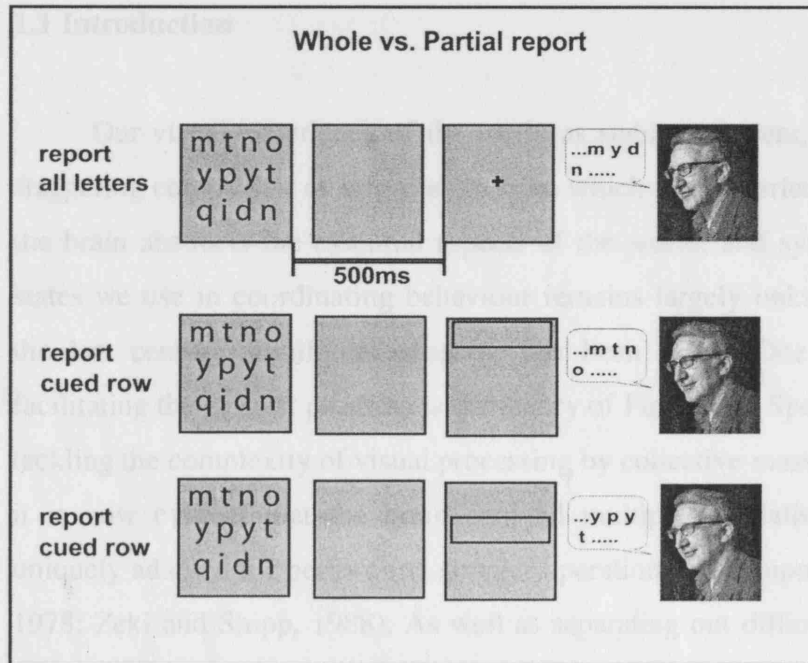


Figure 3. Cartoon depiction of ‘whole report’ and ‘partial report’ experiment of Sperling (Sperling 1960). Time runs from left to right. Each row depicts an experimental trial. The top row depicts ‘whole report’ conditions whereby the subject (in this case the philosopher Ned Block) is instructed to report all of the letters presented in the original array. Under such conditions a random choice of 4 of the 12 letters are reported. The middle and lower rows depict ‘partial report’ conditions whereby the subject reports only the row of letters that is cued after a delay after the offset of the letters. Regardless of where the cue is presented the subject can reliably report all 4 letters. Such an effect was originally demonstrated to be accurate for delays up to 0.5 seconds. It should be noted that the pictured character did not take part in the actual experiment.

1.1. Functional Specialisation of visual cortices

Phrenologists of the 19th century believed that psychological characteristics of the person translated into forms and deformations of the skull. In thinking about the brain phrenologists drew analogy to the muscular system. They assumed different areas of

1. Historical Survey: From Functional Specialisation to the taboo of Consciousness

1.1 Introduction

Our visual experience of the world as stable, coherent, and unified, belies the staggering complexity of information from which that experience is constructed. How the brain abstracts the essential aspects of the world, and synthesises the conscious states we use in coordinating behaviour remains largely unknown. However, within the last century significant progress has been made. One of the core catalysts facilitating the earliest progress is the theory of Functional Specialisation. Rather than tackling the complexity of visual processing by collective mass action (Lashley 1931), it is now evident that the brain evolved multiple specialised regions, which are uniquely adapted for performing simpler operations on components of the task (Zeki, 1978; Zeki and Shipp, 1988). As well as separating out different sensory modalities, Functional Specialisation has been demonstrated and confirmed in virtually all domains of visual processing, from primitive attributes such as colour and motion, to higher-level categories such as faces and houses. Functional Specialisation provides a foundational framework from which even more ambitious questions can be asked. A question central to this thesis is: how does activity in functionally specialised areas of the brain contribute to visual consciousness and how do they interact with the rest of the brain to explain the basic properties of visual consciousness such as modality and reportability.

In the following historical review I will briefly chart the evolution of Functional Specialisation, discussing its conception, its empirical verification and its contemporary application to problems in visual consciousness.

1.2. Functional Specialisation of visual cortices.

Phrenologists of the 19th century believed that psychological characteristics of the person translated into forms and deformations of the skull. In thinking about the brain phrenologists drew analogy to the muscular system. They assumed different areas of

the brain were specialised for different psychological functions, and therefore such areas, if exercised differentially, would have different forms. By this logic psychological characteristics are indicative of increased exercise to that part of the brain and therefore the part of the brain responsible will be enlarged. By studying the skull phrenologists claimed to be indirectly mapping areas of the brain to psychological characteristics. Contemporary science ridicules phrenology as bad science. However it could be argued that a primitive ancestor to the concept of functional localisation was born in phrenology. If this is so then phrenology was vaguely right in its logic, but woefully wrong in method.

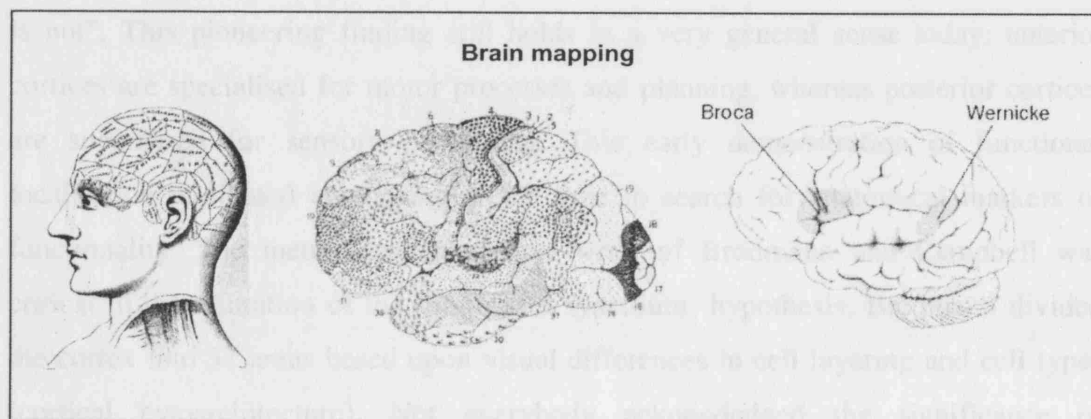


Figure 4. Image on left depicts a phrenologists mapping of functional specialisation with attributes such as ‘trustworthiness’ and ‘tempestuousness’. Middle image depicts the cytoarchitectonics, mapping areas according to cellular composition types and subtypes and laminar characteristics of the cortex. Right image depicts the locations of Broca’s and Wernicke’s areas discovered through the distinct and critical effects damage to such regions has on language abilities. [Illustration taken from Zeki 1973]

Phrenology was by no means the general consensus in the 19th century. Flourens believed that all functions of the brain were intrinsically interwoven in a widely distributed manner, such that no one function could be localised to a single area. The brain according to Flourens was an equipotent ‘syncitium’ delivering a myriad of functions by the collective action of the entire brain. The credibility of this belief was soon to be questioned by Pierre Paul Broca who made the groundbreaking

step of conceptually linking specific damage of the cerebral cortex to a specific behavioural deficit. Broca presented to the world a patient, who following widespread damage to frontal regions (particularly left posterior regions), could understand language but not produce any structured language of his own. “The lesion in the frontal lobe had been the cause of the loss of speech” (Broca 1861). Despite Broca’s revolutionary demonstration of functional localisation, it was not until manipulation of an architecturally distinct region of cortex was shown to precipitate different behavioural manifestations, that functional localisation demonstrated its true power as a scientific theory. Fritsch and Hitzig (1870) observed that stimulating the anterior parts of the cerebral cortex resulted in muscular contractions but the stimulation of the posterior regions did not. “One part of the convexity is motor in function, another part is not”. This pioneering finding still holds in a very general sense today; anterior cortices are specialised for motor processes and planning, whereas posterior cortices are specialised for sensory processing. This early demonstration of functional localisation motivated anatomists of the time to search for anatomical markers of functionality. The meticulous anatomical work of Brodmann and Campbell was critical to the refutation of the ‘brain as a syncytium’ hypothesis. Brodmann divided the cortex into 52 areas based upon visual differences in cell layering and cell types (cortical cytoarchitecture). Not everybody acknowledged the significance of cytoarchitectural differences between areas. Lashley for one, proposed that areas 18, 19 and 17 and they acted altogether as a single unit. This postulation was to form a central principle of his theory of Mass Action. The ‘empirical’ basis of the theory of Mass Action stems largely from one generalisation, that the specific efficacy of a lesion depends more on its size than its location. The observation proved to be true only for complex tasks such as navigational learning, which by virtue of its complexity is necessarily dependent on many different areas of the brain such that other areas can compensate for damage to any one subsystem.

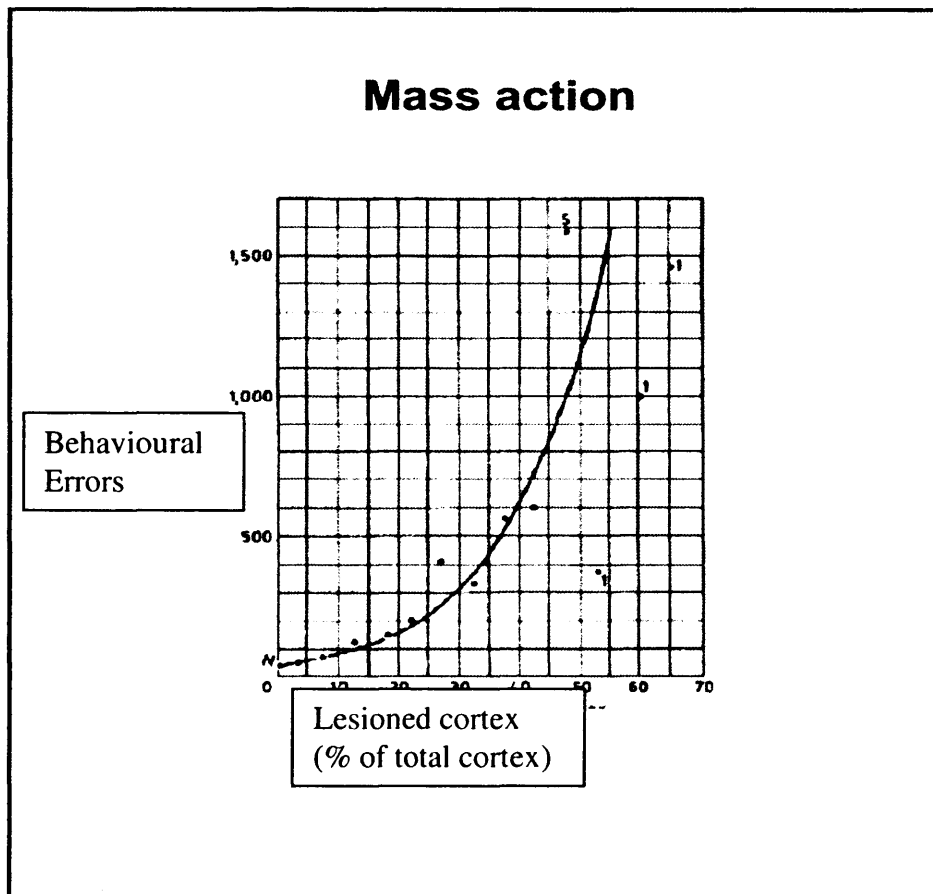


Figure 5. Graph depicting an exponential relationship between the size of any lesion anywhere in the brain (expressed as a proportion of total cortex volume) and the number of behavioural errors made. [Graph reproduced from Lashley 1958].

Early 19th century neurologists were greatly influenced by the speculations of both Lissauer and Munk. Munk had shown that following cortical lesions to the cortex of the dog, the dog retained sight but could not recognise objects. Based on Munk's observations Lissauer (mainly based on his patient study) speculated on the duality of visual processing. The first consisted of reception of 'visual impressions', which he termed 'apperception' (perception; a term borrowed from Leibniz) in which consciousness accepts superficial appearance. The second consisted of connecting concepts by the act of association, by which sensory impressions gained 'meaning'. Studies of patients with visual field defects from scotomas and hemianopias (Holmes 1945) revealed that the critical site of damage was close to or within the calcarine sulcus of the occipital lobe, leading Holmes to posit support for striate cortex being 'perceptive' visual cortex and surrounding regions being association cortex.

Although the body of evidence for functional localisation by this time was significant it was not until the work of the neurosurgeon Wilder Penfield in the 1950's that functional localisation approached its irrefragable status. Penfield took the opportunity to electrically stimulate the brains of epileptic patients during surgery. Consistent with the aforementioned evidence, the effect the electrical stimulation had on behavioural or perceptual synthesis systematically varied with its location. For example, stimulation of the middle temporal lobe invoked vivid memories whereas the same stimulation on the post-central gyrus invoked the perception of touch and pressure. This powerfully demonstrates the genesis of two different experiences resulting from stimulation at two different locations in the brain.

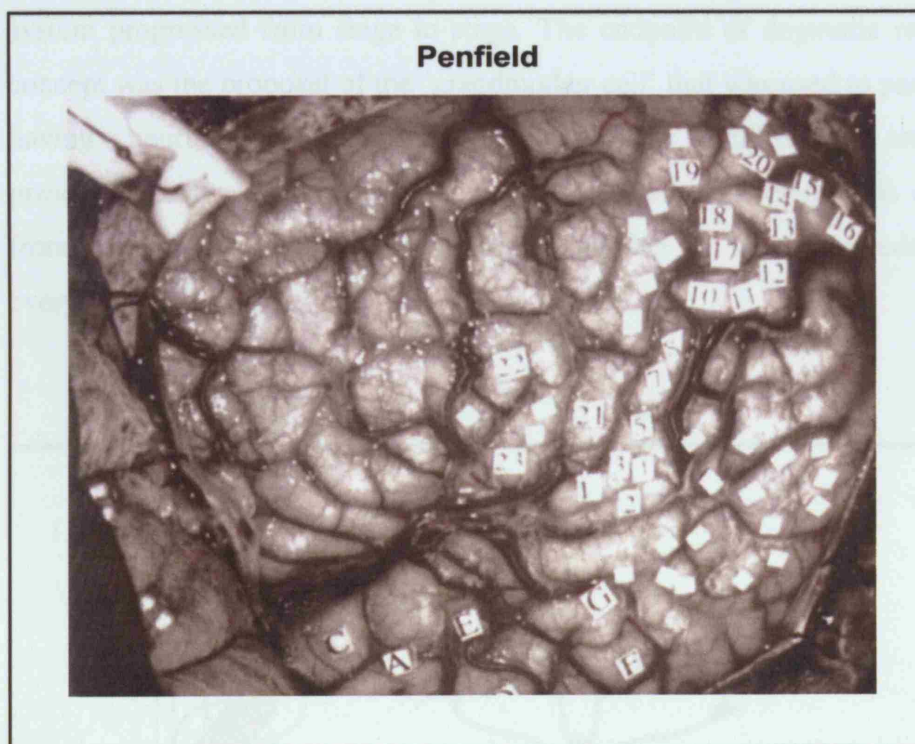


Figure 6. Famous diagram displaying the stimulation sites Penfield used in his seminal but controversial direct stimulation experiments on awake behaving humans. A common joke used by certain neuroscientists displaying this slide to the public goes along the lines of ‘unfortunately the brain does not come with these labels!’ This promotes the belief neuroscientists are not particularly funny. [Picture from Penfield 1968]

By the late 1950's technology had advanced to a state whereby scientists could begin to chart the specific neuronal properties of animals in response to stimulus events in the outside world. The first scientists to do this in visual cortex were Hubel and Wiesel, first in the cat (Hubel and Wiesel 1959) and then in the monkey (Hubel and Wiesel 1968). From their recordings, Hubel and Wiesel were able to map receptive fields of cells (the window of space in which a stimulus effects the firing of a cell) at different anatomical locations within striate cortex. They reported a number of different cell types classed according to their receptive field properties that appeared to embody a hierarchical organisation. Hubel and Wiesel's work had profound influence on research into visual processing that followed. Central to this influence was the notion of increasing complexity in response selectivity as the visual system progressed from stage to stage. The endpoint of dogmatic recursion of this concept was the proposal of the 'grandmother cell' that was used to parody the idea of having a neuron so complex in its selectivity that it would respond only to the visual presentation of your own grandmother. Among other problems this concept suffers from the combinatorial explosion in the number of neurons needed to code for everything you ever saw.

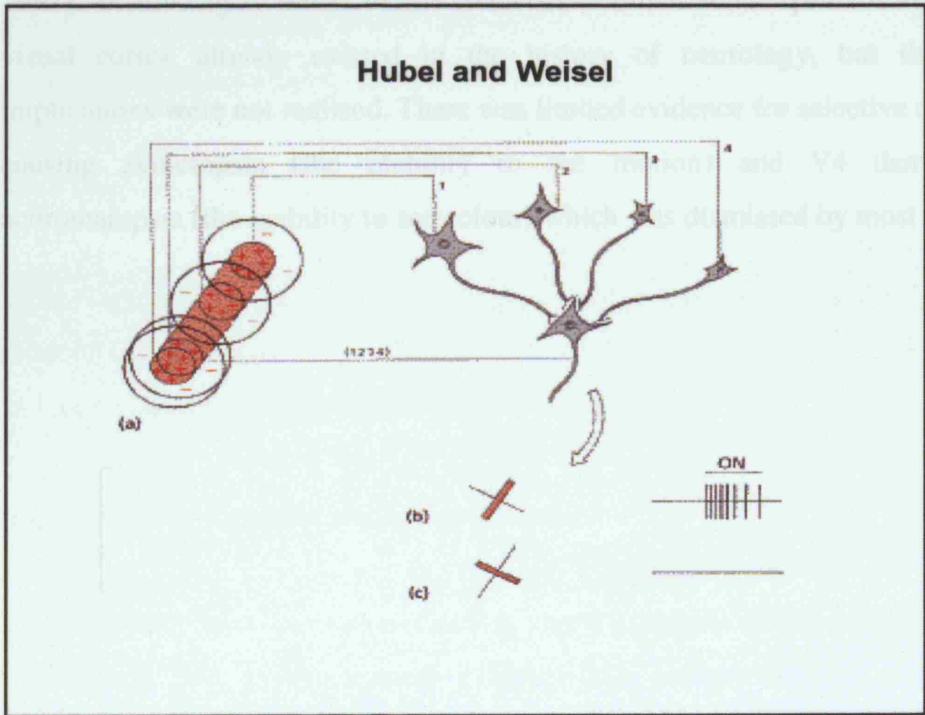


Figure 7. A textbook illustration of the receptive field properties of a V1 simple cell. Officially the story goes that orientation selective cells were serendipitously discovered through the sliding of a cracked filter lens into a projector, causing the cell to fire vigorously. [Illustration taken from Zeki 1993]

The foundation of functional specialisation in the visual brain is derived from the anatomical and physiological work of Zeki, which showed that the response properties of visual neurons in pre-striate cortex varied accordance to the location from which he recorded and that there are many visual areas beyond V1 with different properties. Neurons on the posterior bank of the STS contained neurons selective for motion but not for colour (Zeki 1974). Zeki found another region of cortex in the prelunate sulcus containing neurons which were polar opposite in their selectivities, responding to colour but not to motion (Zeki 1973, Zeki 1978). This discovery in combination to anatomical investigations, lead to a proposition that structures all thought about the visual brain, that the visual system is organised into multiple, parallel, functionally specialised systems. A decade later, Zeki was able to demonstrate using Positron Emission Tomography that the human brain did possess functionally specialised areas for the processing of both colour and motion and that these were homologous to the areas he had discovered in the monkey (Zeki et al 1991). In hindsight, much of the evidence for functional specialisation of human visual cortex already existed in the history of neurology, but the theoretical implications were not realised. There was limited evidence for selective damage to V5 causing akinetopsia (the inability to see motion) and V4 damage causing achromatopsia (the inability to see colour) which was dismissed by most at the time.

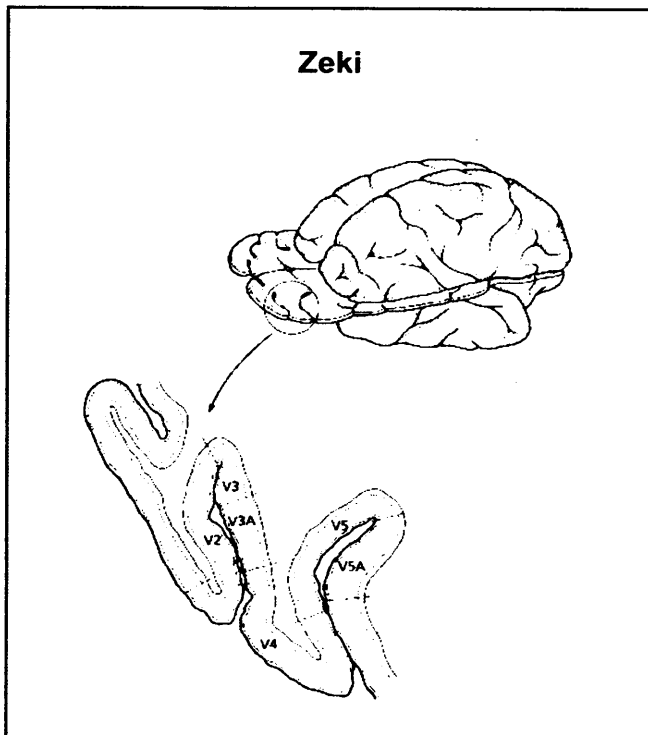


Figure 8. Functionally specialised visual areas seen from a single anatomical slice through posterior cortices. [Illustration taken from Zeki 1993]

1.3 Higher level vision

Having already reviewed the historical evidence for the evolution of the theory of Functional Specialisation I will now review the evidence for its application to the domain of form processing, specifically faces, objects and houses.

1.3.1 Faces and Objects

Extensive evidence for a functionally specialised processing system for faces has been accumulating since the late 1940's, including behavioural studies, functional imaging, and direct recordings of human and non-human primates.

Behavioural and psychophysical evidence

There has long been speculation that faces have a 'special status' in human behaviour. Behavioural and psychophysical evidence has accumulated in support of this speculation. From the moment most primates are born they express a processing bias

for faces. Human neonates visually track faces further and for longer than any other visual object (Goren et al 1975), and discriminate and imitate facial expressions (Field et al 1982). Compared with object recognition, facial recognition has been shown to differ systematically along multiple dimensions. For example, face processing is more sensitive to contrast polarity, direction of illumination, rotation in depth and many other variable image dimensions (Tanaka et al 1993).

Neurological evidence

The first brain-derived evidence came from neurological investigation of patients with focal brain damage. In 1947, Bodamer observed a patient who had a selective impairment in his ability to recognise faces, whilst retaining the ability to recognise objects (Bodamer et al 1947). Bodamer named this syndrome prosopagnosia. Prosopagnosics are classically able to detect the presence of faces whilst being unable to identify familiar faces, or learn new ones. Lesions that cause prosopagnosia are usually found in ventral occipito-temporal cortex, involving the lingual and fusiform gyrus, and are mainly bilateral (Damasio et al 1982). However, some studies suggest that a right unilateral lesion is sufficient for the deficit (Landis et al 1988). Prosopagnosia is one example of a cluster of deficits known as 'object agnosias', in which the ability to recognise objects is damaged in specific and often idiosyncratic ways. Object agnosias arise from damage to inferior temporal cortex and fusiform gyri (Goodale et al 1991). Patients with object agnosia often also display some prosopagnosic symptoms. However, this is likely to be due to the fact accidental lesions rarely respect functional or anatomical boundaries, and does not necessarily imply that both objects and faces depend on the same 'module'. In support of this conjecture, there are patients who present a 'double dissociation' (Shallice et al 2003) between face and object agnosia. For example, patient P.S shows pure prosopagnosia, and patient C.K, who was severely impaired at object recognition, leaving face perception and recognition totally preserved.

Neurophysiology in monkeys

The most direct demonstration of specialised face processing comes from single cell recordings in the Macaque. Whilst recording from the temporal visual area of macaque cortex, Charlie Gross and colleagues found neurons that appeared to respond best to complex visual stimuli such as faces (Desimone and Gross 1979). It was soon found that while some of these neurons could respond to parts of the face, other neurons required several parts of the face to be in the right spatial arrangement. In addition, many of these neurons did not just respond to 'a face', but responded differently according to the identity of the face (Desimone et al 1984). The inferior temporal visual cortex (area TE) where such face selective cells were found can be divided according to cytoarchitecture, myeloarchitecture and afferent input. Within TE neurons responsive primarily to faces are found most frequently in areas TEa and TEm and within the STS in area TPO where they comprise up to 20% of neurons responsive to stationary stimuli. Even within this face-processing network, functional specialisation is evident. Within the STS, clusters of neurons have been found with distinct selectivities, some responding to facial expression (Hasselmo et al 1989), and others selective for direction of gaze. Neurons that are selective for facial identity are primarily found in the Inferior Temporal Gyrus (Bruce et al 1981). Single cell electrophysiology in the macaque has demonstrated neurons tuned to complex object features in inferior temporal cortex (Miyashita 1993). It is these areas that are thought to be the monkey homologue of those areas damaged in the object agnostic human.

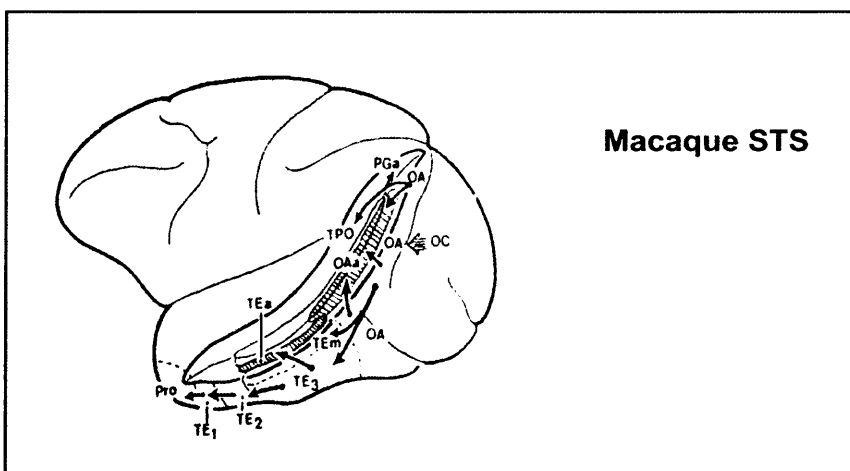


Figure 9. Anatomical diagram depicting the superior temporal sulcus of the macaque opened to reveal intra sulcal sub-regions. Typically cells in TE (1-3)

and TEa, TEm are associated with face and object processing. [Illustration from Rolls 1984]

Brain imaging

The first brain imaging study on high-level visual form perception (Haxby et al 1991) delineated an anatomical distinction between low-level form perception (dot matching) and face perception (face matching). The former activated lateral occipito-temporal cortex, and the latter activated more anterior and ventral regions of the fusiform gyrus. Since then a battery of fMRI studies have identified a number of small areas lying outside early visual cortex that respond more to faces than to objects. Of these, the most robust difference in activity has been located in the middle fusiform gyrus, bilaterally but significantly stronger in the right hemisphere (Haxby et al 1999). This is the region that is claimed to be an essential processing node for face perception and consequently has been named the 'fusiform face area' or FFA (Kanwisher et al 1997). Posterior to the FFA, a region of the inferior occipital gyrus known as the Occipital Face Area or OFA, also responds more to faces than objects (Haxby et al 1999, Gauthier et al 2000). The anatomical borders of striate and extrastriate visual regions are defined in terms of the boundaries of retinotopic maps (areas which contain a topographic representation of visual space employing a reference frame centred on the retinal image); one corresponding to each area. However, the FFA and OFA are not obviously retinotopic in this way (whether they display very crude retinotopy is another question), and so the exact anatomical borders of these regions are harder to define. Their localisation is therefore often based upon functional criteria such as greater responses to faces than houses or objects. Many other methodologies used to study the neural correlates of human face processing such as PET (Haxby et al 1994), intracranial EEG (McCarthy et al 1999) and depth electrode studies all confirm the properties, time course and localisation of face selective responses in the FFA.

Early studies of object processing defined a specific region of lateral occipito-temporal cortex entitled the lateral occipital cortex (LOC) which is more responsive to shape and object form than to texture (Malach et al 1995, Kanwisher et al 1996). The LOC is a large area running from just posterior to V5, antero-ventrally down the

lateral surface of the brain onto the ventral surface and anteriorly along the ventral surface up to and including the fusiform gyrus. The LOC contains neurons that are inferred to be invariant to the size, translation and viewpoint of the object (Grill-Spector et al 1999)

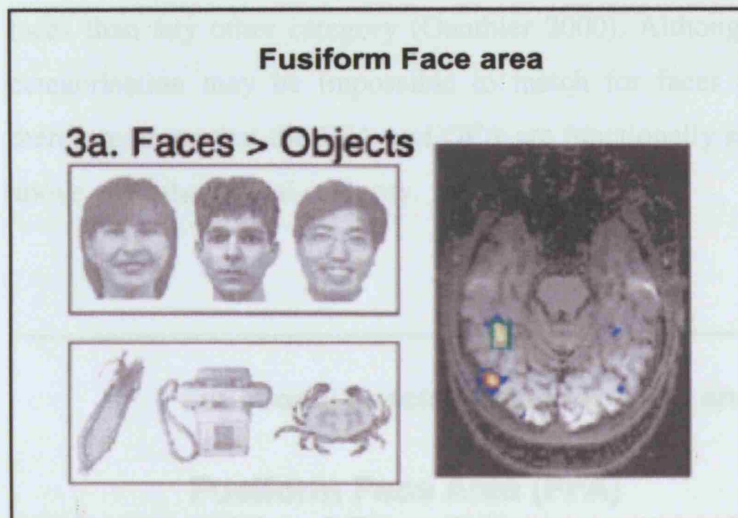


Figure 10. Shows the anatomical location of the first functional delineation of the of the Fusiform Face Area. This was achieved with simple presentation of face and object images. Image depicts regions of the brain which are significantly more active to the presentation of faces than to the presentation of objects. [Figure from Kanwisher et al 1996]

Arguments against Functional Specialisation of FFA to faces

Despite the evidence pertaining to the distinct processing differences between faces and objects of the FFA, there is still considerable debate as to whether the distinction exists at all. Several lines of evidence converge to suggest that the difference in the level of categorisation and expertise in discriminating between faces and objects can largely account for the apparent functional differences in the face and object processing systems. Firstly, if the subjects have expertise in the discrimination of particular objects (i.e. bird spotters) then the qualitative and systematic psychophysical differences between faces and objects drops out (Gauthier et al 1997).

Secondly, non-face objects elicit more FFA activity when matched to specific labels as compared to more categorical labels e.g. ‘ketchup bottle’ as opposed to ‘bottle’ (Gauthier et al 1997). Thirdly, expertise with animal-like unfamiliar objects (‘Greebles’) recruits the right FFA (Gauthier et al 1999). However it still holds that when comparing expertise effects for birds, cars and faces, the FFA is more active for faces than any other category (Gauthier 2000). Although the level of expertise and categorisation may be impossible to match for faces vs. other domains, I would therefore argue that the FFA and OFA are functionally specialised for faces over and above any other visual category.

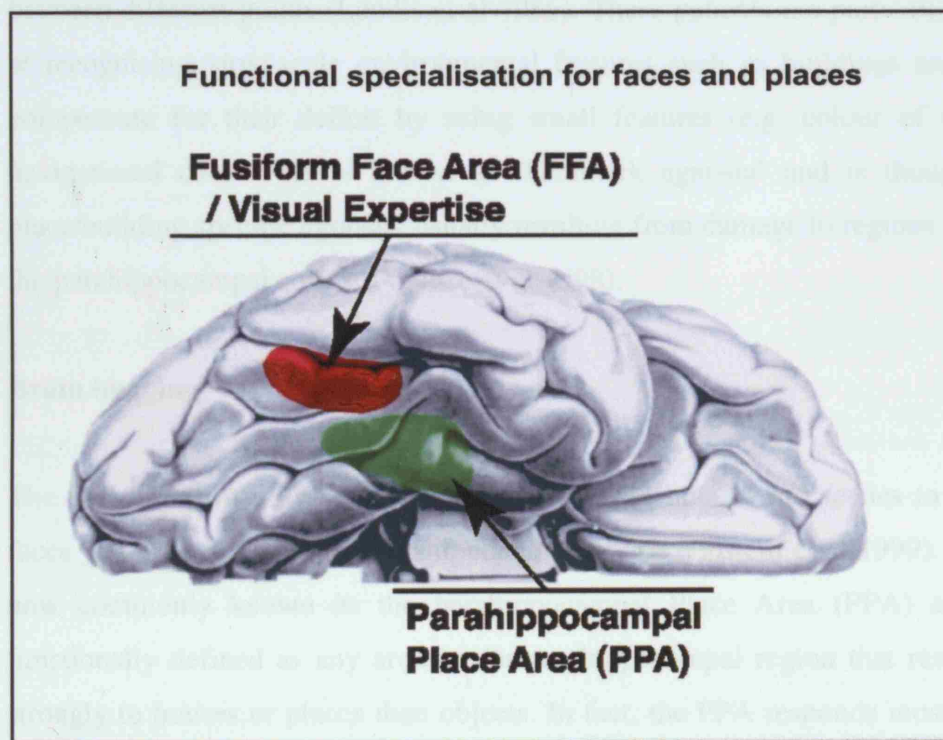


Figure 11. Illustration giving approximate location of FFA and PPA. It should be noted that the location of these areas is not defined by functional localisers employing retinotopic mapping (as is the case for areas like V1 or V2) but by functional localisers contrasting responses between object categories. For example, a common criteria for the FFA is an area of the fusiform gyrus which responds twice as much to faces than any other object [figure from Kanwisher 2000].

1.4 Places and houses

In contrast to faces and objects, there has been much less empirical evidence concerning a functionally specialised processing system for places and houses

Neuropsychological evidence

The first indication of a functionally specialised system for places and houses came from a number of brain damaged patients who lost their ability to recognise places, despite retaining the ability to follow maps and understand the spatial relationships between different points (Landis et al 1986). These patients are particularly impaired at recognising large-scale environmental features such as buildings and can often compensate for their deficit by using small features (e.g. colour of mailbox) as navigational cues. This is known as 'landmark agnosia' and is thought to be a place/building specific agnosia, usually resulting from damage to regions in or around the parahippocampal cortex (Aguirre et al 1998).

Brain imaging

The first brain imaging study to directly compare houses and scenes to objects and faces yielded activity in the Parahippocampal gyrus (Epstein et al 1999). This area is now commonly known as the parahippocampal Place Area (PPA) and is often functionally defined as any area in the parahippocampal region that responds more strongly to houses or places than objects. In fact, the PPA responds most strongly to spatial layouts, and slightly less strongly to single houses or places. It responds weakly to single objects and not at all to faces. The critical factor for this activation appears to be the layout of local space. The PPA does respond to house stimuli alone but not quite as strongly as scenes which depict the layout of local space (Moutoussis and Zeki 2002, Tong et al 1998).

1.5 Visual perception and awareness

Binocular rivalry

Some of the first neuroscientific experiments to claim to have achieved a neural correlate of perception were those utilising a perceptual phenomenon of ‘binocular rivalry’. Binocular rivalry is depicted below, and is achieved by presenting one image to each eye independently, where the two images are incongruous. The brain, it appears, cannot fuse these disparate images into one coherent percept. Instead, what happens is that the brain engages in a rivalrous percept, such that the image that the subject sees stochastically alternates with several seconds between each switch. Binocular rivalry has the rather useful property of providing constant stimulus input, yet resulting in a change in perception. This allows one to dissociate the neural activity that correlates with the stimulus from that which correlates with perception. This dissociation logic is now a central paradigm in contemporary studies of consciousness. In the first binocular rivalry study using awake behaving primates, Logothetis showed that the proportion cells correlating with perception rather than stimulus was significantly lower in V1 (11%) (Leopold and Logothetis 1996) than IT (90%) (Sheinberg and Logothetis 1997). This led to the speculation that only the higher stages of the ventral stream correlate with visual consciousness (Milner and Goodale 1995), a contentious issue that will be discussed in detail later. This study inspired human brain imaging experiments on binocular rivalry, presenting images of faces and houses to alternate eyes. These studies showed that the activity of the FFA, associated with processing of faces, correlates with their perception, and similarly for the PPA and ‘house perception’. From this evidence alone it would seem that the principle of functional specialisation extends into perception.

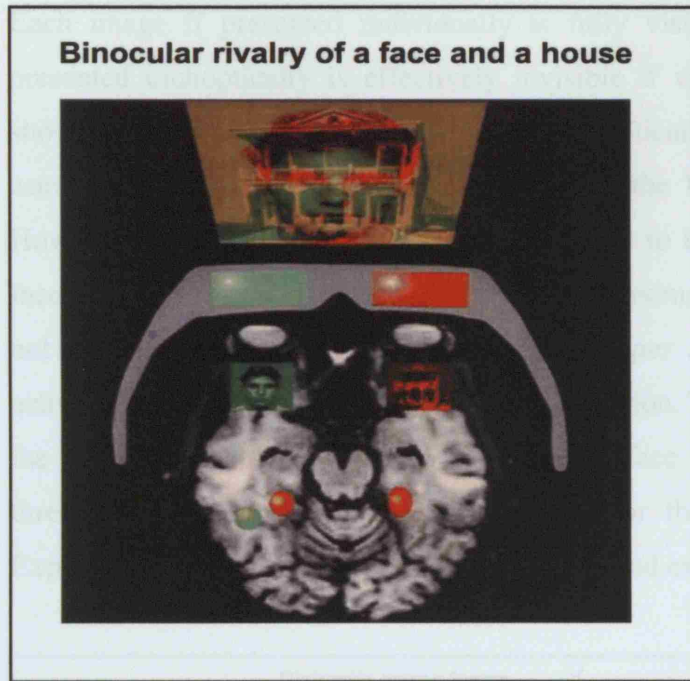


Figure 12. Image depicts the sophisticated stimulus manipulations required to induce binocular rivalry in a scanner. A superimposed image of a face in green and a house in red, when viewed through red-green spectacles delivers each image to each eye independently. The difference between the images cannot be fused by the brain into a coherent percept, resulting in rivalry between the two percepts. Brain activity recorded by the scanner can reveal areas of the brain correlating with each state. [Figure taken from Tong 1998]

Dichoptic colour fusion

The results of the early binocular rivalry studies suggested that the activity of functionally specialised cortical regions correlates with consciousness of those stimuli processed in each region. This was an important but controversial finding, and further investigations were undertaken to strengthen and deepen the interpretation of the result. This was achieved by asking more *specific* questions about the relationship between activity in functionally specialised areas and consciousness. One such question is 'Does *any* activity in the FFA result in a face percept?'. To answer this question Moutoussis and Zeki (2002) used a technique called dichoptic colour fusion. By presenting images of opposite colour contrast to each eye, such as a red face on a green background and a green face on a red background (see figure 13), the resulting percept is that of a homogeneous field of yellow.

Each image if presented individually is fully visible as a face or house, but if presented dichoptically is effectively invisible if viewed binocularly. fMRI data showed that despite being invisible, these dichoptically viewed face and house stimuli caused activity to increase above baseline in the FFA and the PPA respectively. However, the activity of those areas was found to be significantly higher when the face or house was consciously perceived. A parsimonious interpretation is that it is not activity in a functionally specialised area *per se*, but rather the *magnitude* of activity, which correlates with conscious perception. This evidence has further fuelled the speculation that the neural correlate of face perception is activity above a threshold level in the FFA (and similarly for the PPA and house perception). Experiment 1 of this thesis seeks to investigate and expand on this hypothesis.

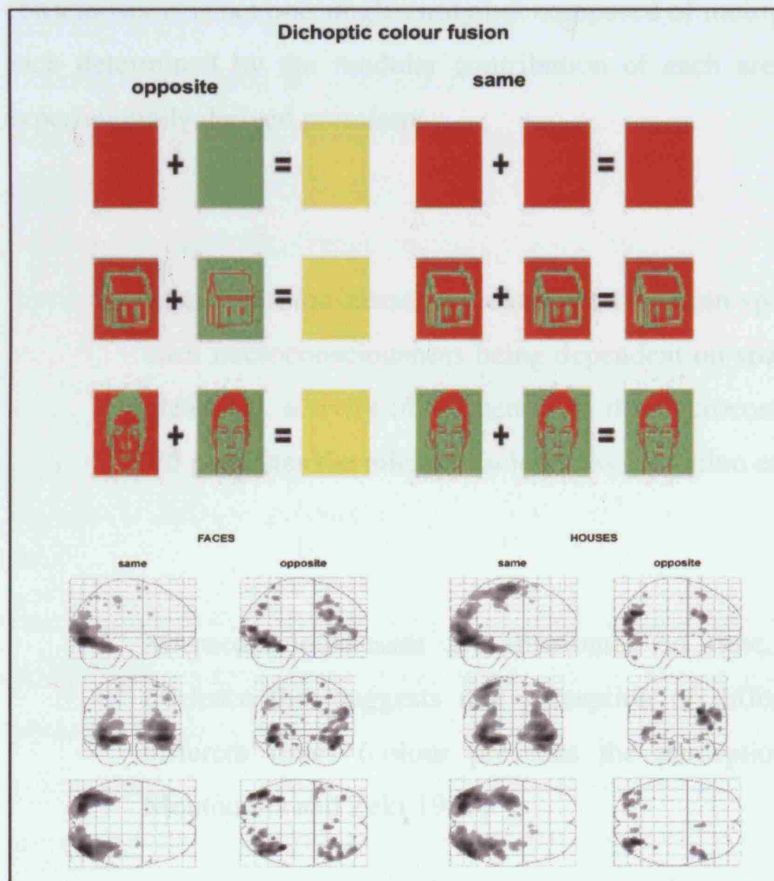


Figure 13. Coloured stimuli depict stimuli presented separately to each eye. If opposite in colour then the resulting fused percept is yellow and if same in colour then the resulting percept is the same as the monocular image. Different stimuli can then result in the same perception, affording the critical dissociation between stimuli and percept. Below are the resultant brain activations for invisible and

visible conditions, showing that overlapping regions of object selective cortex are active whether the object was visible or not. [Figure taken from Moutoussis and Zeki 2002]

2. Theoretical introduction

2.1 Theory of microconsciousness

With functional specialisation established as a principle of brain organisation, and the growing evidence that such areas correlate perceptually with the domain they are demonstrably specialised for, Zeki was motivated to espouse the Theory of Microconsciousness as a natural extension (Zeki and Bartels 2001). The theory of microconsciousness is a modular theory of consciousness that states that consciousness is not one single entity but composed of multiple microconsciousnesses each determined by the modular contribution of each area. It is founded on two experimentally derived premises;

1. Microconsciousnesses are distributed in brain space. This is supported by each microconsciousness being dependent on spatially distinct specialised areas, e.g. activity in V4 generates the microconsciousness of colour and V5 generates the microconsciousness of motion etc.
2. Microconsciousnesses are distributed in time. This is supported by evidence that suggests that perception of different attributes occurs at different times (colour precedes the perception of motion by 80ms; Moutoussis and Zeki 1997)

The Theory of Microconsciousness is a theory concerning the nature of phenomenal consciousness. The position is that activity in visual cortex, if exceeding a threshold,

determines the microconsciousness of the attribute it is specialised for thus causing it to be phenomenally experienced.

2.2 Reportability and consciousness

Arguably one of the most fundamental properties of consciousness is that it is private. Only the subject knows what it is like to perceive what they perceive at any given time. This is an obvious and sizeable obstacle to an empirical investigation into the neural basis of consciousness. The principal problem is how we can measure the quantity of interest – the private experience of the subject. There is no objective access to this quantity that can support a quantitative measure. It has been reasonably argued that the most pragmatic solution is to rely on a subject's report about what they are conscious of, as consciousness is (assumed to be) systematically associated with the ability of the subject to report their perceptual/mental state. Indeed, this property of reportability is so exclusive to conscious information that it is commonly used as an empirical criterion to judge whether a subject was conscious of an information or mental state (Gazzaniga et al 1977, Weiskrantz 1997). In other words, this has resulted in consciousness being operationalised according to whatever can be reported by the subject. This can be characterised by simple logical statements such as;

Report = a therefore percept = a

Report = b therefore percept = b

The problem with this paradigm is that perception doesn't magically leap into a uniquely determined report. There must be a number of attentional, intentional, and motoric operations between perception and the reporting action. These operations will be necessarily be contingent upon the perceptual state but will obviously be subject to influence from many other dimensions. As long as consciousness is operationalised according to reportability, the neural activity that correlates with the percept is likely to correlate equally well with such 'reportability processes'. The problem for neuroscientists then is 'How do we know which is which?' How is it possible to disentangle this convolution of faculties to identify the core process underlying visual perception and its access. Arguably the only way to solve this problem is to dissociate

the components supporting reportability, and in turn investigate the neural underpinnings of each component separately. The second of the two experiments in this thesis is motivated by the quest to dissociate the neural correlates of the component stages of reportability.

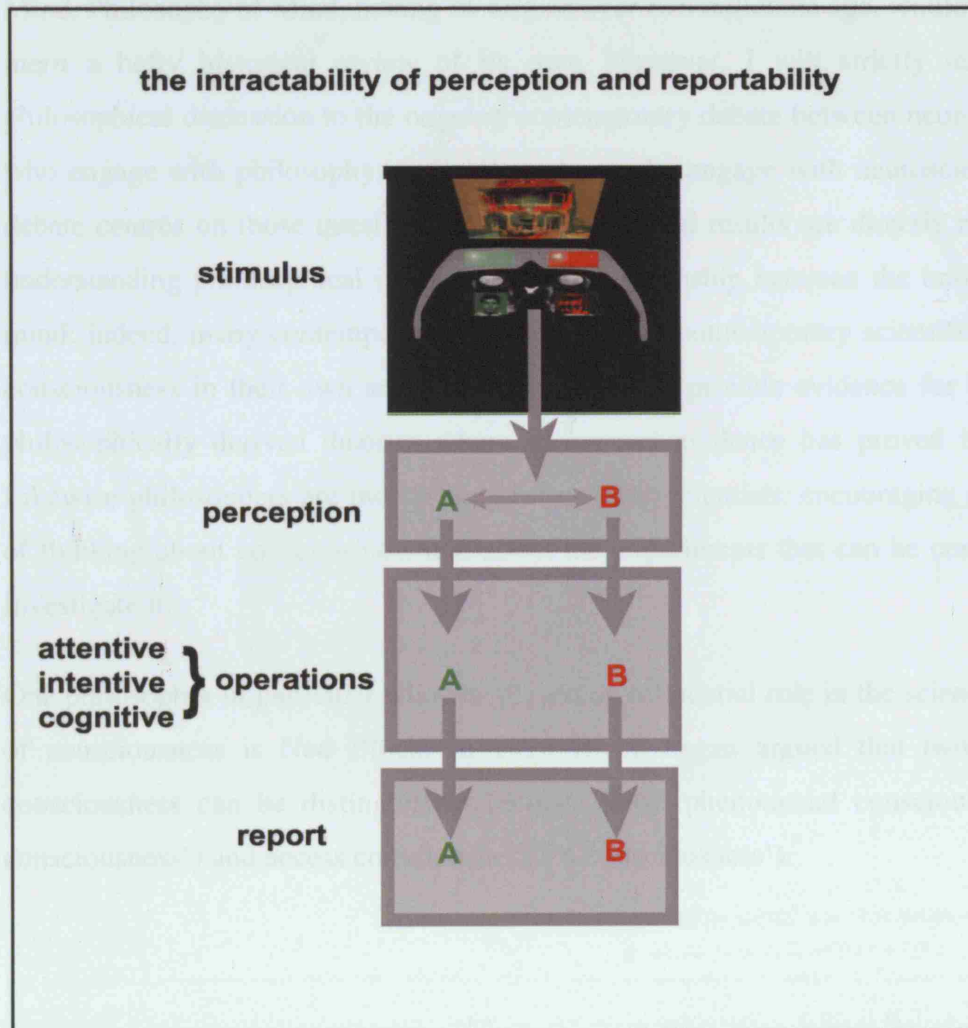


Figure 14. Typical psychological box-and-arrow diagram depicting the multiple operations between perception and the reporting action which is taken as an index of that perception. In experimental design there is often no explicit consideration to the operations that lie between perception and report. Implicitly most experimenters would assume they exist, but that they are constant and reflect with fidelity what is conscious. Such assumptions are rarely if ever made explicit in empirical studies of consciousness.

2.3 Philosophical enquiry

This branch of neuroscience has direct relevance to certain debates in Philosophy of Mind. Philosophy of Mind, having its origins over two millennia ago, would evidently merit a hefty historical review of its own. However, I will strictly restrict any philosophical discussion to the ongoing contemporary debate between neuroscientists who engage with philosophy, and philosophers who engage with neuroscience. This debate centres on those questions where experimental results are directly relevant to understanding philosophical issues about the relationship between the brain and the mind. Indeed, many contemporary philosophers use contemporary scientific work on consciousness in their own argumentation. This can provide evidence for or against philosophically derived theories where theoretical evidence has proved indecisive. Likewise philosophers are increasingly influencing scientists, encouraging new ways of thinking about consciousness and about the experiments that can be conducted to investigate it.

One philosopher in particular who has played an influential role in the scientific study of consciousness is Ned Block. In 1990 Block began argued that two types of consciousness can be distinguished (Block 1990); phenomenal consciousness ('p-consciousness') and access consciousness ('a-consciousness').

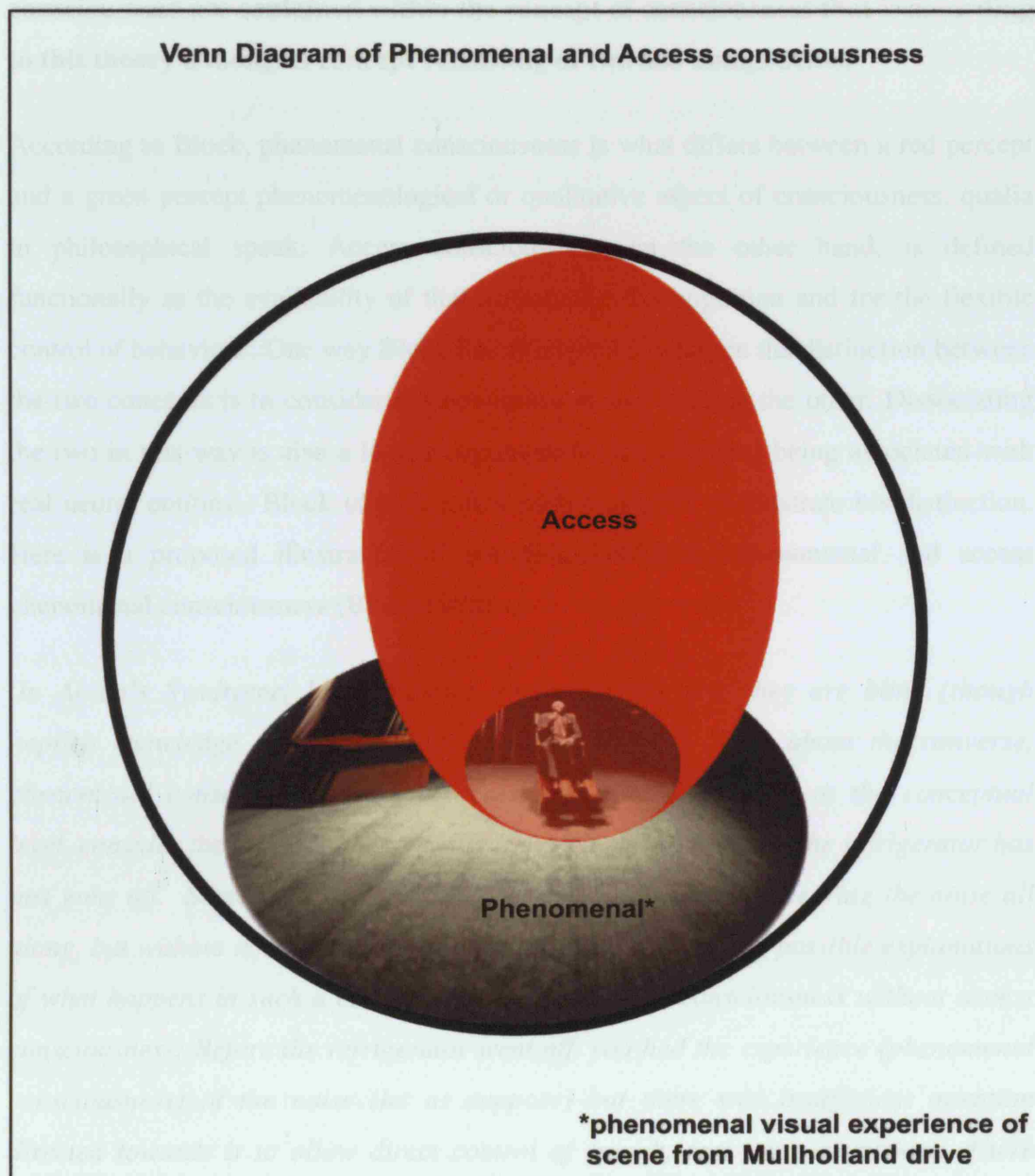


Figure 15. An abstract cartoon Venn diagram of the relationship between phenomenal and access consciousness. Phenomenal visual experience of a scene (from the film *Mulholland Drive*) extends across the visual field. The entire surface (panorama) is part of a simultaneous phenomenal experience. The region of overlap is determined by the location of spatial attention that causes that part of phenomenal experience to be accessed (featural attention can also be used for access which would allow featural rather than spatial access). The information embedded in that location is then available for deeper processing and for storage into working memory such that it can be poised for ‘flexible control of voluntary behaviour and rational thought’. Both p and a-

consciousness are contained within the concept of consciousness that is according to this theory a mongrel concept consisting of two sub-components.

According to Block, phenomenal consciousness is what differs between a red percept and a green percept phenomenological or qualitative aspect of consciousness, qualia in philosophical speak. Access consciousness, on the other hand, is defined functionally as the availability of that information for cognition and for the flexible control of behaviour. One way Block has attempted to sharpen the distinction between the two concepts is to consider the possibility of one without the other. Dissociating the two in this way is also a logical argument for the concepts being associated with real neural entities. Block utilizes a few such examples to illustrate his distinction. Here is a proposed illustration of the dissociability of phenomenal and access phenomenal consciousness (Block 1990).

'In Anton's Syndrome, blind patients do not realize that they are blind (though implicit knowledge of blindness can often be elicited). What about the converse, phenomenal consciousness without access? For an illustration at the conceptual level, consider the familiar phenomenon in which one notices that the refrigerator has just gone off. Sometimes one has the feeling that one has been hearing the noise all along, but without noticing it until it went off. One of the many possible explanations of what happens in such a case illustrates phenomenal consciousness without access consciousness: Before the refrigerator went off, you had the experience (phenomenal consciousness) of the noise (let us suppose) but there was insufficient attention directed towards it to allow direct control of speech, reasoning or action. There might have been indirect control (the volume of your voice increased to compensate for the noise) but not direct control of the sort that happens when a representation is poised for free use as a premise in reasoning and can be freely reported. (It is this free use that characterizes access-consciousness.) On this hypothesis, there is a period in which one has phenomenal consciousness of the noise without access consciousness of it.'

These examples of the dissociation of phenomenal and access consciousness are not particularly scientifically convincing, partly because of their anecdotal nature. At best they serve as examples that direct an intuition that the distinction could be possible. In this sense they support the popular impression that we see the entire visual scene even

if we cannot form reports about every element of it. At worst they offer no solid grounds at all to believe the distinction exists. A better (although still flawed) example is the Sperling partial report experiment discussed in the overview of experiment 2.

After a brief presentation of a 3 x 4 matrix of letters, observers often can report all of the letters in any cued row if the cue occurs immediately after the visual presentation ("partial report"), even though they can report only four to five letters when asked to recall all of the items in the display ("whole report"). Because all rows of the letter matrix are cued with equal probability, reporting all of the items in a randomly cued row implies that the observer has access to all of the items in the matrix at the termination of the display. The partial-report superiority effect (performance advantage of partial-report over whole-report) suggests that a fast-decaying iconic memory exists that can initially hold at least 12 items. Numerous studies have established that iconic memory has a large capacity, decays rapidly, and is destroyed by poststimulus masking (Landman et al 2003). Best reflected in the partial-report superiority effect, the duration of iconic memory has been estimated to be 300–500 ms for young adult observers. Block interprets this as evidence for phenomenal experience without access. This is a more substantial example than the anecdotal evidence he has previously relied on, however it is not without other possible interpretation.

Dehaene and Nacacche postulate 3 levels of accessibility (Dehaene and Nacacche 2001). The two extremes are I_1 (total inaccessibility even if attended) and I_3 (total accessibility) with I_2 (potentially accessible if attended) being intermediate. According to their model the unattended letters in Sperlings paradigm are I_2 , until attended and then they enter I_3 . Up to this point this is perfectly congruent with Block. The difference between Block' position and Dehaene's position lies in the properties they each assign to I_2 and I_3 . Block argues that both I_2 and I_3 are phenomenal conscious whereas only I_3 is access consciousness. Dehaene and Nacacche however claim that only I_3 encompasses phenomenal and access consciousness whereas I_1 and I_2 are unconscious by all accounts.

Dehaene and Nacacche explain that the subjects' mistake seeing the unattended letters for their potential to see them (if attended), in other words they mistake phenomenal

for potentially phenomenal. In keeping with the previous refrigerator theme this is analogous to the refrigerator light that is activated by the door opening. Every time you open the door to check the light is on it is on. This can give the impression it is on all the time, however it is just activated at the point of checking (accessing). This is known (unsurprisingly) as the refrigerator light illusion (Block 2001) an issue we will return to at the end of the thesis.

Thus there is a Blockian model in which unattended letters are a case of P without A and there is the Dehaenian model in which unattended letters are neither P nor A and the apparent P- consciousness of the unattended letters is due to the refrigerator illusion. This issue although seemingly subtle has important ramifications for theories of visual consciousness. Either stimuli outside of attention are conscious or they are not. Whatever one's opinion about the matter will largely determine the interpretation of many different experiments of consciousness, and yet this is rarely discussed in such papers. This is therefore a very important matter for the empirical investigation of consciousness. Unfortunately the argument between Block and Dehaene beyond this point has been pursued in greater depth but with somewhat diminishing returns (Block 2001). This is due to the limitations of Sperlings' paradigm over which this issue hinges. The problem lies in Sperlings use of letters as the targets which subjects are to report. The complication is that there are multiple layers to the phenomenal experience of the letters. For instance subjects can be aware of the non-semantic luminance based form of the letter as well as the semantic aspects of the letter identity. The problem is that the inference about the subjects perception is based upon their ability to report the letter identity. It is perfectly possible that subjects experience one but not the other. Therefore the discussion of what the subjects experienced is ambiguous and rather ill-posed (For this reason we chose the simplest possible stimuli of luminance-based discs for experiment 2). For our purposes I leave the issue as it stands, posing as an example of potential separability between phenomenal and access, with the intention of picking this issue up again in light of experiment 2.

One can enter into philosophical discussion until the proverbial cows come home. Ultimately the neuroscientist has to just delve into the brain at some point. It is not clear that one can state a watertight case for the existence of phenomenal and access consciousness or for that matter any other theory of consciousness. Whilst scientists

working on consciousness will always side explicitly or implicitly with some model what should be most important to neuroscientists in making claims about neural correlates of consciousness, is that they are explicit about their assumptions and premises. It has to be said that so far this has not occurred and as a result there is much confusion and conflict. I will now review some of the neuroscientific theories of consciousness which are pertinent to these issues.

2.4 Neuroscientific theories of consciousness

Block's distinction between p-consciousness and a-consciousness, at first pass, appears to be compatible with the distinction I was describing previously between perception and reportability. Perception in my terms would appear to loosely correspond to phenomenal consciousness and reportability would appear to loosely correspond to Access consciousness. The dissociation between percept and report in the context of a neuroimaging experiment would constitute the very dissociation of the neural underpinnings of p-consciousness and a-consciousness that Block proposes. However, before exploring a neuroscientific approach to dissociating perception and report it is important to review alternative hypothetical stances.

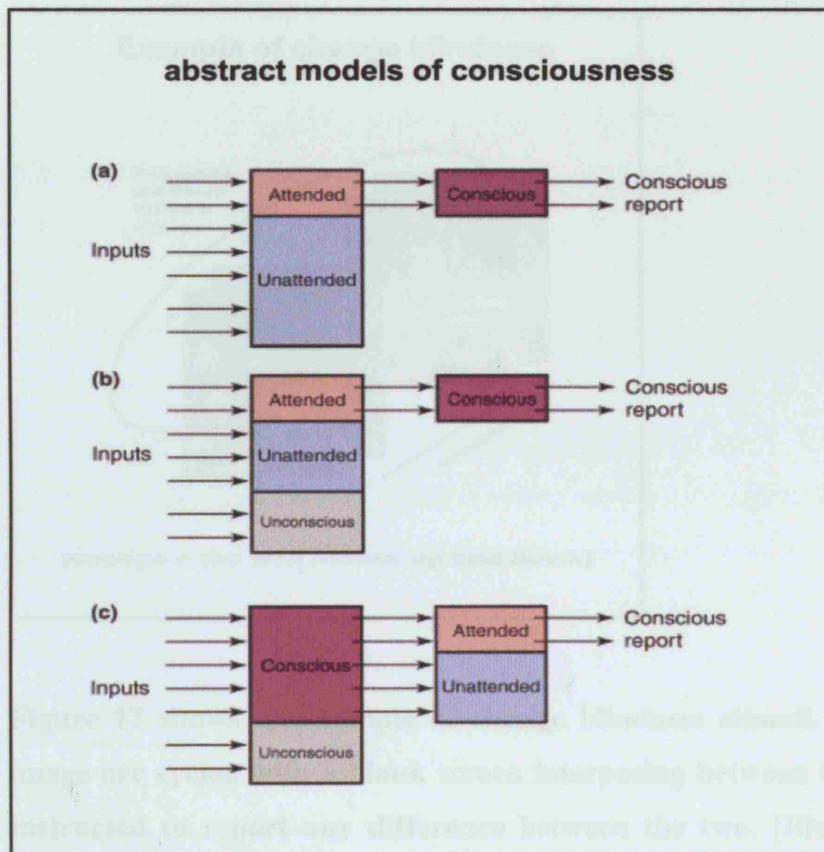


Figure 16. 3 different models of consciousness. a) There is only attended and unattended stimuli and all attended is conscious. b) There is unconscious, attended or unattended and only the attended is conscious. c) There is either conscious or unconscious and the conscious can be either attended or not attended, where only the attended is reported. [Illustration taken from Lamme 2003]

The figure above represents three models for the relationship between consciousness, report, and attention. The following models are characterised in Lamme 2003 and although here they are modified, expanded and elaborated, the skeleton of the arguments is inspired from the original article.

Figure 16 (a) depicts a simple model in which a stimulus can be either attended or unattended, and whatever is attended is conscious and therefore reportable. This is the sort of model that claims support from the results of inattention blindness and change blindness studies. Change blindness can occur when subjects are presented with a rich natural scene that alternates between two different versions in which one small aspect of the scene (colour, object, orientation etc.) changes.

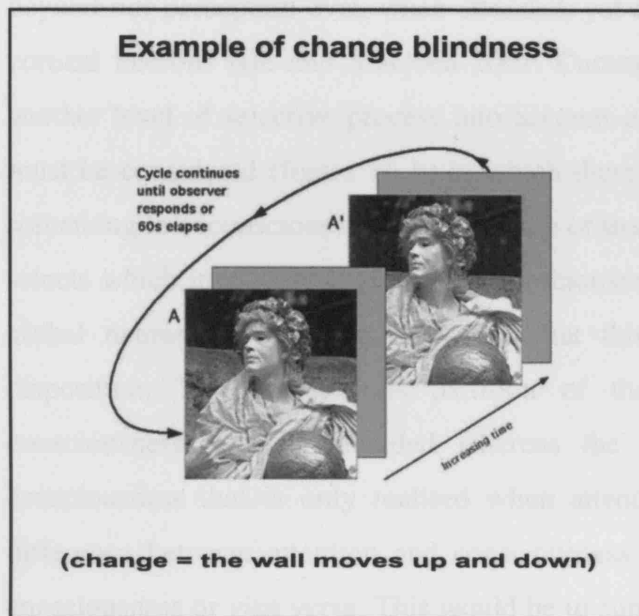


Figure 17 shows an example of change blindness stimuli. Two versions of an image are cycled with a blank screen interposing between the two. Subjects are instructed to report any difference between the two. [Illustration taken from Lamme 2003]

As long as an interposing blank screen masks the onset and offset transients of the changing element, then the change can go unnoticed for up to a several minutes (Rensink 2000, Simons et al 2000, Simons and Rensink 2005). Inattention blindness, on the other hand, can occur when subjects perform an attentionally demanding task whilst changes occur outside of the focus of attention. When asked afterwards subjects cannot report these changes (Newby and Rock 1998). This has been proposed as evidence that we are conscious of only a few privileged items at any one time, despite the fact that there is common folk-psychological intuition that one sees the entire visual panorama. In other words according to this stance, the commonplace impression that we simultaneously see the entire visual scene is an illusion. According to this interpretation, because attention guards the gate by which the privileged items reach consciousness, consciousness takes the form of the schematic in figure 16 a).

Even if the assumptions are true, the real situation cannot be quite that simple. There are obviously other considerations that determine whether something is conscious or not. For instance ultra-high spatial frequencies as well as binocular disparity can exist

beyond our perception even when attended, yet can be shown to selectively activate cortical neurons (He and Macloed 2001, Cumming and Parker 1997). Thus taking another level of selective process into account other than attention, a second model must be considered (figure 16 b) in which there is a neural determinant of whether something is unconscious or not and on top of that there is attentional selection which selects which inputs are promoted to consciousness. This is the position held by the global neuronal workspace. It seems that this model calls on the concept of disposition. The unconscious partition of the model has no disposition for consciousness even if attended whereas the unattended has a disposition for consciousness that is only realised when attended. Under this model there is no difference between attention and consciousness so you may as well call attention consciousness or vice versa. This would be to suppose that there is only one correlate of consciousness and that is one we have already studied extensively i.e. neural correlates of attention.

2.4.1 Lamme

The third model (c) is, as we shall see, influenced by Block's distinction of phenomenal vs. access consciousness, and in this particular elaboration was first espoused by Victor Lamme (Lamme 2003). Lamme is best known for his work on visual feedback, which constitutes his interpretation of the cause of late modulation effects in V1 responses to context (e.g. Lamme et al 1995). One of his theories is that feedback to V1 is necessary for visual consciousness (Lamme and Roelfsema 2000, Lamme 2001). This is a controversial statement that is refuted in its strongest most literal form by the fact that GY, a patient with an ablated V1 in one hemisphere, can experience fast motion (>5 deg/s) in his blind field (Weiskrantz et al 1995) which correlates with V5 activity (Zeki and Ffytche 1998). Additionally when various stimulus parameters are manipulated in GY's good field, it is possible to find stimuli which match GY's experience in his blind field, which has been interpreted that GY indeed does have residual phenomenal vision in his blind field (Stoerig and Barth 2001, see appendix 1 for further discussion). Since GY has no V1 for the parts of the visual field in which he has such crude visual experiences, the feedback to V1 cannot be necessary for those dimensions of consciousness, and therefore the statement that feedback to V1 is necessary for any visual consciousness appears to be unsupported.

There are however caveats to this. Firstly it is perfectly possible that given the extensive training that GY received after the original lesion, that there is sufficient plasticity for a new cortical region to take on V1's role as being the fundamental node for feedback in determining visual consciousness. The second caveat is that there may be disagreement about what actually constitutes seeing, being conscious of, phenomenal experience of X. Since it is not all clear what different experimenters class as conscious, nor is it completely clear what GY defines as conscious, then the whole disagreement could be due a terminological tangle of semantics. For instance if the visual consciousness referred to by Lamme is modal, and the experience of GY is amodal, then it is still possible that the feedback to V1 is necessary only for modal visual consciousness (for further discussion see appendix 1).

The refutations discussed above do not show that feedback is not necessary for visual consciousness in general (since it could be the case that V5 needs feedback for perception of motion); it shows that feedback specifically to V1 is not necessary for all dimensions of visual consciousness. For the purposes of this thesis I am not going to concentrate on what I have termed the essential dynamic for consciousness (principally because my chosen method of fMRI does not have the temporal or spatial resolution for such). Lamme's theory in its full version is couched in terms of recurrent feedback. Whether or not the recurrent feedback holds true as a candidate essential dynamic for consciousness, it is not necessary for the basic elements of Lamme's theory that are most relevant to reportability. For instance one could replace feedback in the following account with any other essential dynamic (e.g. synchrony) and it would still serve as a valid model of 'consciousness'.

Returning to the models in figure 16 model (c) makes an early distinction between conscious and unconscious inputs and the attentive selection process operates at an independent stage. Critically in this model attention does not determine whether inputs are conscious or not but rather determines whether a conscious report about the input is possible. Re-considering the conditions of inattentional blindness and change blindness in this context, recall that it is essential for something to be attended for it to be reported. Since attended items can be held in working memory it could be argued that the occupation of an object in working memory is critical for a change to be reported, not for it to be conscious. Such phenomena therefore do not necessarily

evidence a lack of consciousness, rather a lack of report due to a failure of working memory. Under the framework of this model, change blindness and inattention blindness are failures of conscious memory rather than failures of conscious experience per se (Lamme 2003). To put it another way we could be ‘conscious’ of many inputs simultaneously but without attention, this conscious experience cannot be accessed by working memory (and therefore cannot be reported) and is quickly erased and forgotten.

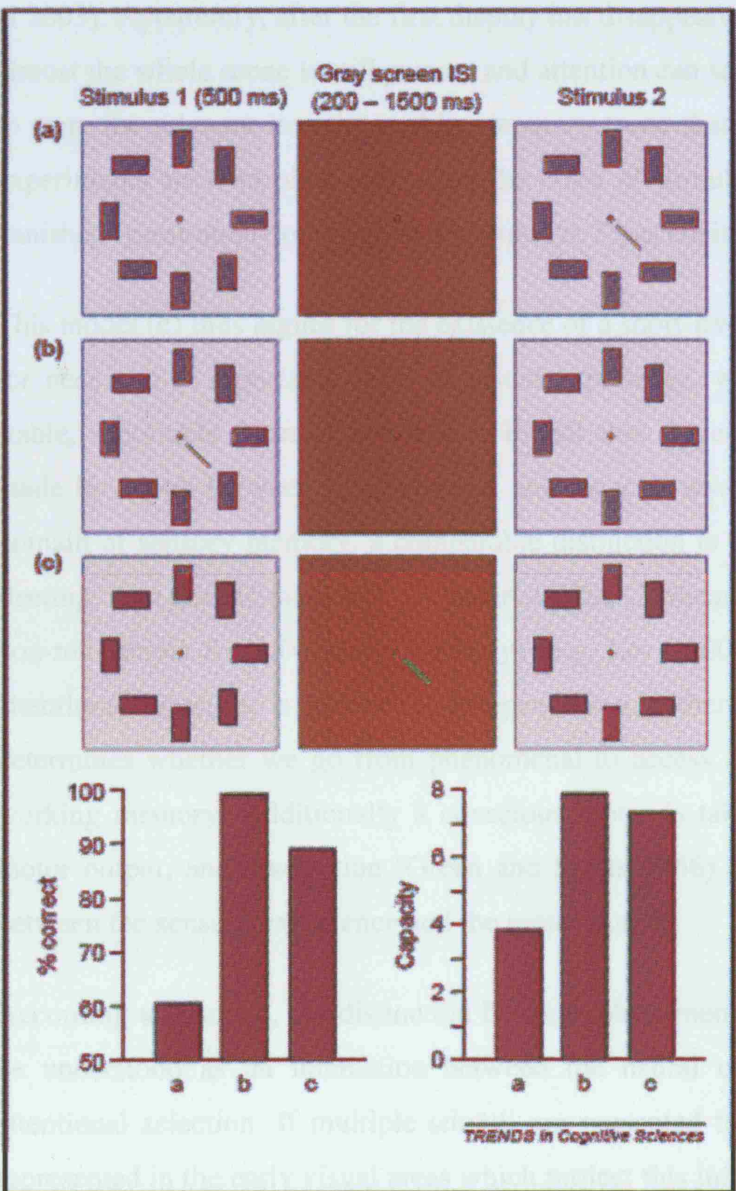


Figure 18. Figure shows that cueing 200-1500ms after stimulus offset results in performance approaching that of cueing at the time of stimulus onset. This suggests that the claim that subjects have only a very sparse representation of

the scene is incorrect, since subjects accurately report on the basis of iconic memory for a long time after stimuli have disappeared. [Taken from Landmann et al 2003.]

We have noted that cueing the item that might change in a display of many objects protects it from change blindness. Surprisingly, however, cueing the relevant item long after the first stimulus has disappeared, but before onset of the second stimulus, also protects it from change blindness (see figure 18; Becker et al 2000, Landmann et al 2003). Apparently, after the first display has disappeared, a neural representation of almost the whole scene is still present and attention can select from this representation to store the relevant item in working memory (note that retinal afterimages in these experiments are controlled for). After the onset of stimulus 2, this representation has vanished (presumably overwritten by stimulus 2), as cueing at that time does not help.

This model (c) thus argues for the existence of a short-lived, vulnerable and not easily (or necessarily) reportable form of visual experience, which contrasts with a more stable, reportable form of awareness. Recall this is very similar to the distinction made by Block between 'phenomenal' and 'access' awareness (Block 1990). In the domain of sensory memory, a comparable distinction is made between a retinotopic, fleeting form (iconic memory) (Coltheart 1980a, Coltheart 1980b) and a more durable non-retinotopic form (working memory) (e.g. Levy 2000). According to this view, attentional selection is inherently independent of either awareness or memory, but determines whether we go from phenomenal to access awareness or from iconic to working memory. Additionally a conscious report is taken exactly for what it is, a motor output, and a selection (Green and Swets 1966) or decision process that sits between the sensory experience and the motor output.

According to Lamme, the distinction between phenomenal and access awareness can be understood as an interaction between the neural correlates of awareness and attentional selection. If multiple stimuli are presented to the observer, these are all represented in the early visual areas which project this information up to higher stages of the visual hierarchy. However, at successively higher extrastriate areas, receptive fields (RFs) become larger and larger such that competition between multiple stimuli starts to arise i.e. different stimuli competing for the same receptive field (Beck and Kastner 2005, Kastner and Ungerleider 2001, Desimone 1998). Not all stimuli can be

processed in full by these receptive fields. This results in crowding phenomena, where individual awareness of closely spaced items is impaired. Attentional selection might resolve this competition, most likely through mechanisms resembling a biased competition model (Desimone 1998, Deco and Rolls 2005). In the end, only a few stimuli reach the highest levels, up to and including so-called “executive” areas. Meanwhile, the early visual areas have started to engage in recurrent interactions. In the process visual features are related to each other, binding and segregation can occur, and perceptual organization evolves. In my terms, this is what Lamme claims is the essential dynamic underpinning phenomenal awareness. Because at low levels (of the processing hierarchy) there is relatively little competition between stimuli (unless they are close to each other), many different coalitions of recurrent interactions representing multiple stimuli are possible. We can thus have phenomenal awareness of many items in a scene as long as that awareness is sufficiently shallow (low-level and not processed to higher levels e.g. abstract semantic representations).

When these recurrent interactions grow more and more widespread, and eventually include areas in executive or mnemonic space (frontal, prefrontal, temporal cortex), the visual information is put into the context of the systems' current needs, goals and full history. This is what, according to the model, is the essential dynamic that produces access awareness of a-consciousness (in Blocks terms). There is considerable competition, however, for interaction with these higher levels and therefore only a limited number of recurrent coalitions can span the range from visual to more frontal (parietal, temporal) areas. Therefore, access awareness is limited to a few items in the scene. Depending on the extent to which recurrent interactions between visual areas incorporate interactions with action or memory-related areas, awareness evolves from phenomenal to access awareness. Whether this occurs depends on attentional selection mechanisms, via influences on both the feed-forward sweep and recurrent interactions.

2.4.2 Global workspace models

Workspace models are another family of theories that are distinct from those already discussed. The original conception of the theory was by Bernie Baars and later taken expanded by Stanislas Dehaene and colleagues.

2.4.3 Baars

To explain the differences between conscious and unconscious processes, Baars recruits and adapts a popular model in artificial intelligence, which he calls the 'distributed-processing systems' model. Many artificial intelligence systems rely on a globally accessible block of working memory to mediate communication and novel interaction between the various individual processors. Baars proposes that a similar structure exists in the human brain, and that it supports conscious experience. The structure, which he terms the 'global workspace', is accessible by most processors (presumably a brain processor corresponds simply to a brain area), meaning that most processors have their contents potentially accessible by working memory. The global workspace can also "broadcast" its contents globally in such a way that every other processor receives or has access to the 'conscious' content. The Global Workspace is serial and of limited capacity in that at any given time, only one processor's contents may be "broadcast" over the global workspace to the other processors. Baars acknowledges that an important part of a complete global workspace theory of consciousness would be an explanation of how a given processor gains access to the global workspace. In the absence of such a mechanism the workspace model remains a box and arrow theory. Presumably there are a multitude of processors that could have their contents broadcast by the global workspace. However, according to the theory, only one processor's representations are broadcast at any one moment (consciousness is therefore serial), and in order for the global workspace to be functionally advantageous, as consciousness is assumed to be (epiphenomenalists notwithstanding), there must be an explanation for why one processor gains access to the global workspace and another does not. Baars describes two possible systems for controlling access to consciousness, and they in themselves are large enough in scope for extra thesis chapters. I will not discuss these to maintain a reasonable focus, and we will proceed under the assumption that an access system exists in conjunction with the global workspace.

The global workspace theory is receiving a rather large consensus amongst disparate fields from philosophers (Daniel Dennett) to neuroscientists (Stanislas Dehaene, Gerald Edelman) in that it is by many accounts the most popular family of contemporary theories of consciousness. Although important in terms of the genealogy of workspace theories, it is hard to pin down exactly what Baars' original version is really saying over and above the fact that consciousness involves a workspace in which processors broadcast their contents to one another. To discover the true potency of the workspace theory it is necessary to turn to a more recent neurally inspired incarnation of the global workspace.

2.4.4 Global Neuronal Workspace

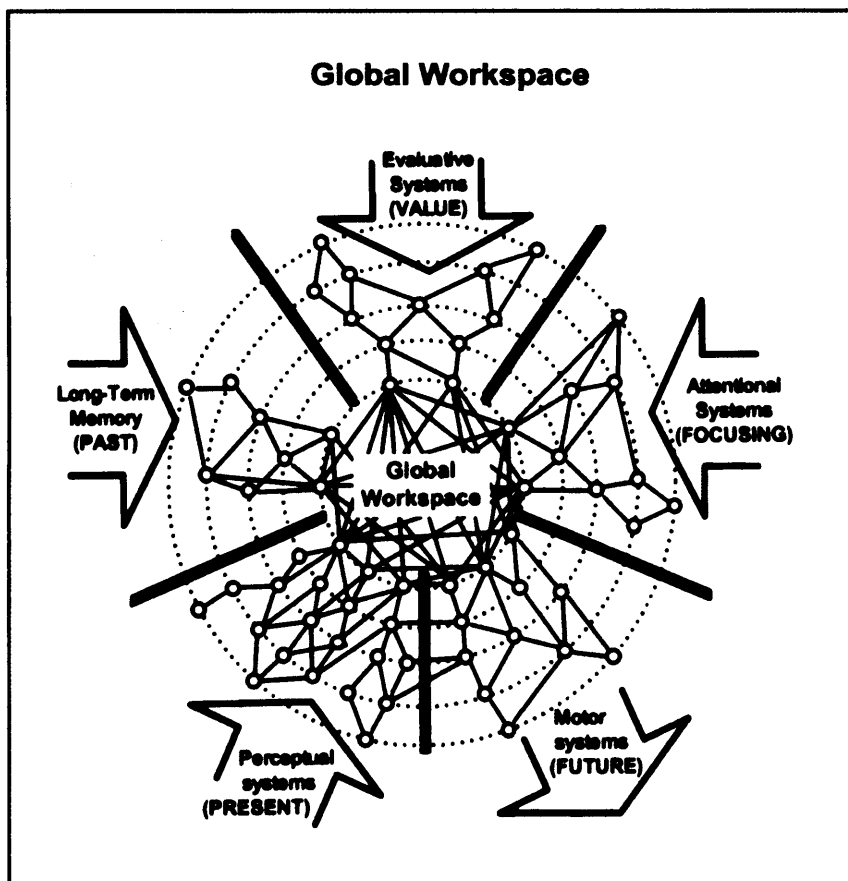


Figure 19. Cartoon depiction of the global neuronal workspace as schematised by Stanislas Dehaene (Dehaene and Naccache 2001). Key to the model is the principle of a core network of areas (i.e. the workspace) that can flexibly incorporate any number of modules from any of the five subsystems (attention, perception, motor etc.). The workspace takes a different shape according to the

modules it mobilises through the ‘ignition’ process (see below). [Figure from Dehaene and Naccache 2001]

Dehaene has elaborated the global workspace hypothesis of Baars and translated it into a much more sophisticated neurally based hypothesis called the ‘Global Neuronal Workspace’. This framework postulates that, at any given time, many modular cerebral networks are active in parallel and process information in an unconscious manner. Information becomes conscious, however, if the neural population that represents it is mobilized by top-down attentional amplification into a brain-scale state of coherent activity that involves many neurons distributed throughout the brain. The long-distance connectivity of these ‘workspace neurons’ can, when they are active for a minimal duration, make the information available to a variety of processes including perceptual categorization, long-term memory encoding, evaluation, and intentional action. The postulation is that this global availability of information through the workspace is what we subjectively experience as a conscious state. The data from the brain-imaging literature argues for a major role of prefrontal cortex, anterior cingulate, and the areas that connect to them, in creating the postulated brain-scale workspace.

The workspace model comprises a distributed neural system with long distance connectivity that can potentially interconnect multiple specialised brain areas in a coordinated manner. Through the workspace modular systems that do not directly exchange information in an automatic fashion can nevertheless gain access to each other’s content. The workspace therefore provides a communication protocol through which a particularly large potential for combination of multiple input, output and internal systems become available. The connections between the workspace and motor and language systems underpin verbal and non-verbal reportability of conscious states.

The critical role of attention in consciousness is made explicit in the model in that it is stated that top-down attentional amplification is the mechanism by which modular processes can be temporarily mobilised and made available to the workspace and therefore to consciousness. The modules that the workspace mobilises determine which modules contribute to the content of consciousness. To be conscious the module must be amplified and maintained over a sufficient duration for it to become

accessed by other parts of the workspace. Therefore the workspace is anatomically 'slippery' in that any region that is mobilised, and any region with the appropriate long-range connections and distribution can play its role in the workspace. This is similar in essence (but not in all details) to the dynamic core hypothesis of Edelman (Edelman and Tononi 2001). They insist that no region is essential for consciousness (other than the supporting systems that maintain arousal such as reticular activation systems, Intralaminar nucleus of the thalamus) but rather the type of the activation that is the determining factor (long range mobilisation through top-down amplification in the case of workspace theories and re-entry for dynamic core theory).

The dependence on non-linear positive feedback imposed by the top-down attentional amplification leads to the postulation of two thresholds, the first being that necessary to cause differential neural processing and a second being that which exists for a module to be amplified enough that it enters the workspace through the establishment of a self-sustaining long-distance loop. There is some evidence to support this (Sergent and Dehaene 2004). In contradiction to the models supported by Lamme and Block (Model c of figure 16), the position that Dehaene takes is that attention is a prerequisite of consciousness. Dehaene therefore subscribes to the model (b) in the figure 16 where only what is attended is conscious, and there is no consciousness beyond that. He believes that only what is accessed has phenomenal experience associated with it and therefore implicitly rejects any real distinction between the two.

Having introduced a representative, but by no means exhaustive, range of viewpoints pertaining to the neural correlates of consciousness, we shall proceed to introduce some of the technical concepts which are relevant to the experiments which follow.

3.1 Introduction to Signal detection theory.

Some of the key concepts of the second experiment of this thesis rely on a basic understanding of signal detection theory. Central to reporting a percept is that a decision must be taken about the content of the percept in order to determine the content of the report. In reality nearly all such decisions are made in the presence of some uncertainty due to the omnipresence of noise (in the sense organs, in the brain and in the world outside). Signal detection theory provides a precise language for analysing decision-making in the presence of uncertainty. Signal detection theory was developed in the 1950's and 1960's by the collaborations of Tanner and Swets (Tanner and Swets 1954) and Green and Swets (Green and Swets 1966) among others, having had its antecedents in Thurstone's (Thurstone 1927) work on comparative judgement. It makes little sense to expand on the history of signal detection theory when it is, for the purposes of this thesis, part of the analytical approach. I will therefore focus on a general 'textbook-like' introduction to signal detection theory (some of which is borrowed from David Heeger's Signal detection theory course) and then move onto an example of its application to an fMRI experiment:

Let us suppose a subject is making a simple decision about the presence or absence of a flash of light in a dark room. Because the light flash is very dim there will be uncertainty in the subject's decision and consequently they will make mistakes. It follows therefore that there are 4 types of outcome when the subject is asked 'did you see a flash?'

1. The light was present and the subject reports 'yes' = HIT
2. The light was present and the subject reports 'no' = MISS
3. The light was absent and the subject reports 'yes' = FALSE ALARM
4. The light was absent and the subject reports 'no' = CORRECT REJECTION

To optimise performance the subject must try to maximise the number of hits and correct rejections and to minimise the number of misses and false alarms. There are

hypothesised to be two main stages to the decision process, (1) acquisition of sensory information and (2) decision criteria.

Acquisition of information corresponds in this example to the strength of the light and the ability of the visual system to detect and process the occurrence of the light flash; the stronger the light, or the more sensitive the visual system, then the better the acquisition of the relevant sensory information. The stronger or more reliable the information is, the higher the likelihood the subject will have of making the right response and therefore maximising the number of hits and correct rejections.

The second component of the decision process is the criterion that the subject uses to decide whether to report 'yes' or 'no'. In plain English this corresponds to the bias in the judgement of the subject. For instance, if the flash of light were very infrequent, say once every two hours, then the probability of a flash would be very low. In this case the subject would be more cautious in saying yes given that it is overwhelmingly more likely that there was no flash. The converse bias would occur in the same subject however if the flash were every two seconds. They would be much more likely to report yes on any one trial. Such a bias will only ever occur in the presence of uncertainty. If the event is salient enough, such as a thunder, then no matter what the prior probability of it occurring, it will be recognised as thunder with great certainty. If the same thunder were very far away, then the same observer will be less certain and therefore more prone to such biases. So where does the uncertainty come from?

The uncertainty of a subject can come from two sources: internal and external noise:

External noise would correspond to the flash of light being presented on an old TV monitor, which had a noisy display. This would cause physical variability in the luminance of the monitor whether the flash was present or absent.

Internal noise refers to the well-established fact that the neural processing stream of the visual system (from retina to the visual cortex and thalamo-cortical circuits) is noisy. For instance the same flash may cause in a particular visual neuron 10 spikes/sec on one occasion and on another 30 spikes/sec. The variability in the

distribution of internal responses of a neuron in the presence and absence of a signal (e.g. flash) is depicted in the following diagram.

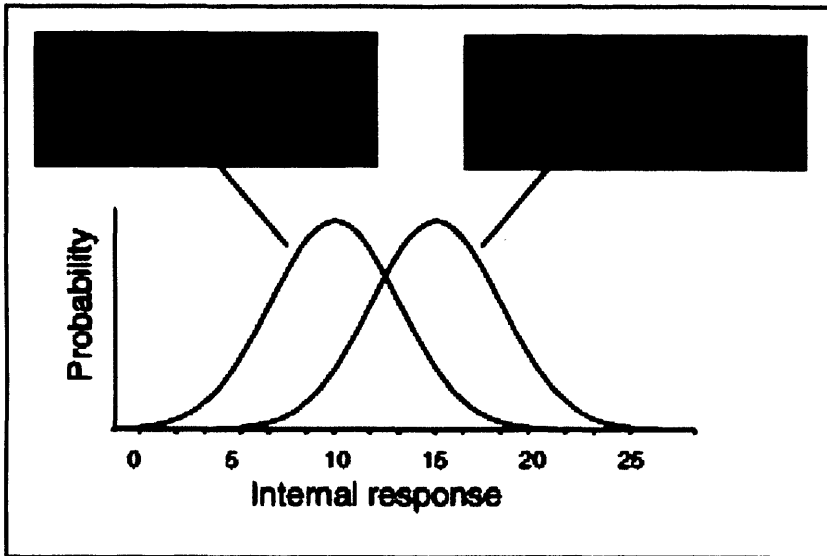


Figure 20. Depiction of two overlapping distributions of internal responses. Internal response is deliberately non-specific but could map onto any index of neural activity (single unit rates, Local Field Potentials etc). This is a biologically plausible distribution of 'brain responses' since it has long been demonstrated that all are inherently noisy resulting in an overlap of the distributions.

The curve on the left reflects the variability in, say, the baseline firing of a particular neuron in the absence of the flash and the curve on the right reflecting the variability in the response when the flash was present. It is important to notice that the curves may overlap such that the response when the signal is absent may exceed the response when the signal is present. Thus for any area making a decision based on the activity of this neuron will find it very hard to distinguish the signal (signal plus noise) from noise (noise alone) all of the time. Therefore mistakes will inevitably be made resulting in false positives and misses. The simplest model of decision-making in this circumstance is to set a criterion location along the internal response axis. Whenever the internal response exceeds this criterion the decision made is yes, and no if the response fails to reach this criterion.

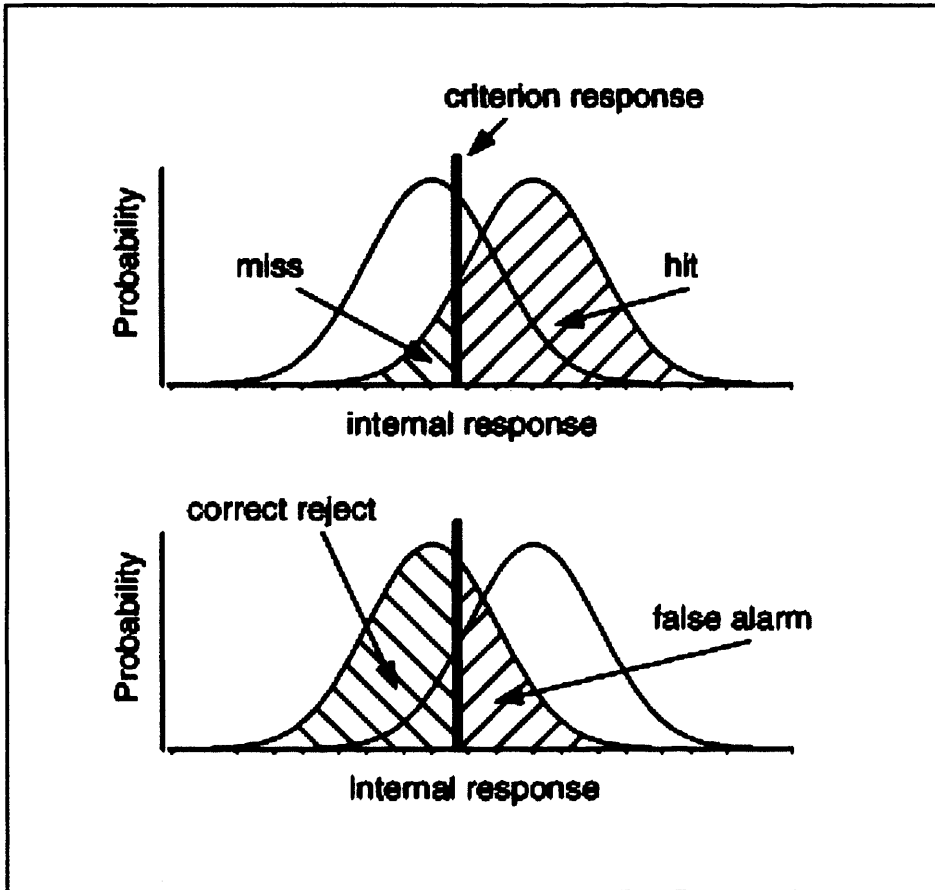


Figure 21. A logical consequence of the overlapping response distributions is that any system that makes a decision based on whatever variable the response represents, will do so with uncertainty for the range in which there is overlap. If there is any overlap then it is impossible for a decision system accessing that information to make a decision that does not involve some mistakes being made.

Now consider the consequences of shifting the criterion according to a change in bias.

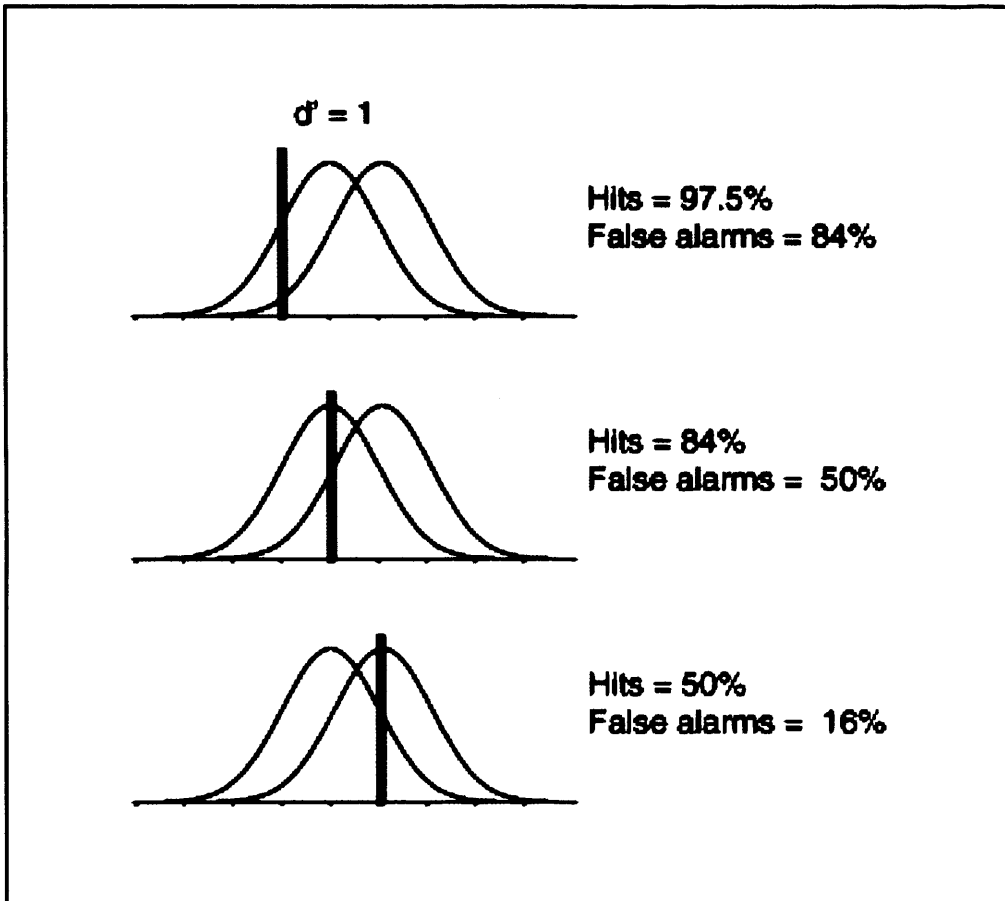


Figure 22. This figure demonstrates the effects of changing the criterion a decision system uses. By changing the criteria the relative proportion of hits and false alarms changes. One could set a criteria so low that the hit rate is almost 100% however this would be associated with the cost of a higher false alarm rate. Alternatively one could minimise false alarms but this would be at the expense of the hit rate. The optimal solution to minimise mistakes is to opt for a criteria at the point the two distributions meet.

In the top profile, the subject has shifted their criterion so low that they say yes to nearly every trial. This means that the hit rate will be high so will the false alarm rate. Conversely by shifting the criterion higher the subject now will make much fewer false alarms but their hit rate will fall. Note that no matter where the subject puts their criterion they will make mistakes. Thus with distributions which overlap subjects cannot ever be 100 percent right all of the time. They are destined to make mistakes and all they can do is to adjust the kind of errors that they make by adjusting their criterion.

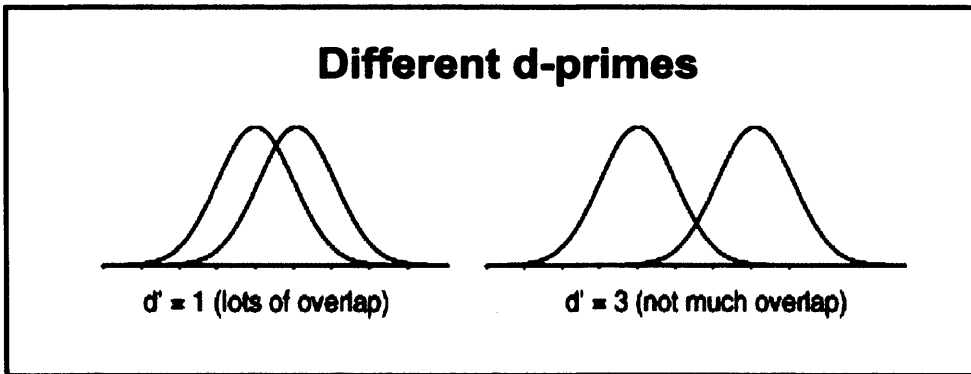


Figure 23. Figure depicts how increasing the separation increases the discriminability of the signal and therefore is reflected in the higher d-prime (d-prime=separation/spread)

Adjusting the criterion can change the relative proportion error types that are made, but what determines the propensity for error making in general? In plain English what determines whether the decision is easy or hard?

There are two things to consider.

1. Intuitively, acquiring more information makes the decision easier. In our example this would increase the strength of the flash such twice as many photons were emitted. In this case, the subject's internal response strength will, on average, be stronger. In less abstract language this would mean that the response of V1 would be increased (on average). Pictorially, this will have the effect of shifting the probability of occurrence curve for signal-plus-noise trials to the right, a bit further away from the noise-alone probability of occurrence curve. Figure 23 shows two sets of probability of occurrence curves. When the signal is stronger there is more separation between the two probabilities of occurrence curves. When this happens the subject's choices are not so difficult as before. They can pick a criterion to get nearly a perfect hit rate with almost no false alarms.

2. The amount of noise. The lower the noise the less the distribution of response is spread out. For example, consider the two 'probability of occurrence' curves in Figure 23. The separation between the peaks is the same but the second set of curves are much skinnier. Clearly, the signal is much more discriminable when there is less spread (less noise) in the probability of occurrence curves. So the subject would have an easier time setting their criterion in order to be right nearly all the time.

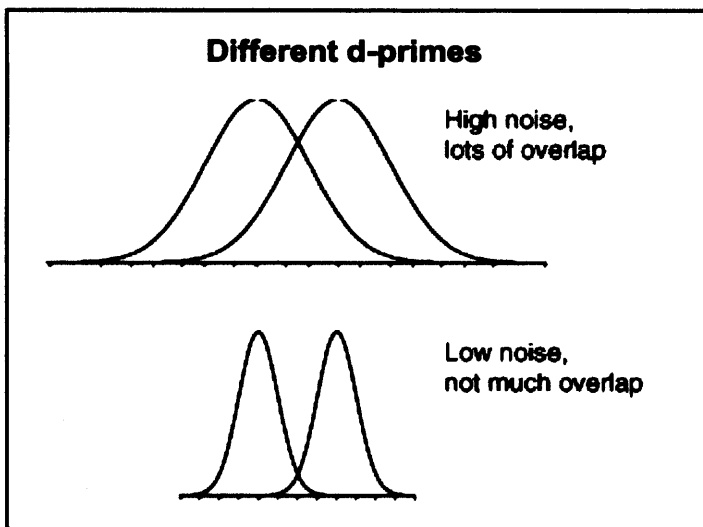


Figure 24. Alternative to increasing the spread, increased discriminability can result from decreasing the noise.

In reality, we have limited control over the amount of internal noise. But it is important to realize that decreasing the noise has the same effect as increasing the signal strength. Both reduce the overlap between the 'probability of occurrence' curves.

D-prime as a measure of discriminability

From the two considerations above it is clear that the discriminability of a signal depends both on the separation and the spread of the noise-alone and signal-plus-noise curves. Discriminability is made easier either by increasing the separation (stronger signal) or by decreasing the spread (less noise). In either case, there is less overlap between the 'probability of occurrence' curves. To write down a complete description of how discriminable the signal is from 'no-signal', we want a formula that captures

both the separation and the spread. The most widely used measure is called d-prime (d'), and its formula is simply:

$$d' = \text{separation} / \text{spread}$$

This number, d' , is an estimate of the strength of the signal. Its primary virtue, and the reason that it is so widely used, is that its value does not depend upon the criterion the subject is adopting, but instead it is a true measure of the strength of internal response that is independent of the criterion. d' can be estimated from recording the proportion of hits and false alarms alone (see appendix 2).

3.2 Neuroimaging studies employing Signal Detection Theory

One advantage of signal detection theory (SDT) is that it allows one to link psychophysics with recordings from the brain. An important demonstration of this approach is a study by Ress and Heeger (Ress and Heeger 2003). When human subjects performed a simple contrast discrimination task their responses could be partitioned according to the 4 signal detection categories. Ress and Heeger partitioned the evoked fMRI responses according to the same signal detection categories to find the signal detection profiles for each visual cortical region.

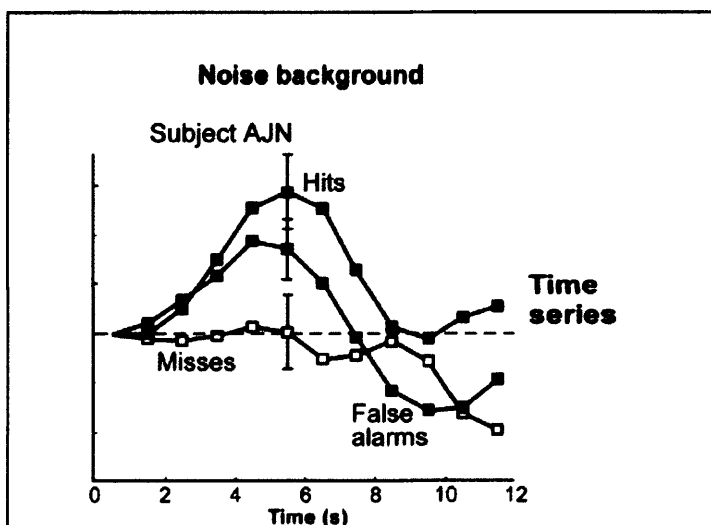


Figure 25. Peri-stimulus time histogram of the BOLD signal caused by the different trial types categorised post-hoc according to the subjects perceptual reports. The responses have the baseline correct rejection trials subtracted from

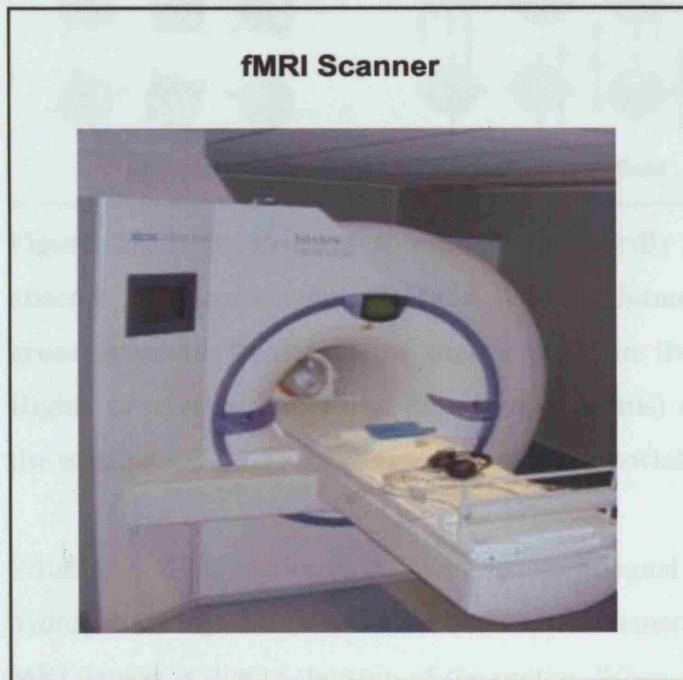
them. Note how the response reflects the report not the stimulus presence. Since the experiment assumes perception and report are in full agreement (perception determines report without any mistakes) and that V1 reflects perception not motor components of the report this response is posited as a correlate of perception. [Figure from Ress and Heeger 2003]

The graph in the figure above demonstrates how the response to each signal detection category ranks with hits >>false positives>misses. This shows that irrespective of whether the stimulus difference was present or not the report of the subject correlated with the response of visual cortex. This was true for V1, V2 and V3. The conclusion of this then was that these early visual areas correlate with the perception irrespective of the stimulus. There is however a problem with this experiment stemming from the aforementioned problems stemming from the use of reportability as an index of perception. Firstly the experiment assumes that that the report is an index of perception. Given this assumption, how can you tell the difference between a true neural correlate of perception and a neural correlate of a motor response since they both yield exactly the same predictions for this specific paradigm? If one found the pattern of the V1 response in motor cortex then noone would claim it to be a neural correlate of perception. Therefore one would attribute the pattern of the response to the report yet there would be no actual difference in the data other than where in the brain it was found. It is only because of where the activity is in the brain that leads the interpretation that it is a perceptual response. Whilst I believe that this is most likely, there is nothing in the data to support this. The interpretation of this experiment shall be discussed in greater depth in the discussion section of experiment 2.

4. Methodological introduction.

4.1 Introduction to functional magnetic resonance imaging

Finally I introduce some of the basics of fMRI method.



**Figure 26. The Sonata scanner of the Functional Imaging Laboratory, UCL.
[Photo taken from FIL website www.fil.ion.ucl.ac.uk]**

Functional Magnetic Resonance Imaging (fMRI) is a technique that can measure brain activity through changes in blood flow that accompany neural activity. To explain how this works it is necessary to first describe how MRI (magnetic resonance imaging) machines work.

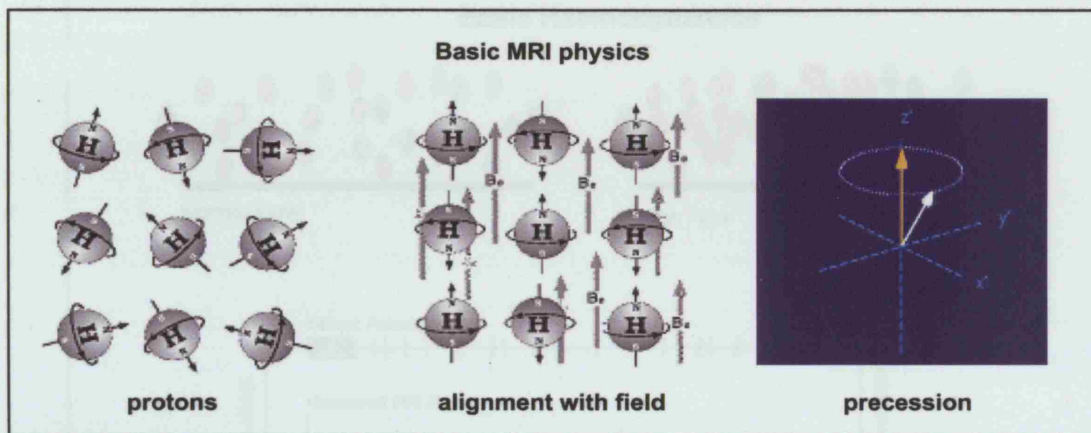


Figure 27. Left: Protons spinning haphazardly with no net alignment in the absence of a gross magnetic field. Middle: Same protons in the presence of a gross magnetic field such as that induced in the bore of the Sonata scanner. Right: precession (‘turning’ in layman’s terms) of a proton around the axis of the magnetic field. [Figure taken from Frackowiak et al 2004]

What does MRI measure? MRI measures a signal from protons (the nucleus of a hydrogen atom), which is the most common element in brain tissue. The signal that MRI detects is due to the spin of the proton. When a magnetic field is applied to the brain the protons align with the magnetic field and turn (technically known as precess) around the magnetic field axes. Each type of nucleus in the brain has a unique frequency of spin that is proportional to the strength of the field otherwise known as its resonance. If a second magnetic field is applied at right angles to the static field at the precessing frequency of the atomic nuclei, the nuclei will all precess around the static field together in-phase. A nearby coil can detect this in-phase precession by picking up the resultant current induced through it with its own distinct frequency. When this field is turned off, the nuclei will progressively de-phase with time. It is this change in signal, which is detected in fMRI. The next question then is ‘How does this link to brain activity?’

Blood Oxygen Level Dependent signal = BOLD

The central assumption of fMRI is the linear transfer model that states that the BOLD signal is approximately proportional to the neural activity summed over several millimetres and several seconds. The details of the link between activity and BOLD

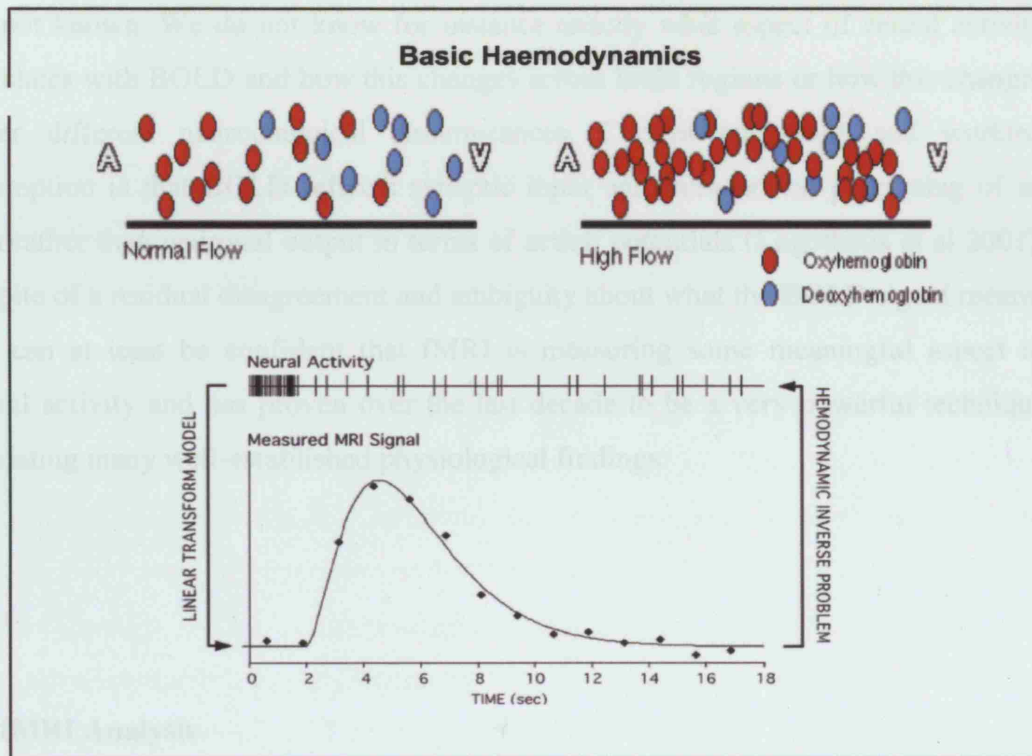


Figure 28. Upper diagram depicts the change in the distribution of oxy and deoxy haemoglobin induced by the metabolic response to neural activity. Since ‘deoxy’ are paramagnetic but ‘oxy’ is diamagnetic this change in proportion results in a detectable change in the local magnetic field strength. Lower diagram depicts the neural cause and the canonically modelled effect it has on the MRI signal that constitutes the BOLD signal. [Figure taken from Heeger 1999]

Neural activity causes a local increase in oxygen consumption due to increased metabolic demand. After 2 s there is increased blood flow to the active area, which overcompensates for the amount of oxygen extracted. This effective increase in blood oxygenation causes a change in the local magnetic field strength and it is this which is measured in fMRI:

Blood Oxygen Level Dependent signal. = **BOLD**

The central assumption of fMRI is the *linear transfer model* that states that the fMRI signal is approximately proportional to the neural activity summed over several millimetres and several seconds. The details of the link between activity and BOLD

are not known. We do not know for instance exactly what aspect of neural activity correlates with BOLD and how this changes across brain regions or how this changes under different neurochemical circumstances. Current consensus and working assumption is that BOLD reflects synaptic input and intracortical processing of an area rather than neuronal output in terms of action potentials (Logothetis et al 2001). In spite of a residual disagreement and ambiguity about what the BOLD signal means, one can at least be confident that fMRI is measuring some meaningful aspect of neural activity and has proven over the last decade to be a very powerful technique validating many well-established physiological findings.

4.2 fMRI Analysis

The ultimate goal of fMRI data analysis is to detect correlations between brain activation and the sensory/cognitive/emotional/motor operations the subject performs during the scan. The BOLD signature of activation is relatively weak, however, so other sources of noise in the acquired data must be carefully controlled. This means that a series of processing steps is usually performed on the acquired images before the actual statistical search for activation can begin.

For a typical fMRI scan using an EPI pulse sequence the 3D volume of the subject's head is imaged every one or two seconds, producing a few hundred to a few thousand complete images per scanning session. The nature of MR imaging is such that these images are acquired in Fourier transform space, so they must be transformed back to image space to be useful. Because of practical limitations of the scanner the Fourier samples are not acquired on a grid, and scanner imperfections like thermal drift and spike noise introduce additional distortions. Small motions on the part of the subject and the subject's pulse and respiration will also effect the images.

Software in the scanner platform itself then performs the reconstruction of images from Fourier transform space. During this stage some information is lost. Some types of artifacts, for example spike noise, become more difficult to remove after

reconstruction, but if the scanner is working well these artifacts are thought to be relatively unimportant.

After reconstruction the output of the scanning session consists of a series of 3D images of the brain. The most common corrections performed on these images are motion correction and correction for physiological effects.

4.2.1 SPM

SPM is the name of statistical software used in the experiments of this thesis to analyse fMRI data. SPM stands for *Statistical parametric mapping* and is generally used to identify functionally specialized brain responses and is the most prevalent approach to characterizing functional anatomy and disease-related changes (Friston et al 1990, Friston et al 1991). The characterisation of regionally specific effects in the brain is achieved through classical statistical estimation and inference. Inferences in fMRI may be about differences expressed when comparing one group of subjects to another or, within subjects, changes over a sequence of observations. In order to assign an observed response to a particular brain structure, or cortical area, the data must conform to a known anatomical space. Before statistical modelling the time-series of images are realigned and mapped into some standard anatomical space (*e.g.* a stereotactic space).

The statistical approach taken by SPM is a mass-univariate (univariate meaning only one dependent variable, mass meaning many univariate tests in one brain) analysis, whereby experimental effects (*e.g.* stimulation by experimenter) are tested at every single voxel independently. The basic approach is that of multiple linear regression, whereby known experiment variables (*e.g.* the onset of a visual event) and nuisance variables (head movements or scanner drift) are assigned regressors, which are then regressed onto the fMRI time series. Parameters are estimated such that they are scaled to optimally minimise the global error between the model and the data for each voxel (independently). This estimation stage results in each regressor being assigned a Beta (regression co-efficient) value that is the magnitude by which each regressor needs to be scaled (along with all other regressors) to minimise the residual error. The betas then correspond to a measure of the peak BOLD response to a particular event.

Because this is a mass univariate approach this multiple regression is performed for every single voxel, leaving one Beta (parameter estimate) for every single regressor at every single voxel. Statistical operations can then be performed such as t-tests and f-tests to test the significance of the effect of any combination of regressors, known as contrasts, which when performed across all voxels results in a statistical parametric map of the voxels which show a significant effect for the contrast (this is what the blobs of activity are in the pretty imaging figures).

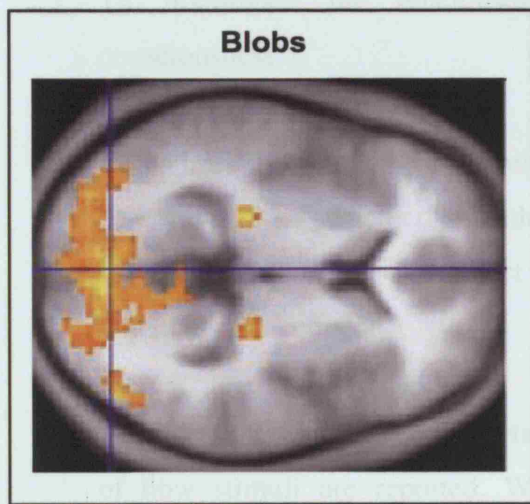


Figure 29. SPM blobs

5. Reiteration of aim

Having introduced the history of functional specialisation, the relevant contemporary empirical and philosophical literatures as well as the basic technical concepts used in this thesis the reader should have sufficient background with which to understand the context for the following experiments. Before describing the experiments in detail I will briefly reiterate my objectives.

1. To investigate the relationship between neural activity and visual consciousness.
2. To address the problem of visibility and invisibility by looking at the neural determinants of whether something can be seen (or not). We therefore investigate the response of object selective areas when presented with objects that undergo occlusion.
3. To address the problem of reportability by looking at the neural determinants of how stimuli are reported. We used a 'partial report methodology' to dissociate the neural correlates of these two processes in an fMRI study of object detection.

By consideration of both findings I attempt to synthesise (with reasonable contrivance) a perspective that unites the neural correlates of visibility invisibility and reportability into one framework. The following sections will be presented in the following order:

Methods and results for experiment 1

Methods and results for experiment 2

Discussion of experiment 1

Discussion of experiment 2

Discussion of both experiments in the context of wider theoretical issues

Speculative discussion of a theory of consciousness

Figure 30A Schematic depiction of the 6 conditions for faces (equivalent for objects) (not pictured). Time runs from top to bottom. There are 3 stimuli types, Occluded (OC), Visible (V) and Segment (SG), each with

6. Methods and Results

6.1 Methods: Experiment 1

Vanishing (OCv), Visible Vanishing (Vsv), Segment Vanishing (SGv). The vanishing controls are identical to their counterparts for the first 1.5 s after which the object (in this example the face or the face segment) vanishes and is not seen for the remaining 7.5 s. The stimulus sequence is depicted with cartoon vanishing faces that are for illustration purposes only.

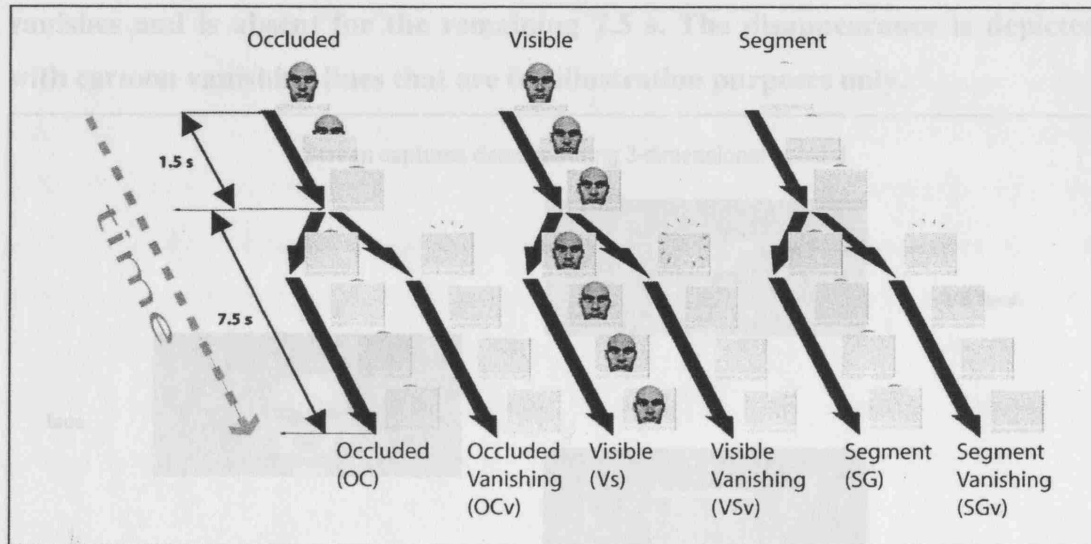


Figure 30A Schematic depiction of the 6 conditions for faces (equivalent conditions for house stimuli not pictured). Time runs from top to bottom. There are 3 stimuli types, Occluded (OC), Visible (VS) and Segment (SG), each with their own control Occluded Vanishing (OCv), Visible Vanishing (VSv), Segment Vanishing (SGv). The vanishing controls are identical to their counterparts for the first 1.5 s after which the object (in this example the face or the face segment) vanishes and is absent for the remaining 7.5 s. The disappearance is depicted with cartoon vanishing lines that are for illustration purposes only.

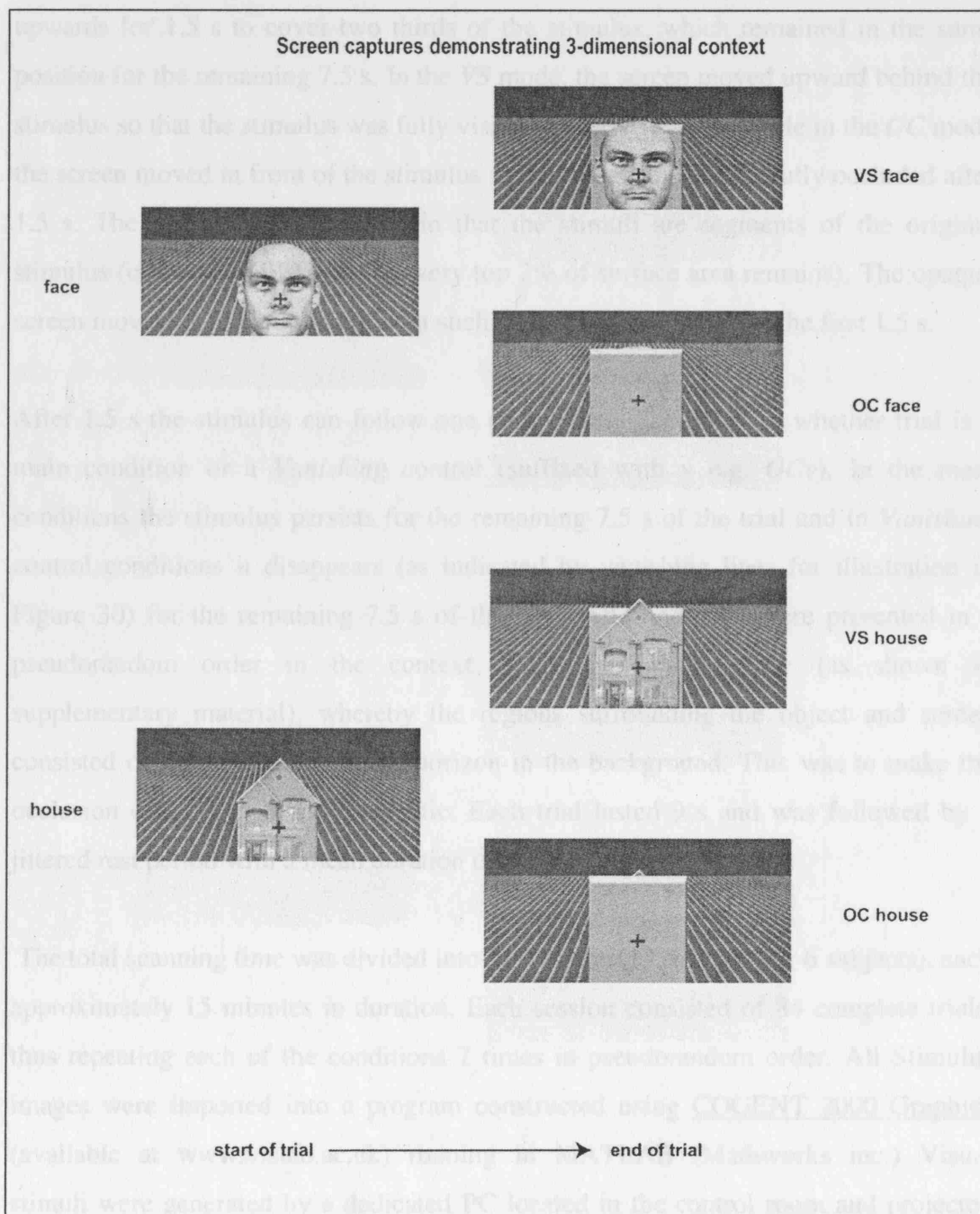


Figure 30B shows screen captures of the difference between occluded and visible conditions.

Stimuli consisted of achromatic photographs of 10 faces (provided by the Max-Planck Institute for Biological Cybernetics, Tuebingen, Germany) and 10 houses (photographs taken specifically for the experiment) in frontal view. The object stimuli occupied 5° of visual angle. Figure 30 illustrates the 6 different conditions for each object type. Each of the conditions could be displayed in one of three modes *Visible (VS)*, *Occluded (OC)* or *Segment (SG)* mode. In each mode, an opaque screen moved upwards for 1.5 s to cover two thirds of the stimulus, which remained in the same position for the remaining 7.5 s. In the *VS* mode, the screen moved upward behind the stimulus so that the stimulus was fully visible for the first 1.5 s. while in the *OC* mode the screen moved in front of the stimulus so that it became almost fully occluded after 1.5 s. The *SG* mode is different in that the stimuli are segments of the original stimulus (cropped so that only the very top 2% of surface area remains). The opaque screen moves up toward the segment such that it is fully visible for the first 1.5 s.

After 1.5 s the stimulus can follow one of two fates according to whether trial is a main condition or a *Vanishing* control (suffixed with v e.g. *OCv*). In the main conditions the stimulus persists for the remaining 7.5 s of the trial and in *Vanishing* control conditions it disappears (as indicated by vanishing lines for illustration in Figure 30) for the remaining 7.5 s of the trial. All conditions were presented in a pseudorandom order in the context of 3-dimensional space (as shown in supplementary material), whereby the regions surrounding the object and screen consisted of depth contours and a horizon in the background. This was to make the occlusion event appear more realistic. Each trial lasted 9 s and was followed by a jittered rest period with a mean duration of 3 s.

The total scanning time was divided into 2-3 sessions (3 sessions for 6 subjects), each approximately 15 minutes in duration. Each session consisted of 84 complete trials, thus repeating each of the conditions 7 times in pseudorandom order. All Stimulus images were imported into a program constructed using COGENT 2000 Graphics (available at www.vislab.ac.uk) running in MATLAB (Mathworks inc.) Visual stimuli were generated by a dedicated PC located in the control room and projected

via a LCD projector onto a transparent screen positioned over the subjects head, and viewed through a tilted mirror fixed to the head coil.

Subjects

Thirteen subjects (8 males and 5 females, mean age 25.4 years, 7 right handed) with normal or corrected to normal vision were scanned. All gave informed consent in accordance with the Declaration of Helsinki, and the Ethics Committee of the National Hospital for Neurology and Neurosurgery, London, UK, granted ethics approval for the study.

Task

To ensure that subjects maintained fixation and attention they were instructed to fixate a central cross at all times. To ensure that subjects attended to the stimuli they were instructed to respond via button press if there was a change in the fixation cross. This occurred in only 15% of trials (randomly distributed across trials and randomly timed within a trial). If no change occurred they did not press any button. The performance of all subjects was recorded and found to be 95.5% (mean) correct. The hand with which each button was pressed, and handedness, was counterbalanced across all 13 subjects. Subjects had to perform the same task on every trial regardless of condition. Eye-tracking data was monitored to check correct fixation.

Scanning details

Scanning was done in a 1.5T Siemens Sonata fMRI scanner with a head-volume coil (Siemens, Erlangen, Germany). A gradient echo-planar imaging (EPI) sequence was used to maximise blood oxygen level dependent (BOLD) contrast (TE = 90ms, repeat time 3.42s). Each brain image was acquired in a descending sequence comprising 38 axial slices each 2mm thick with 1mm gaps in-between, consisting of 64 x 64 voxels, and thus covering nearly the whole brain. Each session consisted of 247 volumes with the first 5 volumes being discarded to allow for T1 equilibration effects. Each subject was also scanned after the main scanning session with a T1 weighted structural sequence, using a phased array head coil to obtain a high-resolution structural image.

Data analysis

We analysed the data using SPM2 software (Wellcome Department of Imaging Neuroscience, London, U.K.). The echo planar-imaging images were realigned spatially, normalized to the Montreal Neurological Institute template provided in SPM2, smoothed spatially with a 12-mm isotropic Gaussian kernel, and filtered temporally with a band-pass filter of low-frequency cut-off period of 256 s. Global changes in activity were removed by proportional scaling. The data was realigned in time with sinc interpolation, and in space with spatial normalization. The experiment was event-related with each event being modelled with a separate boxcar function, convolved with SPM2's canonical hemodynamic response function. The analysis was performed at the random effects level whereby the reliability of the measurements was assessed in relation to the between subject variance (Friston et al 1999). Each of the 12 different conditions was modelled separately in the design matrix and treated as effects of interest. The 15 % of trials in which a button response was made, and the head movement parameters, were modelled independently as effects of no interest.

The resultant parameter estimates for each regressor (at each voxel) were compared using *t* tests to determine whether there were significant differences in activation between conditions. The statistical results given are based on a single-voxel *t* statistic corresponding to $P < 0.05$ corrected for multiple comparisons (unless otherwise

stated). The co-ordinates of all activations are reported in Talairach space. The actual BOLD response of interesting voxels revealed by the above approach was assessed using an event triggered averaging approach. The raw signal was pre-processed in an identical fashion to the data used in the SPM analysis and was mean-corrected. The BOLD responses for each epoch type were then aligned to the start of each event and binned into 2.5s bins. These values were then used to calculate the average BOLD response across subjects from the beginning of the block until 30s after block-onset.

Object specific regions of interest were localised implicitly with the face specific localiser (*VS face > VS house*) and the house specific localiser (converse of the face localiser).

Psychophysiological Interaction analysis

We also investigated the neural interactions between the regions of interest (identified with orthogonal localisers as described above) and other cortical or subcortical regions through the application of psychophysiological interaction (PPI) analysis (Friston et al 1997). PPI analysis allows the detection of regionally specific responses in terms of the interaction between the input from another brain region and a psychological variable (such as occluded vs. visible objects). PPI employs one regressor representing the deconvolved activation timecourse in the regions of interest, which constitutes the physiological variable, a second regressor representing the psychological variable (e.g. a regressor for *OC face vs. VS face*) and a third regressor representing the cross product of the previous two (the psychophysiological interaction term). An SPM can then be computed to reveal areas where activation is predicted by the psychophysiological interaction term, treating the physiological and psychological regressors as effects of no interest. We therefore extracted the deconvolved time course of activity in the regions of interest (5mm radius spheres centred on hottest voxel from the face and house localisers). The product of each activation timecourse and the psychological variable term was calculated to create the psychophysiological interaction term. PPI's were carried out for each ROI in each subject and entered into a random effects group analysis (thresholded at $P < 0.001$ and a cluster size of 5 voxels).

6.2 Results: Experiment 1

The experimental aim was to compare brain activity produced by a visual stimulus that is perceived (*VS*) and the same stimulus when it is not perceived directly because it is occluded (*OC*). We were especially interested in looking for common activations produced by the two conditions. The comparison is not straightforward, however. To achieve realistic occlusion, the stimulus needs to be visible before it is occluded. *OC* stimuli are therefore necessarily compound stimuli consisting of visible and occluded components. Because of this, it is not valid to simply compare the *OC* conditions with the *VS* conditions. Given the slow time-course of the BOLD response it was potentially problematic to distinguish, for instance, brain activity corresponding to the period in which the stimulus is occluded from that corresponding to the initial period in which it was visible. We therefore controlled for the initial visibility of the *OC* stimulus by using *Vanishing* controls. Subtracting the appropriate *Vanishing* control conditions from the main conditions should give a measure of the neural activity corresponding to the permanence of the stimulus, whether visible or occluded, independently of the response to its initial visibility. The *SG* conditions were designed to control for the visual appearance of the stimuli when occluded (*OC*) without any stimulus being occluded. For example in the *OC face* condition, when the face is occluded the very top segment (2.5% of the surface area) of the head is still visible. The *SG face* condition therefore acts as a control for any neural activity corresponding to the appearance of the top of the head. (There were equivalent segment controls for house stimuli).

Activity common to visible and occluded stimuli

To investigate the first question, of which areas were commonly active in *VS* and *OC* conditions (whilst controlling for initial exposure), we performed a conjunction analysis using SPMs of the minimum T-statistic over 2 orthogonal contrasts. Inference was based on $p < 0.05$ adjusted for the search volume using random field theory (Nichols et al 2005, Friston et al 2005). The null distribution for the minimum statistic is described in Friston et al (2005). This enabled us to infer a conjunction of multiple effects at significant voxels.

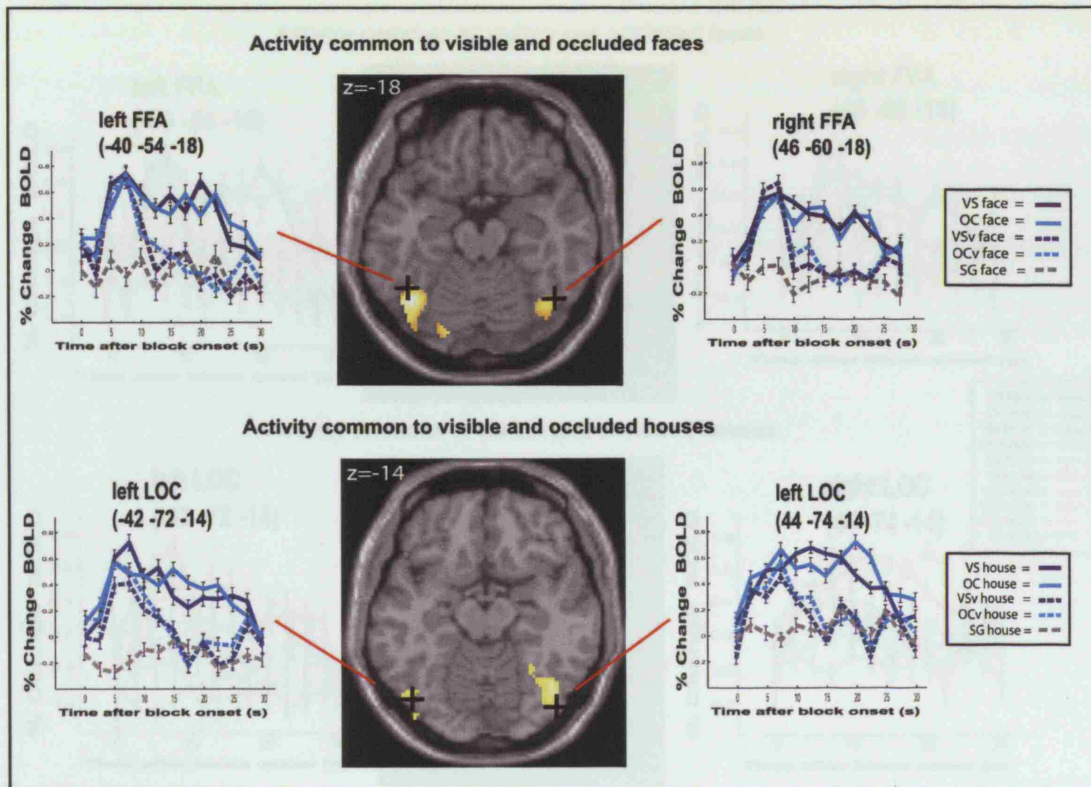


Figure 31A. Central column depicts SPM images of brain activity common to visible and occluded faces (above) and houses (below). Top brain: Voxels which show significant effects for both visible AND occluded faces expressed in the conjunction $[OC - OCv]$ AND $[VS - VSv]$ (for face stimuli only). Common activity was restricted bilaterally to the Fusiform gyrus. Bottom brain: Voxels that show significant effects for both visible AND occluded houses expressed in the conjunction $[OC - OCv]$ AND $[VS - VSv]$ (for house stimuli only). Common activity was restricted bilaterally only to the Lateral Occipital Cortex LOC. Timecourse data extracted from hottest voxel of regions of interest (defined by face and house localisers, marked on brain with black crosses) for multiple conditions. Plots show the average BOLD response across all subjects aligned to the start of each event from the local maxima of the regions of interest. The error bars show the random effects standard errors across subjects. The pattern of activity for FFA and LOC is qualitatively similar; both display sustained responses to the visible and occluded objects (for which they are selective) which is conditional on the object remaining present throughout the trial. For vanishing trials the FFA and LOC show a more transient response. For segment trials there is little response at all.

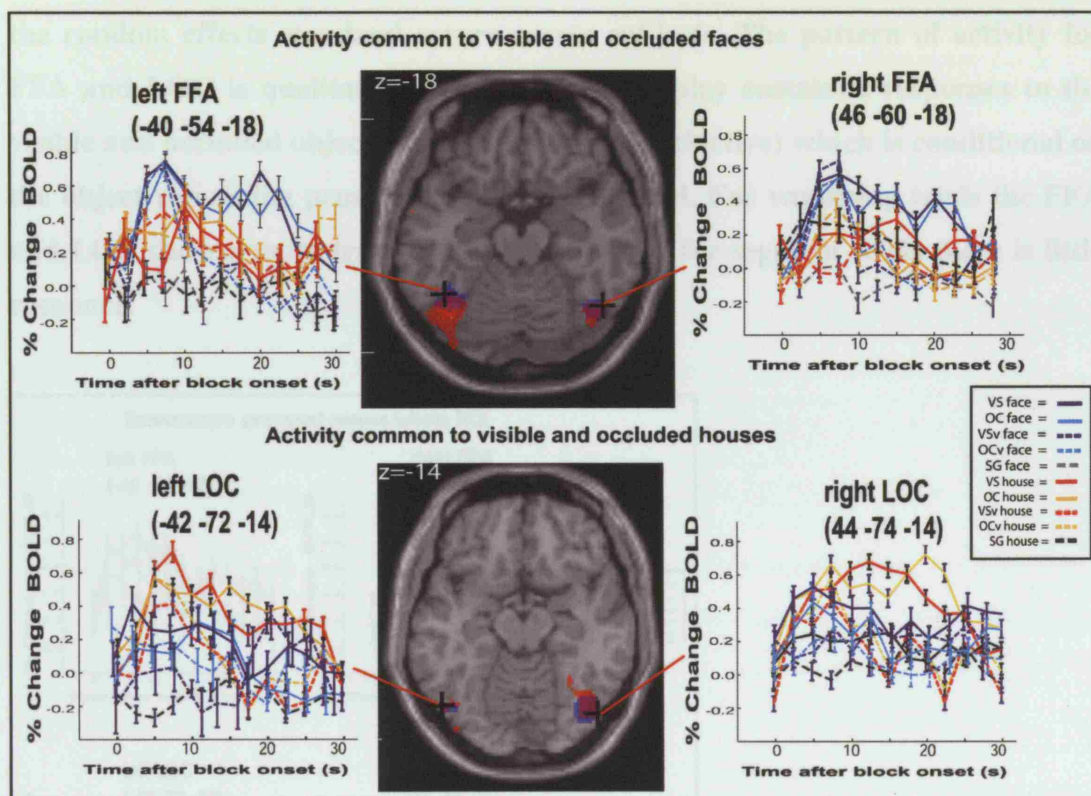


Figure 31 B. Central column depicts SPM images of brain activity common to visible and occluded faces (above) and houses (below). Top brain: Red voxels show significant effects for both visible AND occluded faces expressed in the conjunction $[OC > OCv]$ AND $[VS > VSv]$ (for face stimuli only). Blue voxels show significant effects for the Face specific localiser (VS faces $>$ VS houses). Purple voxels show overlap between conjunction and localiser. Common activity was restricted bilaterally to the Fusiform gyrus. Bottom brain: Red voxels show significant effects for both visible AND occluded houses expressed in the conjunction $[OC > OCv]$ AND $[VS > VSv]$ (for house stimuli only). Blue voxels show house specific localiser (VS houses $>$ VS faces). Purple voxels show overlap between conjunction and localiser. Common activity was restricted bilaterally to the Lateral Occipital Cortex LOC. Timecourse data extracted from hottest voxel (marked on brain with black crosses) defined by orthogonal face and house localisers. Since this data was spatially smoothed the hottest voxel should be considered not as taken a single voxel but as an ROI averaged across a Gaussian kernel (12mm FWHM) centred on the hottest voxel (Friston et al 2006) (see figure 3 for data averaged across the whole ROI). Timecourse plots show the average BOLD response across all subjects aligned to the start of each event from the local maxima of the localised regions of interest. The error bars show

the random effects standard errors across subjects. The pattern of activity for FFA and LOC is qualitatively similar; both display sustained responses to the visible and occluded objects (for which they are selective) which is conditional on the object remaining present throughout the trial. For vanishing trials the FFA and LOC show a more transient response whilst for segment trials there is little response.

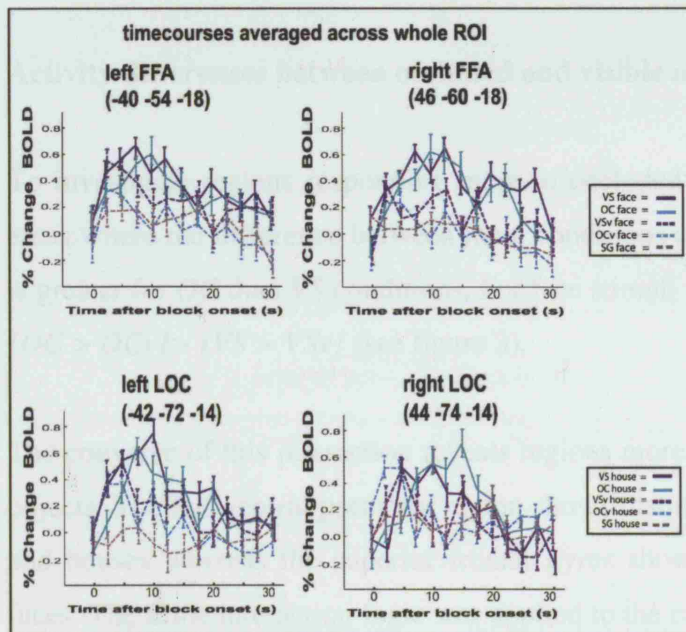


Figure 31C. Timecourses shown averaged across the whole ROI as defined by face and house specific localisers. The timecourses of the whole ROI show a similar but noisier response compared to the hottest voxel analysis, particularly for right FFA and right LOC, however the results are less clear cut for left FFA or left LOC.

To locate voxels that show significant effects for both *VS* AND *OC* faces we performed a face specific conjunction of [*OC* – *OCv*] AND [*VS*– *VSv*] comparisons. This revealed that bilaterally, the fusiform gyrus was significantly more active in main conditions than in *Vanishing* controls (See Figure 31) for both *VS faces* and *OC faces*. The face localiser contrast identified the location of the peak FFA response (left FFA=-40 -54 -18, right FFA= 46 -60 -18) to be in close proximity to the peak response revealed in the conjunction for visible and occluded faces (left FFA=42 -66 -18, right FFA= -38 -62 -20). The same logic was applied to finding regions

responding to visible AND occluded houses. The Lateral Occipital (LO) region was significantly more active in main conditions than in *Vanishing* controls for both *VS house* and *OC house* stimuli.

To reveal areas commonly activated by *OC face* AND *OC house* we performed the occlusion specific conjunction analysis: [*OC face* – *OCv face*] AND [*OC house* – *OCv house*]. This revealed that left ventral premotor cortex (see figure 4) was significantly more active during *OC* conditions than any other condition.

Activity differences between occluded and visible objects

To investigate regions responding more to occluded than visible objects we located areas where the difference between main conditions and *Vanishing* control conditions is greater for *OC* than *VS* conditions. For face stimuli we performed the interaction: [*OC* > *OCv*] > [*VS* > *VSv*] (see figure 3).

The converse of this interaction reveals regions more active for visible than occluded objects. The left ventral premotor region shows distinct specificity for occluded faces and houses whereas the superior frontal gyrus shows specificity only for occluded faces. The same interaction logic was applied to the category of houses. The premotor cortex showed specificity for both *OC face* and *OC house* stimuli, whereas the left prefrontal was active for *OC house* but not *OC face*.

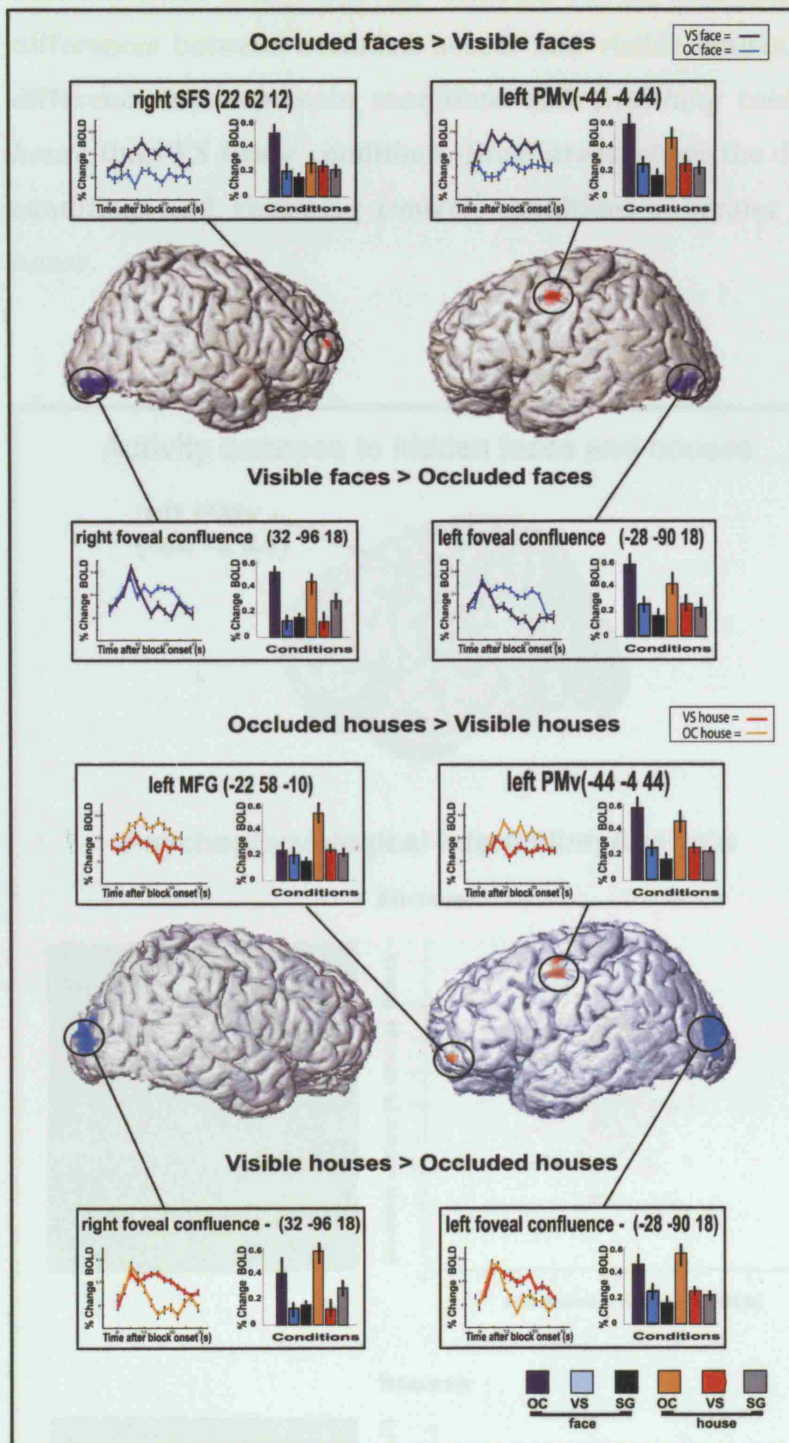


Figure 32. Upper brains show activity differences between occluded faces and visible faces. Red: areas where the difference between main conditions and *Vanishing* controls is greater for *OC* than *VS* conditions. Blue: areas where the difference between main conditions and *Vanishing* controls is greater for *VS face* than *OC face* conditions. Barplots show difference in parameter estimates

between main and *Vanishing* controls for all conditions: Lower brains show differences between occluded houses and visible houses. Red: areas where the difference between main conditions and *Vanishing* controls is greater for *OC house* than *VS house* conditions. Blue: areas where the difference between main conditions and *Vanishing* controls conditions is greater for *OC house* than *VS house*.

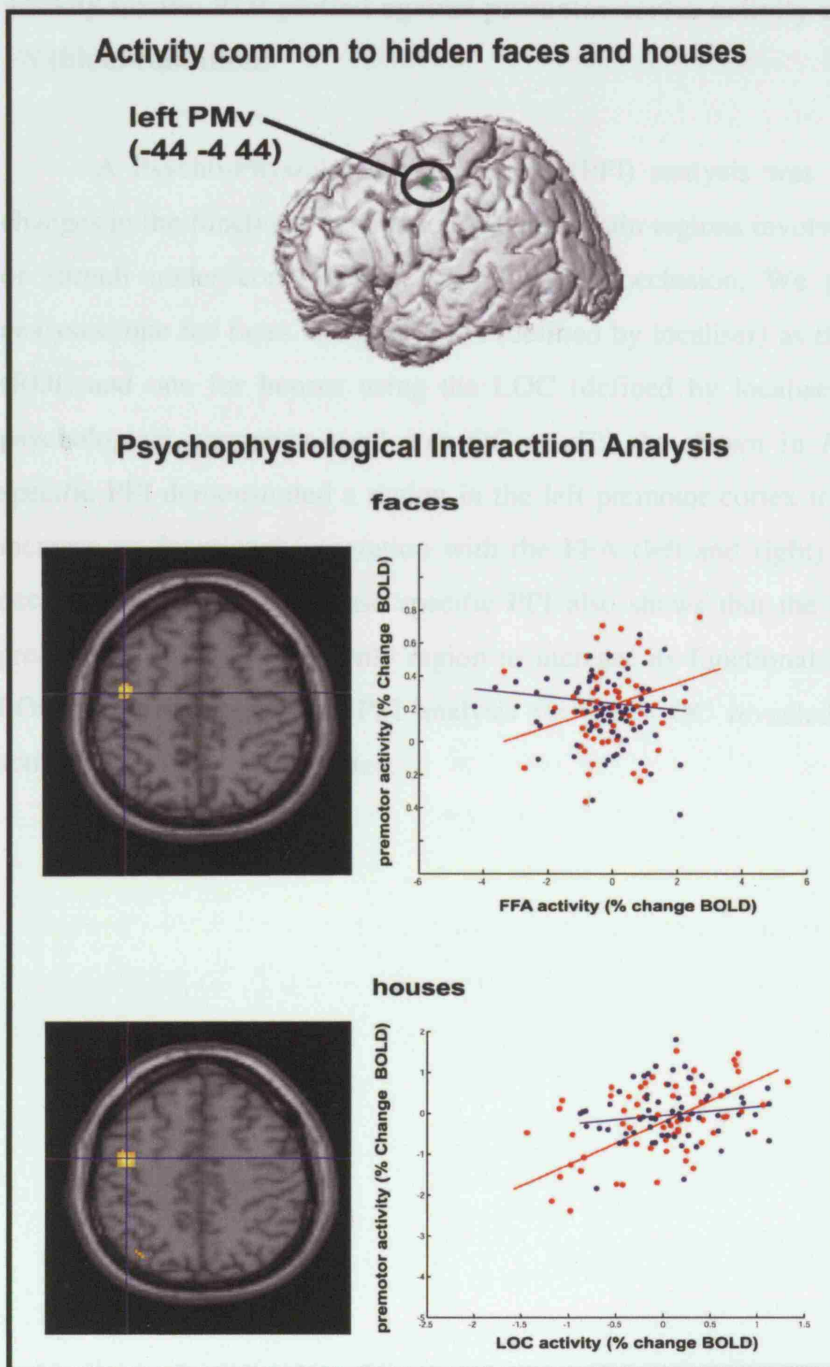


Figure 33. Upper diagram shows brain activity common to occluded faces and houses. Voxels (Green) that show significant effects for both *OC face* AND *OC house* expressed in the conjunction [*OC face* – *OCv face*] AND [*OC house* – *OCv house*]. Common activity was restricted to Premotor Cortex. This shows that premotor cortex response is significantly more active for occluded faces and houses than visible faces and houses. 2nd and 3rd row of figure shows Psychophysiological interactions. Left images show the SPM computed for the Psychophysiological interaction term. Right graphs show the extracted BOLD activity for the ROI plotted against premotor cortex activity under *OC* (red) and *VS* (blue) conditions

A Psycho-Physiological Interaction (PPI) analysis was employed to reveal changes in the functional interaction between brain regions involved in the processing of stimuli under conditions of visibility and occlusion. We performed two PPI analyses; one for faces using the FFA (defined by localiser) as the region of interest (ROI) and one for houses using the LOC (defined by localiser) as the ROI. The psychological parameter used was *OC vs. VS*. As shown in *Figure 33*, the face specific PPI demonstrated a region in the left premotor cortex to be the only one to increase its functional integration with the FFA (left and right) when the face was occluded. Similarly the house specific PPI also shows that the same region of left premotor cortex to be the only region to increase its functional integration with the LOC (left and right). The PPI analysis for *VS > OC* revealed no suprathreshold activity at $P < 0.001$ uncorrected.

6.3 Methods Experiment 2

6.3.1 Psychophysics

Purposes

Psychophysical testing was performed for different purposes. For each subject there was a parameter-setting sessions using a staircasing procedure such that contrast could be set so that performance was within the required intervals for the short delay condition. Once contrast was set we characterised the effect of different delay durations on performance. Subject specific delays were set such that performance was within required intervals for medium and long delays. Similar setting protocols were employed at the start of each scanning session to ensure subject performance met the required standard.

Stimuli and task

At all times throughout the experiment a central fixation cross was present upon a grey background. At the onset of each trial a disc array, consisting of 8 discs, appeared for a duration of 200ms, randomly positioned at any 8 of 16 radial positions of in a circular arrangement. Each disc was defined by a sinusoidal change in luminance over time as depicted in figure 34. This is best described as an initial decrease in luminance relative to background, from grey to black, followed by an increase, from black to white, followed by a final return to grey. Thus the discs were defined by salient luminance transients of short duration. All discs were presented at an eccentricity of 3° (parafoveal since radius of fovea is 2.6°) of visual angle at a distance 50 cm outside of the scanner and 63 cm inside the scanner.

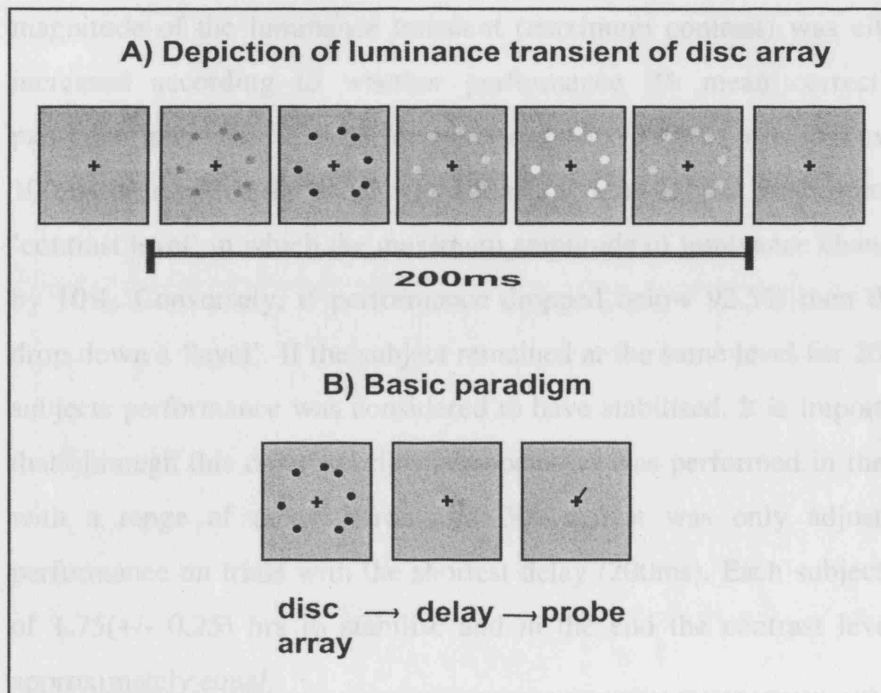


Figure 34. A) Schematic depiction of a disc array defined by a 200ms duration luminance transient. Subjects were instructed to fixate the cross at all times throughout the experiment. B) Depicts the basic structure of a trial. Following the cueing stage the subject makes their response. Trials are separated by a random duration jitter period between the subject's response and the onset of the next trial.

A central cue was then flashed up at the location of the central cross for 100ms. The cue extended to 1° of visual angle from fixation. The orientation of the cue pointed pseudo-randomly to one of 16 disc positions indicating which position the subject was to base their report on (as described previously, the partial report method Sperling 1960). The subject reported via button press whether the disc was present or absent.

Psychophysical training phase

The aim of the psychophysical training phase was to set the contrast parameters for each subject such that performance for the short delay was within the required range. For the training stage the cue was always presented at a delay of between 100ms and 3 s after the *offset* of the disc array. Performance was monitored throughout and the

magnitude of the luminance transient (maximum contrast) was either decreased or increased according to whether performance (% mean correct) deviated from particular intervals. If the performance exceeded 97.5% correct on average for the last 100 trials in which the delay was 200ms then the subject would progress to the next 'contrast level' in which the maximum amplitude of luminance change was decreased by 10%. Conversely, if performance dropped below 92.5% then the subject would drop down a 'level'. If the subject remained at the same level for 2000 trials then the subjects performance was considered to have stabilised. It is important to emphasise that although this contrast adjusting operation was performed in the context of trials with a range of delays (from 100-3000ms) it was only adjusted according to performance on trials with the shortest delay (200ms). Each subject took an average of 1.75(+/- 0.25) hrs to stabilise and in the end the contrast levels reached were approximately equal.

Subjects were asked on a scale of 1 (not at all sure that there was a flashed array of discs): 10 (absolutely certain there was a flashed array of disc) how certain they were that they saw each disc on any trial for the shortest cue delay, in the context of blank catch trials in which subjects were cued without being presented with a disc array. All 3 subjects reported 10, indicating that the contrast values that were adjusted to be salient events of very high visibility. The same ratings were given whether the whole array or just a single disc was presented. If cued at the time of onset the task appeared to be trivially easy with all subjects being able to perform at a level of 100% over 100 trials, indicating that the task is easy if cued at the time of onset. An average decrease of ~5% in performance then is incurred on average by cueing after 200ms.

	He	Wy	Cg
Visibility rating zero delay	10	10	10
Visibility rating 200ms delay	10	10	10
Time to stabilise (to nearest 30mins)	1.5hrs	1.5hrs	2hrs
Mean performance at 0ms	100%	100%	100%
Mean performance at 200ms	94.3%	96.4%	92.5%

Table 1 shows performance values, visibility ratings and stabilisation times for the different delay conditions in the contrast calibration phase.

Psychophysical characterisation phase

The aim of the psychophysical characterisation phase was to characterise the performance-delay decay function such that the correct delay timings could be selected for each subject to ensure performance for the middle and long delay was within the required intervals and approximately equal. The secondary aim was to ensure that performance had stabilised and that no further learning was taking place. It was important that performance had stabilised since it would be potentially problematic to have to adjust stimulus parameters by any large degree in the scanner to control performance.

The psychophysical characterisation phase involved the same paradigm and same subject-specific contrast parameters attained in the previous phase of testing. This involved 2000 further trials, which allowed accurate plotting of performance-time decay function that allowed an assessment of how performance decreases when the

delay of the cue is increased. By looking at the performance decay functions for all 3 subjects one could chose subject specific delay values for the medium at which performance was between 72.5% and 77.5% accuracy. Subjects were asked before proceeding to the next stage to rate on the same scale of 1:10 how visible each disc array onset was for all delays. Visibility was again rated as 10 for all subjects with no effect of delay.

	He	Wy	Cg
Visibility rating 200ms delay	10	10	10
Visibility rating 1200ms delay	10	10	10
Visibility rating 2200ms delay	10	10	10
Mean performance at 200ms	96.1%	93.2%	94.3%
Mean performance at middle duration delay	74%	76.5%	72.5%
Middle duration selected (ms)	1000	1100	900
Mean performance at 2200ms	56%	53%	58%

Table 2 shows visibility ratings and performance values for different delay values

Each subject performed 10 sessions of 600 trials making a total of 6000 trials. Performance at all delays was stable in time between sessions.

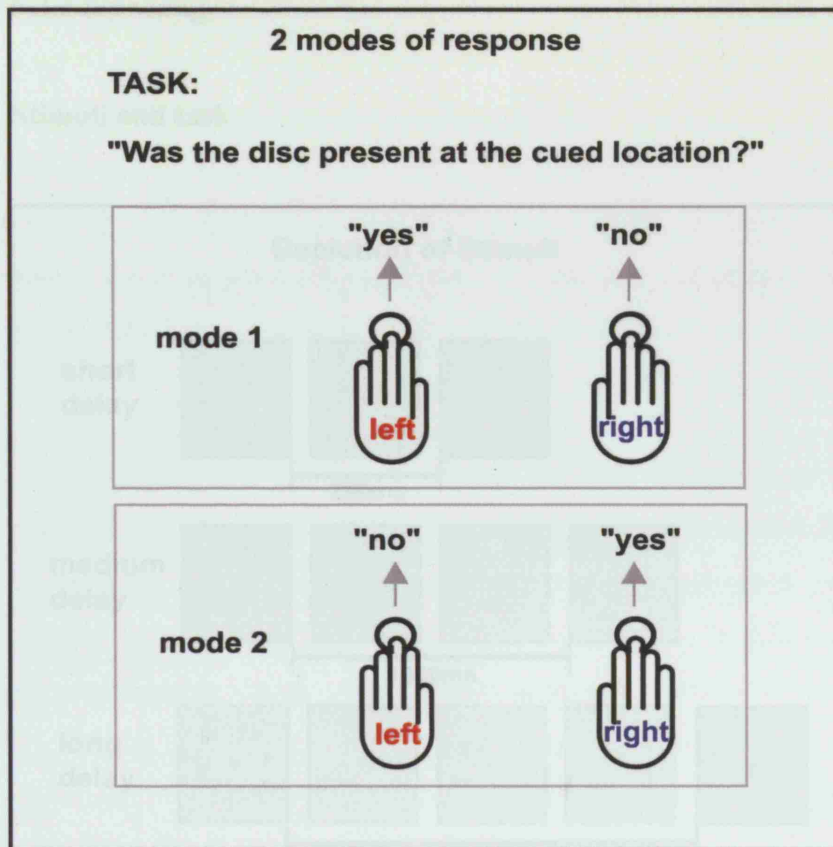


Figure 35 depicts the two response modes, each used to answer the same question of whether the disc is present or not at the cued location.

Every session the 'meaning of the report' ('present' or 'absent') was switched such that each of the two hands was re-assigned to report the opposite percept (see figure x). For instance for the first set the subject would report 'present' with the left hand and 'absent' with the right, and for the second set this would be switched (left=absent, right=present), and so on, switching between sets in a pseudorandomised order. This was to control for any possible bias between the hand used to report and the actual meaning of the report (e.g. subjects are more likely to report 'present' with left than right hand) and it would allow one to dissociate the motor act of reporting from the perceptual decision (or meaning of the report, 'present' or 'absent') in the subsequent fMRI experiment. Subjects were trained and tested for 6 hours each prior to scanning on the variable delay task.

6.3.2 Scanning

Stimuli and task

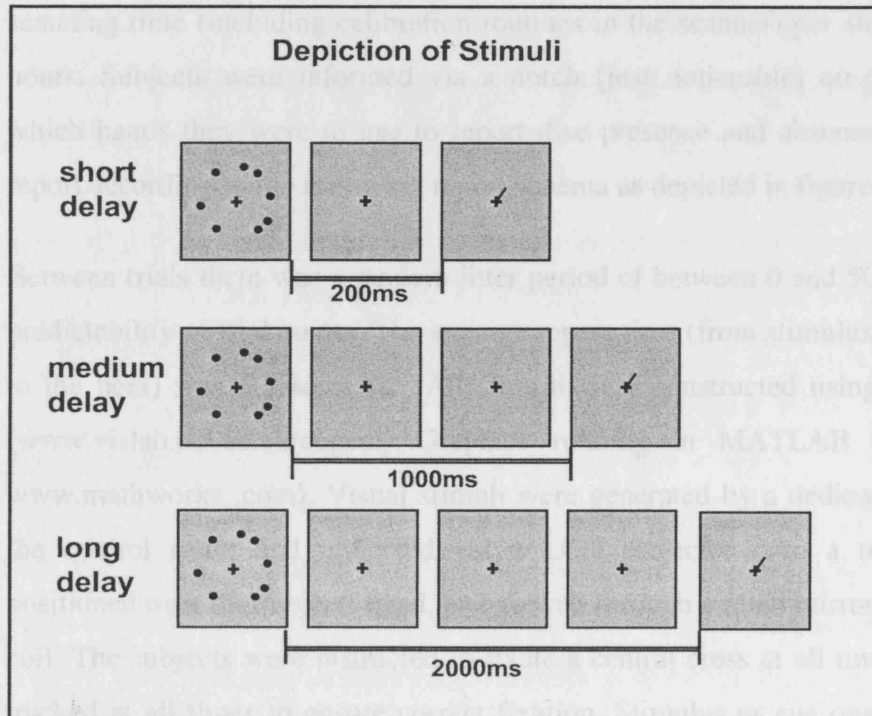


Figure 36. The psychophysical task for the fMRI experiment was identical to that outside scanner with the only modification being that the cue delay is only one of three values (short, medium or long) as opposed to a continuous variable.

The same stimuli were used as in the psychophysical characterisation (see above) with the exception that only three delays were used, short (200ms) medium (900-1100ms) and long (2200ms). The contrast levels for the disc array were recalibrated at the beginning of every session to account for differences in the display system used in the scanner, and any learning effects that may have occurred. The same training program was used to adjust the contrast levels of the discs such that performance was approximately 95% (+/- 2.5%) accuracy at the short delay. The staircasing procedure was applied to the entire experimental session in the scanner, only adjusting contrast to maintain performance between 92.5% and 97.5% for the short delay only. This was to clamp performance to a stable level by accounting for natural performance variations within a session, e.g. due to tiredness. The starting contrast for each session

was identical for the beginning of each session suggesting learning effects had reached ceiling. The experiment was split into 8 sessions, each consisting of 600 trials, making a total of 4800 trials per subject. Each session was approximately 30 min in duration (not exact due to random jitter and variable response times) with total scanning time (including calibration routines in the scanner) per subject exceeding 6 hours. Subjects were informed via a notch (just noticeable) on the fixation cross which hands they were to use to report disc presence and absence. Subjects would report according to the following report schema as depicted in figure 35.

Between trials there was a random jitter period of between 0 and 500msec to prevent predictability of trial onsets. The average repeat time (from stimulus onset of one trial to the next) was 3.5 seconds. All stimuli were constructed using COGENT 2000 (www.vislab.ucl.ac.uk/cogent) Graphics running in MATLAB (Mathworks inc. www.mathworks.com). Visual stimuli were generated by a dedicated PC located in the control room and projected via a LCD projector onto a transparent screen positioned over the subjects head, and viewed through a tilted mirror fixed to the head coil. The subjects were instructed to fixate a central cross at all times and were eye-tracked at all times to ensure correct fixation. Stimulus or cue onset did not induce observable saccades as indexed by an in-scanner eyetracker. Subjects were instructed not to explicitly attempt to memorise any particular location or set of locations nor use any mnemonic strategies.

Subjects

Three subjects who were all experienced as psychophysical observers and fMRI subjects (2 males 1 females, mean age 23 years, all right handed) with normal vision were scanned. The inclusion criterion for all subjects was an ability to maintain correct fixation for the duration of each session (OH was excluded on this basis). All gave informed consent in accordance with the Declaration of Helsinki, and the Ethics Committee of the National Hospital for Neurology and Neurosurgery, London, UK, granted ethics approval for the study.

Scanning details

Scanning was performed using a 3T Siemens Allegra fMRI scanner with a head-volume coil (Siemens, Erlangen, Germany). A gradient echo-planar imaging (EPI) sequence was used to maximise blood oxygen level dependent (BOLD) contrast (TE = 65ms, repeat time 2.47s). Each brain image was acquired in a descending sequence comprising 38 axial slices each 2mm thick with 1mm gaps in-between, consisting of 64 x 64 voxels, giving nearly whole brain coverage. There were 8 sessions, each consisting of an average of 730 volumes with the first 5 volumes being discarded to allow for T1 equilibration effects. All subjects had been scanned previously with a T1 weighted structural sequence.

Data analysis

We analysed the data using SPM2 software (Wellcome Department of Imaging Neuroscience, London, U.K., www.fil.ucl.ac.uk/spm). The echo planar-imaging images were realigned spatially, normalized to the Montreal Neurological Institute template provided in SPM2. Due to spatial resolution issues pertaining to the induced responses of small regions of visual cortex it was necessary to analyse spatially unsmoothed data. Since the use of Gaussian field theory to adjust statistical thresholds assumes spatially smooth data we supersampled all functional images to 1mm cubic voxels. All data was filtered temporally with a band-pass filter with a low-frequency cut-off period of 128 s. Global changes in activity were removed by proportional scaling. The data were "realigned in time" by using sinc interpolation before spatial normalization. This temporal realignment adjusted the whole-brain images to approximate those that would have been obtained if the whole image had been acquired instantaneously, at a time halfway through the acquisition of the volume (scanning from top to bottom). The experiment was event-related in which each event was modelled as a compound event with a stick function and convolved with SPM2's canonical hemodynamic response function. The analysis was carried out as case studies at the individual subject level. The paradigm was a 4 way factorial design, consisting of signal detection category (4 levels), cue location (16 levels), report mode (2 levels) and cue delay (3 levels). Additionally we specified a regressor to model presence of the 'uncued' discs (locations of discs which were not cued) at each of the

16 locations and two regressors to model the two button presses. The co-ordinates of all activations are reported in MNI co-ordinates. A linear 1st order parametric modulator was specified to model the variations in contrast deviation within sessions

6.4 Results Experiment 2

In this section I report the results from experiment 2. The psychophysical results consist of the psychophysics performed inside the scanner, which was used to ensure that the appropriate performance levels were maintained during the acquisition of fMRI data. As is often the case with fMRI studies, the analysis of the results can be complex, especially for this experiment given the design (it is a 4-way full factorial design with 384 factor combinations). To summarise: The experiment performed was a case study involving the subjects used in the preliminary psychophysical studies. This means that any results obtained can only extend to the subjects studied here and cannot strictly be extrapolated to the general population. The fMRI results are analysed in order to dissociate the neural correlates of stimulus processing, attentional selection, decision-making and report using a factorial design, manipulating cue delay, report mode and cued position with post-hoc sorting of data according to signal detection categories. The aim of the analysis is to look specifically at the medium delay condition and to use formulations of signal detection profiles as formal hypotheses for dissociating the neural correlates the components underlying stimulus reportability.

Psychophysics inside the scanner

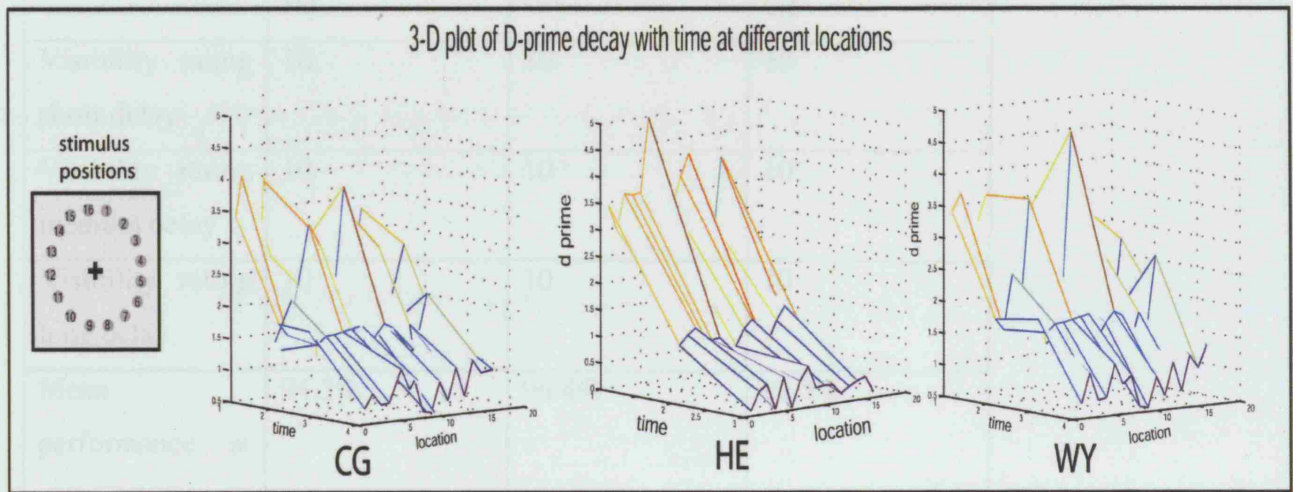


Figure 37. 3-d mesh plot of the performance defined by d-prime with cued location and delay as experimental variables. Inset in the left-hand diagram is a reminder of the locations in stimulus space of each position number.

Table 3 shows performance and stability ratings for the different delay conditions of in-scanner psychophysics.

We plotted the d' rating of each subject against the delay for each stimulus position to reveal the decay in performance with increased cue delay. As described previously, the task of this type, the decay appears to be approximately exponential in nature (Laird et al. 2005). As can be seen, at the shortest delay of 200ms d' is high for all subjects which is expected given the fact that contrast parameters were adjusted to hit accuracy of ~95% for the shortest delay. As can be seen the d' approaches but does not reach zero after 2 s.

	He	Wy	Cg
Visibility rating short delay	10	10	10
Visibility rating medium delay	10	10	10
Visibility rating long delay	10	10	10
Mean performance at short delay	94.3%	96.4%	92.5%
Mean performance at middle delay	76.6%	73.2%	74%
Mean performance at long delay	58%	59%	57%

Table 3 shows performance and visibility ratings for the different delay conditions of in-scanner psychophysics.

We plotted the d-prime of each subject against the delay for each stimulus position, to reveal the decay in performance with increased cue delay. As described previously for tasks of this type, the decay appears to be approximately exponential in nature (Lu et al 2005). As can be seen, at the shortest delay of 200ms the d' is high for all subjects which is expected from the fact that contrast parameters were calibrated to set accuracy at ~95% for the shortest delay. As can be seen the d' approaches but does not reach zero after 2.2 s.

Scanning results

The experimental rationale was to analyse the fMRI data according to the predefined signal detection categories of hit (H) false positive (F) miss (M) and correct rejection (C) described above. This approach allows us to test for areas whose activity conforms to the signal detection profiles as described previously.

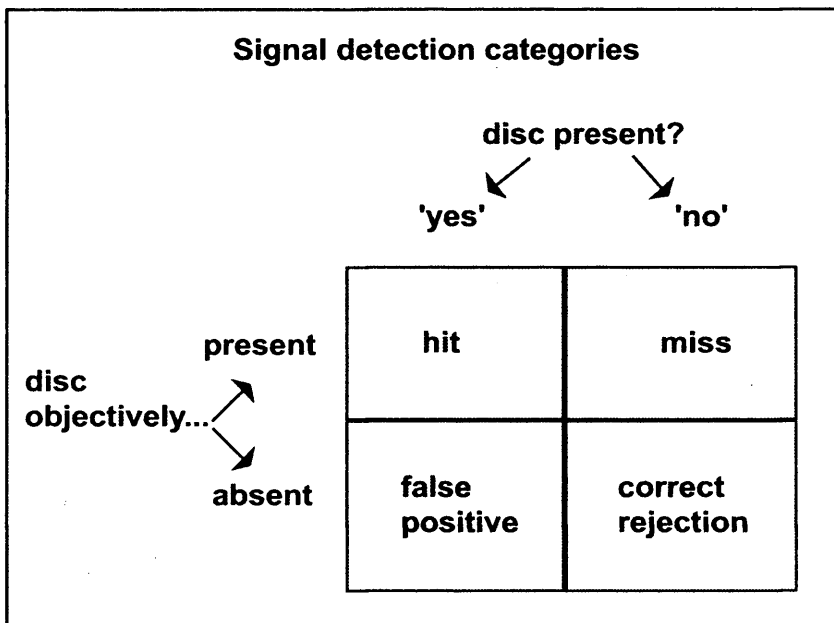


Figure 38. Depiction of the 4 signal detection categories that are labels for each possible combination of report and stimulus presence.

The hypothesis is that stimulus reportability is achieved by a convolution of operations encompassing sensory processing, decision making, attention and motor action generation. Thus stimulus reportability is expected to engage a cohort of brain areas, including sensory, decision, attentional and motor areas. The hypothesis is that according to the different roles played by each area, there should be a unique 'signal detection profile' for each. This means that they should express a distinct profile activity for the four signal detection categories; H, F, M, and C (see figure 39 below). We therefore partitioned the data according to these categories in order to look at the signal detection profiles across the areas of interest. This process of 'partitioning' involves assigning different regressors according to whether the trial was a hit, miss

etc. In doing so it is possible to create contrasts that conform to particular signal detection profiles that formalise our hypothesis testing as specified in figure 39.

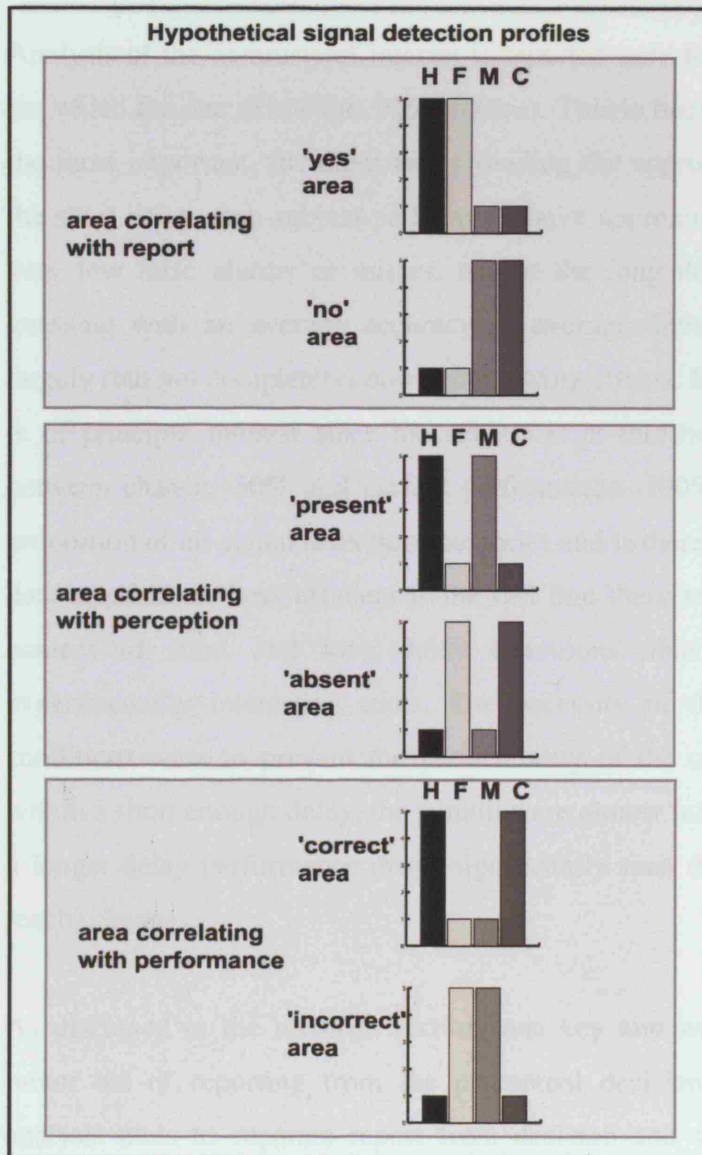


Figure 40. Depicts the parameter estimate 'fingerprints' of hypothetical areas that conform to signal detection profiles. In other words, these profiles depict the expected parameter estimates for the different signal detection categories for different hypothetical areas. For example a report area will have a greater response for trials in which the subject reports 'yes' than when they report 'no' independently of whether it was there or not.

We could then compute statistical parametric maps for each hypothetical signal detection profile, for the factors of interest. This enabled us to look for regions whose

effects matched the characteristics of the hypothetical regions (stimulus processing area, report area etc).

Analysis of the contrasts of interest is reported only for the medium delay condition (in which the cue delay was 900-1100ms). This is because the middle delay factor is the most important, the other two providing the appropriate context for the task. At the short delay each subject performed above approximately ~95% accurate, leaving very few false alarms or misses, and at the long delay each subject was mainly guessing with an average accuracy of average 58%, rendering the interpretation largely (but not completely) down to guessing effects. It is thus the medium delay that is of principle interest since the subject is at threshold, (defined here as halfway between chance -50% and perfect performance- 100% correct). This gives a good proportion of all signal detection categories and is therefore more efficient in terms of data acquisition. Less efficient is the fact that these trials must be embedded in the context of short and long delay conditions, thus diluting the proportion of experimentally interesting trials. The necessity of the short and the long delay conditions were to prevent the predictability of the cue and to prove that, if cued within a short enough delay, the stimuli were almost fully reportable and, if cued after a longer delay performance drops significantly such that it approached (but did not reach) chance.

As discussed in the methods section, one key aim was to isolate the effect of the motor act of reporting from the perceptual decision. However, even though the analysis aims to separate report from decision and stimulus processing, it is still necessary to separate the motor act of reporting from the perceptual decision that precedes it. Two report modes were therefore used to dissociate motor related activity from decision related activity. This would not be possible with the standard procedure of modelling button press responses since the same motor act would be associated with the same decision and thus the two would be inextricably linked as it is in most NCC experiments. We therefore asked subjects to make the same response *with different hands* in different experimental sessions. Different hands were used rather than different fingers to maximise the difference in brain response for the two different types of report, and to give a strong *a priori* handle (left hand right motor

cortex mapping etc) on which brain areas to expect. The modes of report are depicted in figure 41 of the methods.

Combining the two modes of report with the signal detection profiles gives one signal detection profile per mode. Since each signal detection profile relates to four response categories, this leads to signal detection profiles 8 columns long for each area of interest, as illustrated below. The figure illustrates the hypothetical profiles of our hypothesised functional areas. With the exception of the 'report area', these will resemble the signal detection profiles illustrated in figure 40 copied once and concatenated together. Because of the difference in the report area across the two report modes, one can discern a report area from the 8 columns signal detection profile that could not be discerned by just inspecting the 4-column signal detection profile. Thus the function of the 8-column profile is two-fold. Firstly it shows whether a signal detection profile is report invariant (the same pattern is repeated for each report mode); secondly, it shows the areas that differ according to the report mode, allowing us, as described above, to dissociate report from decision.

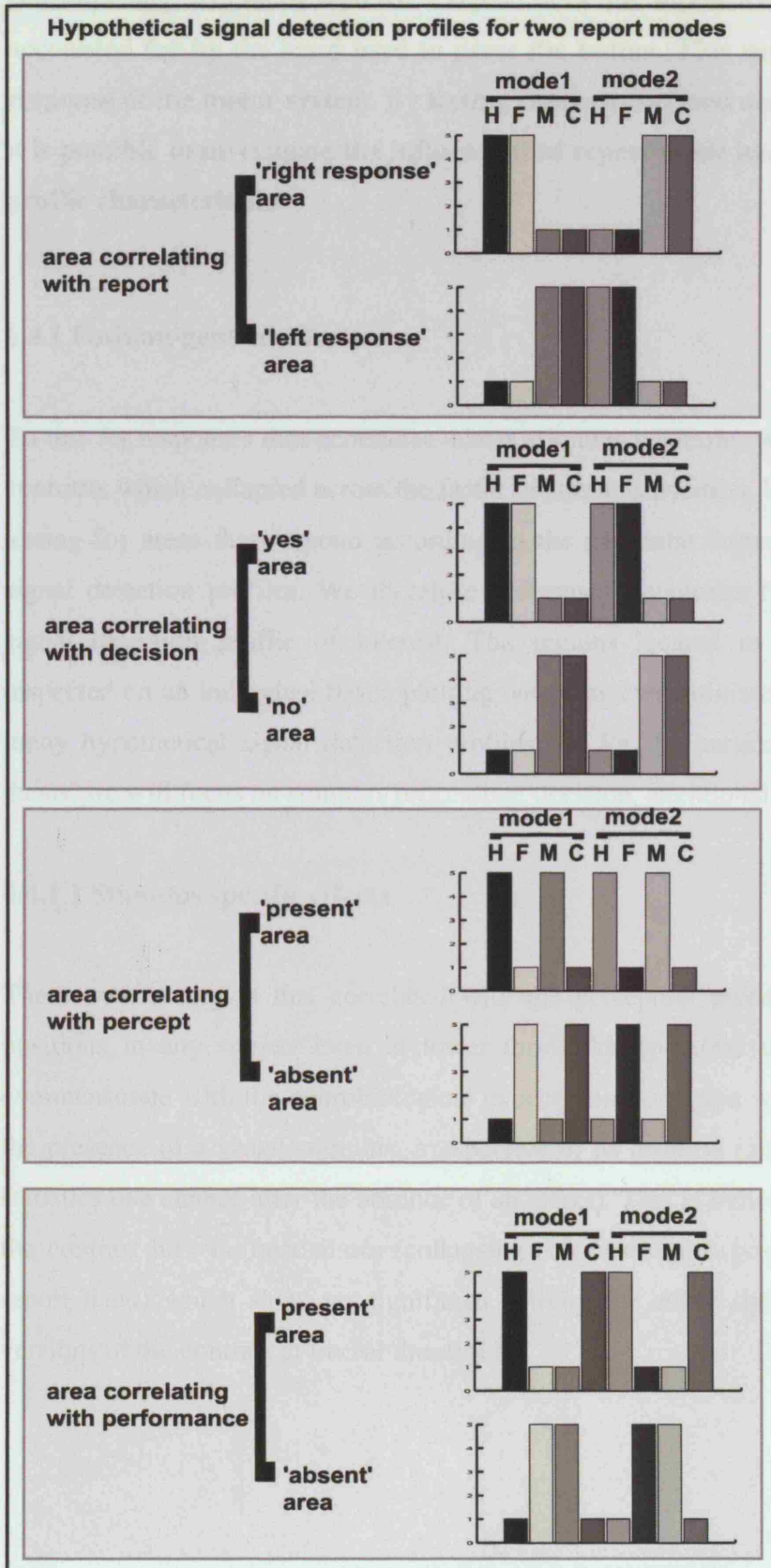


Figure 41 depicts a richer signal detection profile by splitting the profile according to report mode; this effectively doubles the profile such that it is 8

columns long. As can be seen for a report area the response of the area is entirely accounted for by the hand used to press the button. This would be the expected response of the motor system. By testing contrasts defined across the 8 conditions it is possible to investigate the influence that report mode has on signal detection profile characteristics

6.4.1 Position-general effects

To test for responses that generalise across stimulus positions, we performed a set of contrasts which collapsed across the factor of stimulus position. We were interested in testing for areas that respond according to the particular forms of the hypothetical signal detection profiles. We therefore performed t-contrasts for each hypothetical signal detection profile of interest. The regions located in each contrast were inspected on an individual basis, plotting out parameter estimates for each. There are many hypothetical signal detection profiles but for the particular purposes of this thesis, we will focus on stimulus processing, decision, attentional and report profiles.

6.4.1.1 Stimulus specific effects

There was no region that correlated with the perceptual profile invariantly for all positions in any subject even at lower thresholds ($p < 0.005$ uncorrected). This is commensurate with the neurobiological expectation no region would be selective for the presence of a visual stimulus, irrespective of its position (although with classical statistics one cannot infer the absence of an effect). This is indicated by the results of the contrast hits vs. false alarm (collapsing across stimulus position, controlling for report hand) which show no significant activity for either right hand or left hand versions of the contrast at liberal thresholds.

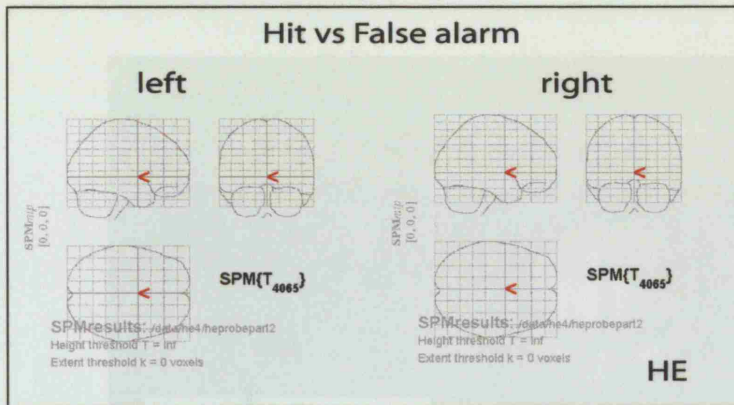
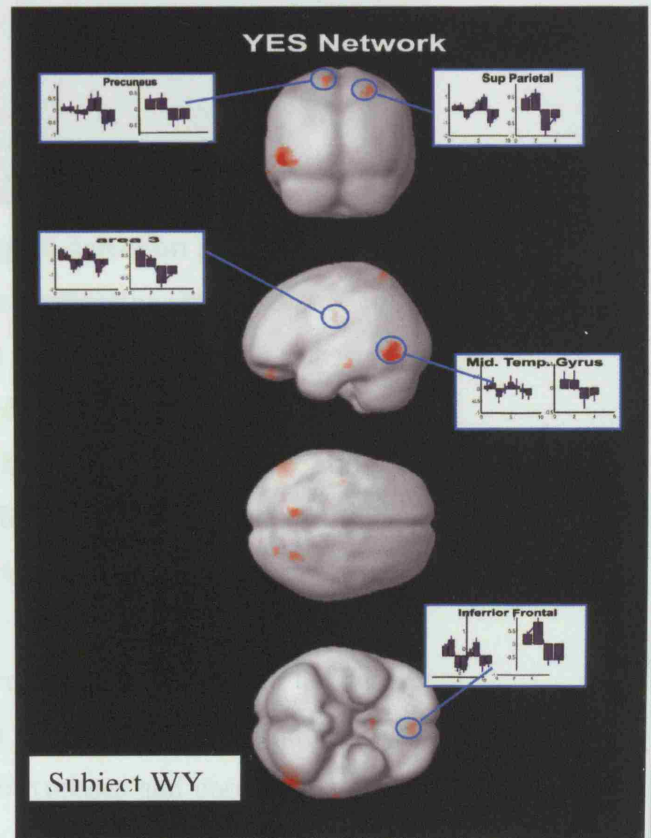
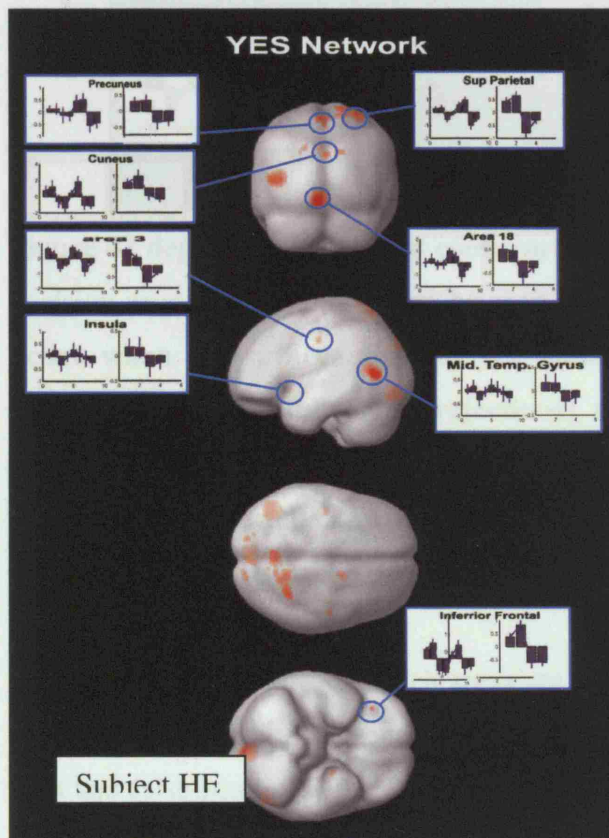


Figure 41 b figure shows no significant difference in any voxels for a liberal threshold ($p < 0.05$ uncorrected) for the difference between hits and false alarms collapsing across conditions. The same was true for all other subjects and for all other contrasts in which stimulus position was collapsed across, including present vs. absent.

6.4.1.2 Decision effects



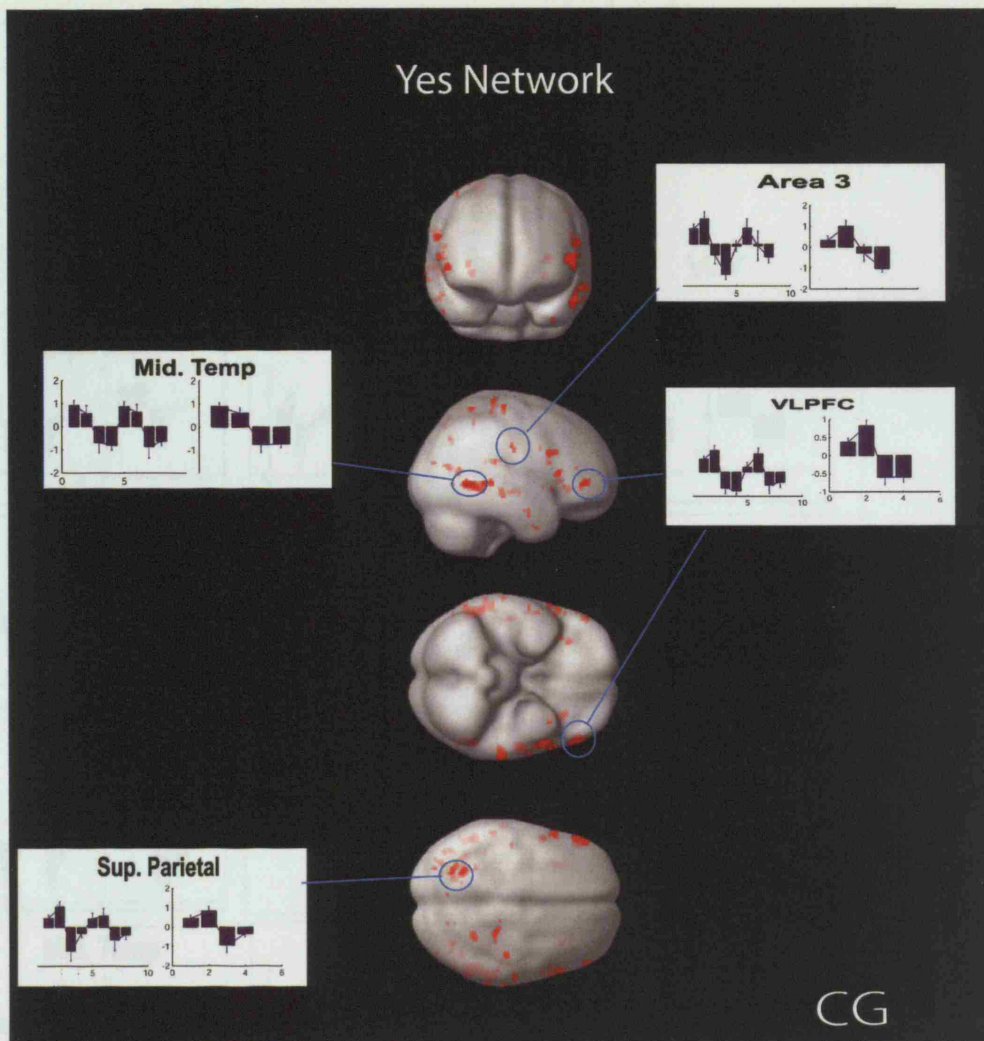


Figure 42 depicts all areas that conform to the signal detection profile for a ‘yes’ decision area. All regions as can be seen from the parameter estimates, are more active when the subject decides ‘yes’ (hits and false alarms) than when they decide ‘no’ (misses and correct rejections), regardless of report mode, regardless of stimulus presence. Since stimulus presence best determines perception (as defined by what can be reported at the shortest delay) the invariance of the ‘yes’ network to perception can be inferred from its invariance to stimulus presence.

To address the question of which areas are selectively more active when the subject decides (the answer to the question of was there a disc at the cued location) ‘yes’ (hits and false positives) than when they decide ‘no’ (misses and correct rejections), independently of stimulus processing and report, we performed a t-contrast ‘yes > no’ for the middle delay averaging across all cue positions. This revealed in all three

subjects a diverse and wide-ranging network of areas. Areas commonly activated for all three subjects included superior parietal, area 3 and middle temporal gyrus and inferior PFC.

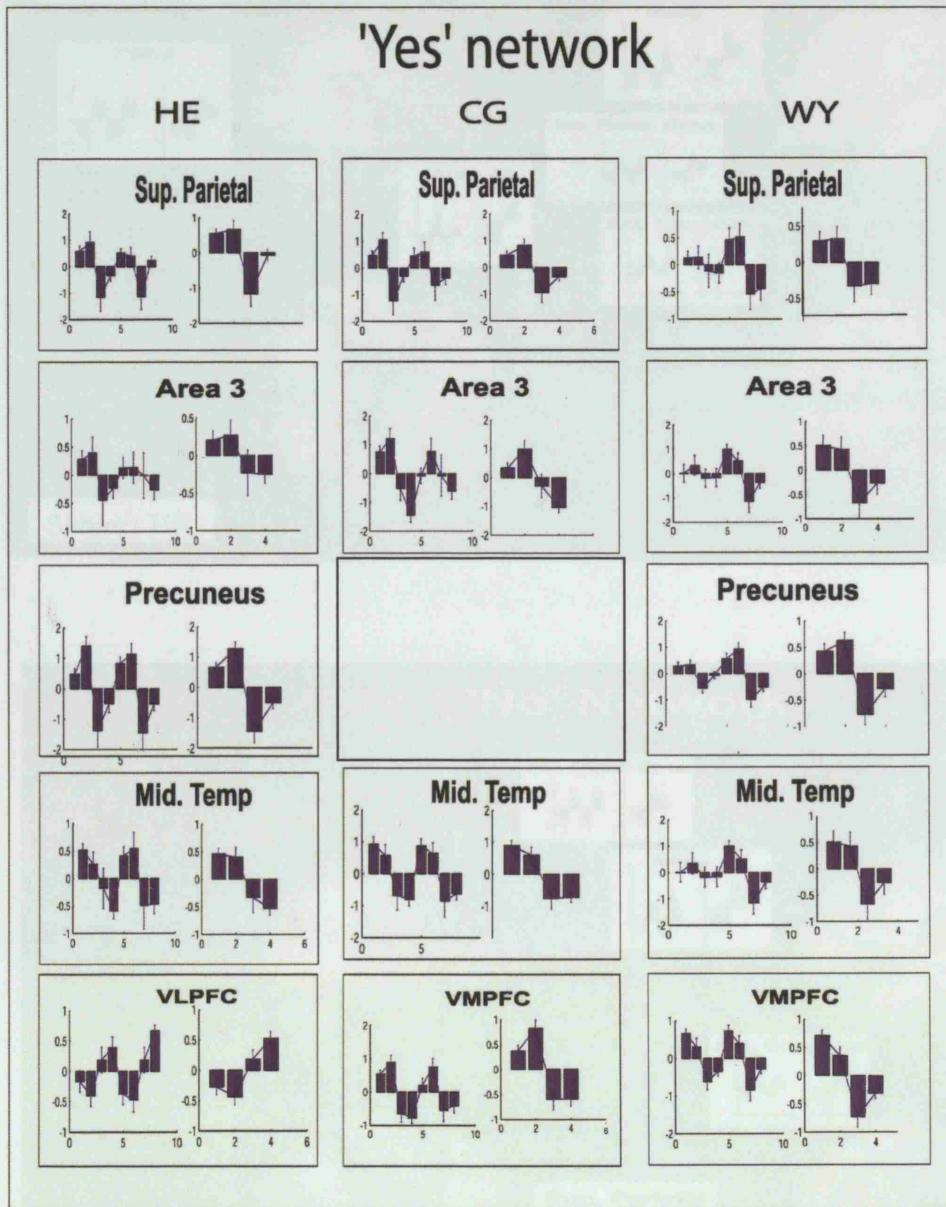
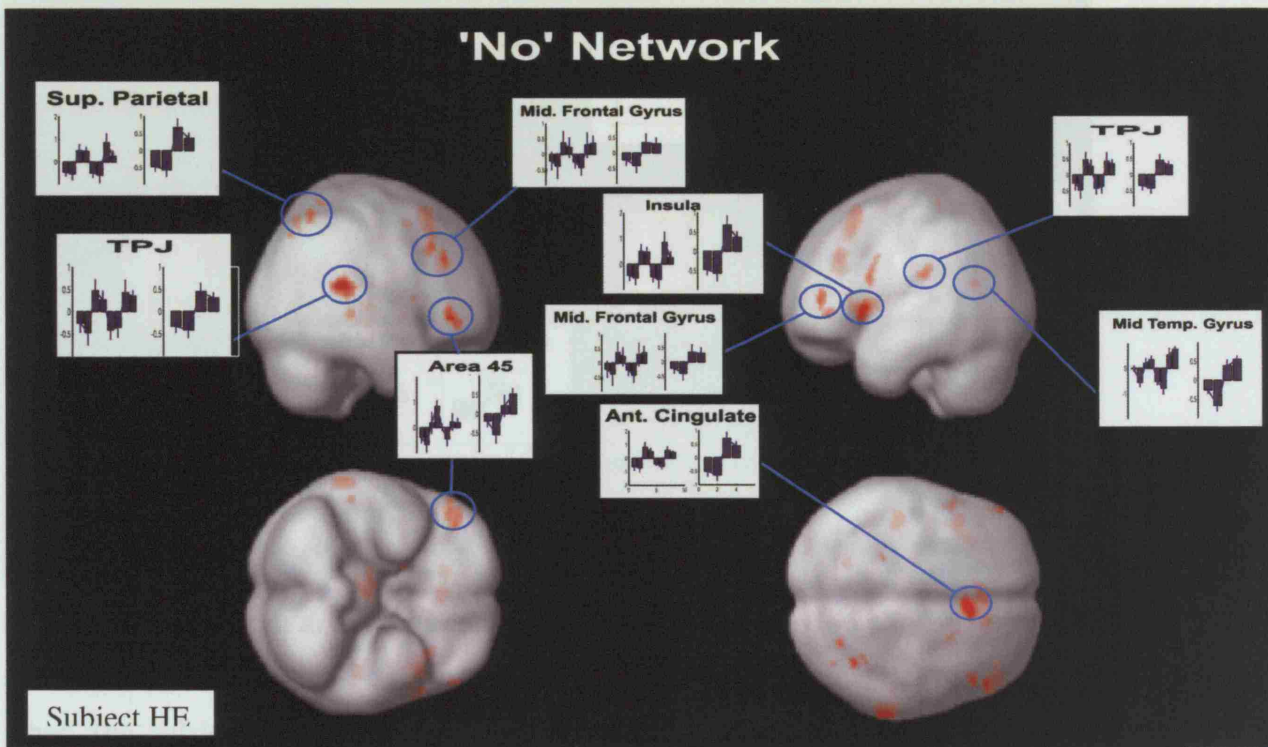
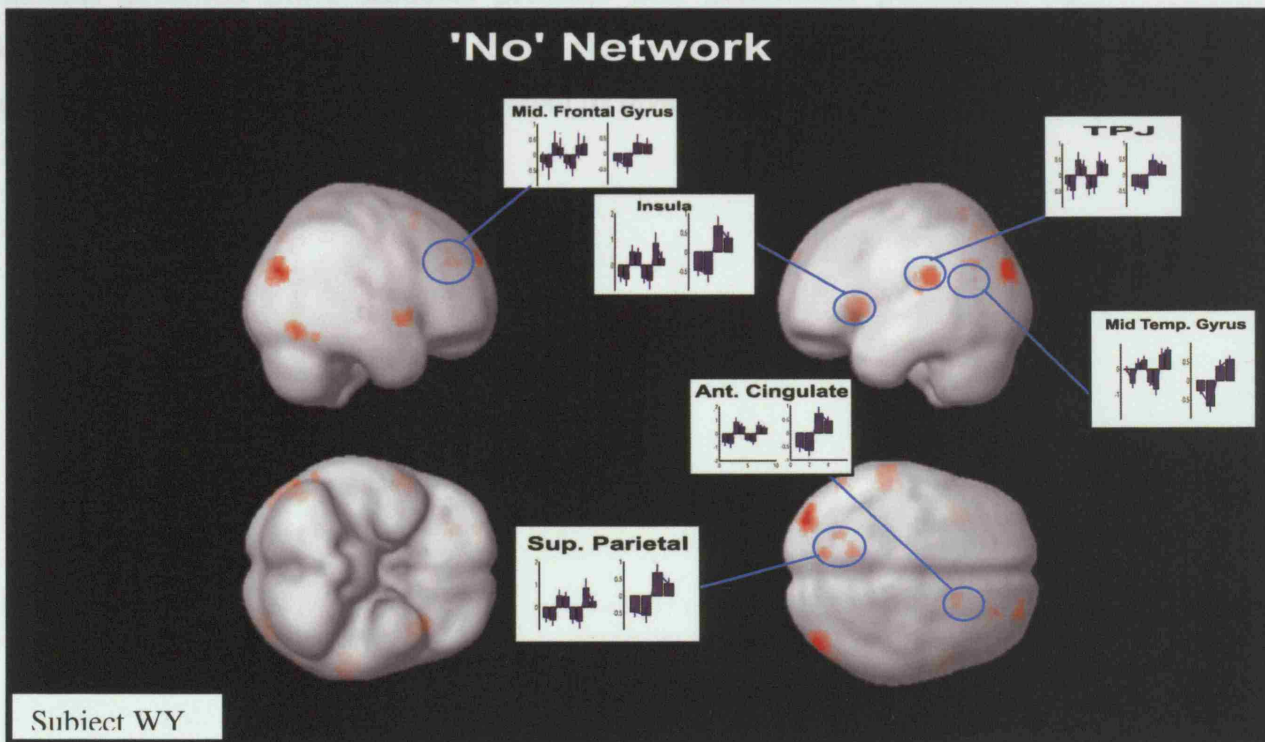


Figure 43. Parameter estimates from 'yes' network regions. Commonality of region judged according to anatomical location, through rendering on to individual structural scans and through use of anatomy toolbox. Because of anatomical variability, only a reduced set of core areas were common to all subjects, whereas in the individual cases there were many additional areas activated with a similar profile.



Despite the different data format, regardless of report mode, regardless of

display options, since stimulus generation is independent of



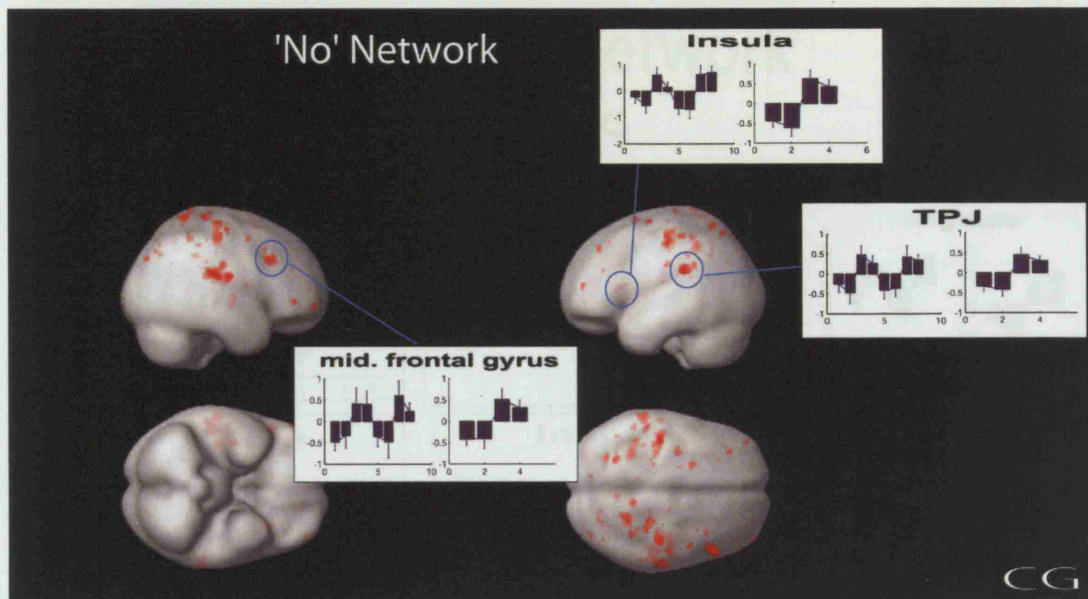


Figure 44. Depicts all areas that conform to the signal detection profile for a ‘no’ area. All regions, as can be seen from the parameter estimates, are more active when the subject decides no (misses and correct rejections) than when they decide ‘yes’ (hits and false alarms), regardless of report mode, regardless of stimulus presence. Since stimulus presence best determines perception (as defined by what can be reported at the shortest delay) the invariance of the ‘no’ network to perception can be inferred from its invariance to stimulus presence.



Figure 45. Parameter estimates from ‘no’ network regions. Commonality of region judged according to anatomical location, through rendering on to individual structural scan and through use of template surfaces. Because of anatomical variability, only a reduced set of core areas were common to all subjects, whereas in the individual cases there were many additional areas activated with a similar profile.

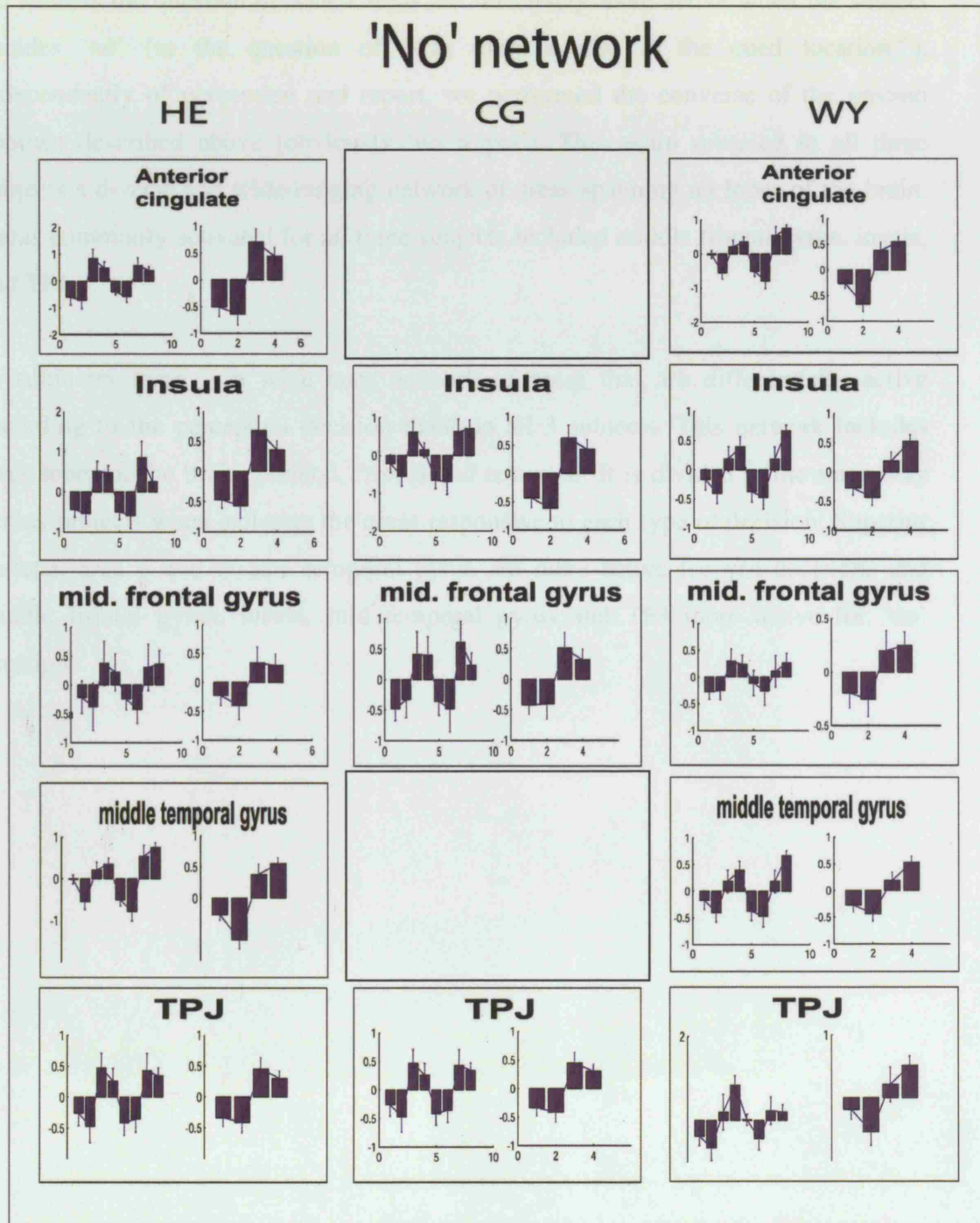
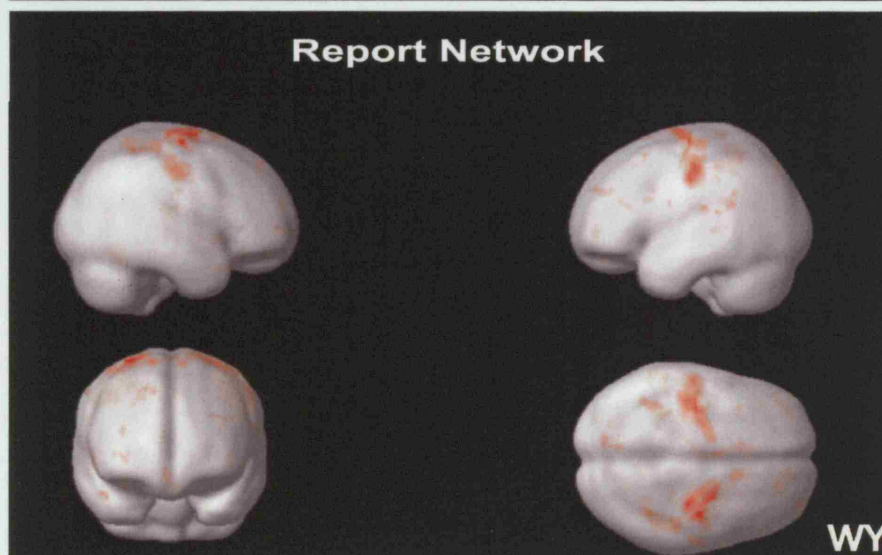
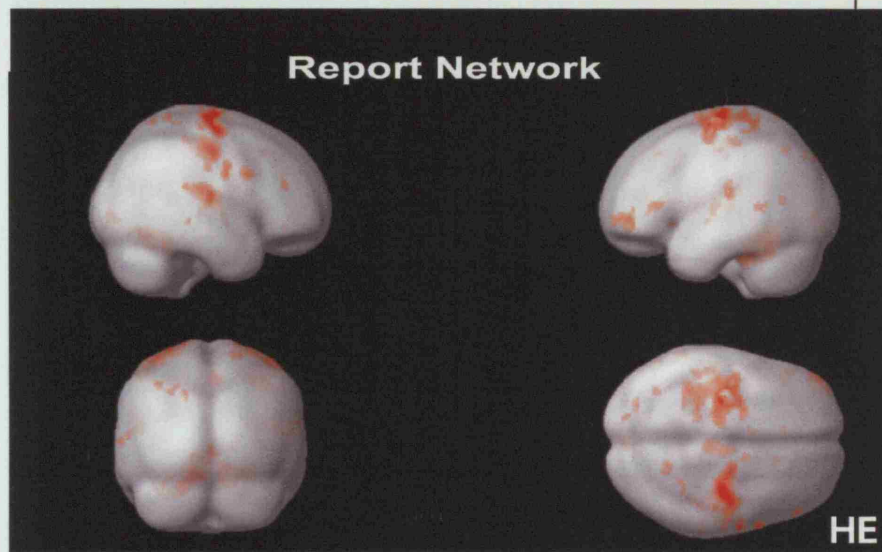


Figure 45. Parameter estimates from 'no' network regions. Commonality of region judged according to anatomical location, through rendering on to individual structural scans and through use of anatomy toolbox. Because of anatomical variability, only a reduced set of core areas were common to all subjects, whereas in the individual cases there were many additional areas activated with a similar profile.

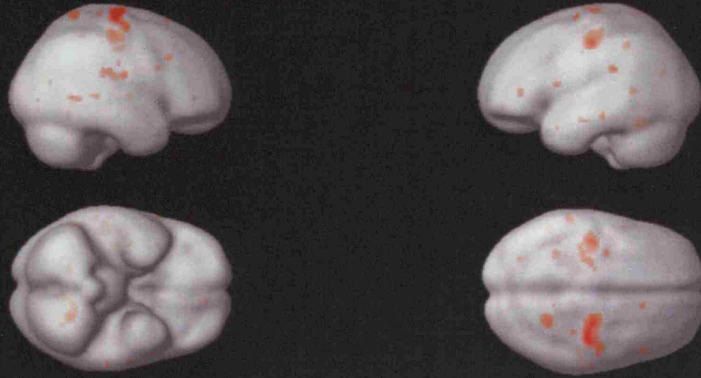
To address the question of which areas are selectively more active when the subject decides 'no' (to the question of 'was there a disc at the cued location?'), independently of perception and report, we performed the converse of the yes>no contrast described above (obviously 'no > yes'). This again revealed in all three subjects a diverse and wide-ranging network of areas spanning all lobes of the brain. Areas commonly activated for all three subjects included middle frontal gyrus, insula, and TPJ.

In summary there is a wide core network of areas that are differentially active according to the perceptual decision made in all 3 subjects. This network includes three lobes of the brain, parietal, frontal and temporal. It is divided in the same way across subjects when isolating the areas responsive to each type of decision. Superior parietal, area 3 and middle temporal gyrus are more active for yes decisions and middle frontal gyrus, insula, mid temporal gyrus and TPJ more active for 'no' decisions.

6.4.1.3 Report effects



Report network



CG

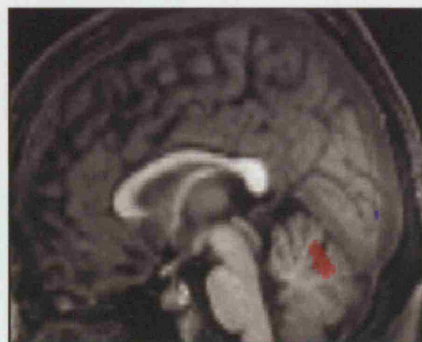
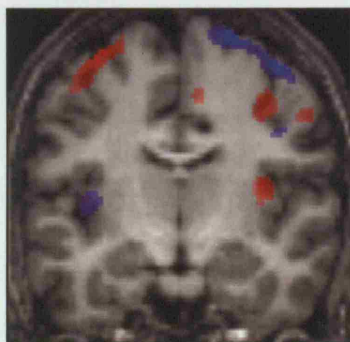
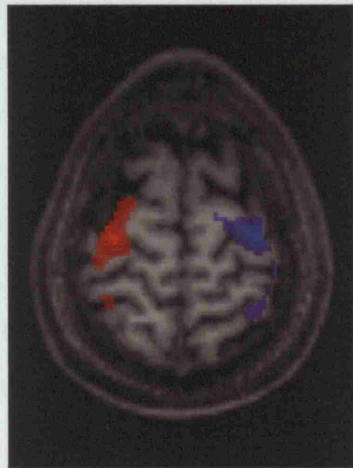
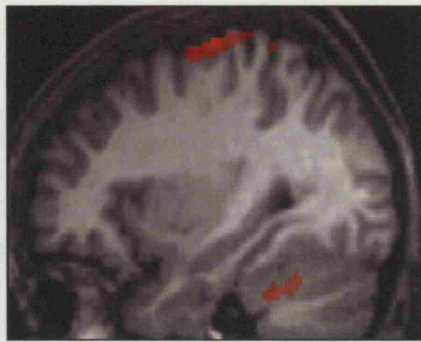
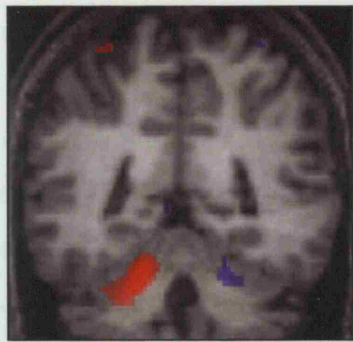


Figure 46 shows brain activity corresponding to a differential response according to the hand used to report the stimulus.

To address the question of which areas are selectively more active when the subject makes a report independently of their decision or perception, we performed the t-contrasts 'left response > right hand response' and its converse. The left response is all trials in mode 1 where subjects responded yes ('hits and false alarms') and all trials in mode 2 where subjects responded no (misses and correct rejections). For illustration we display the f-contrast for left > right and for right > left. This indicates all areas that are differentially active according to the motor act of reporting. This revealed in all three subjects consistent activations of area 4 (primary motor cortex), area 6 (premotor cortex), TPJ (temporo-parietal junction) and cerebellum. This is classical motor related activity with the exception of TPJ which is usually associated with attention. In addition there was inconsistent prefrontal and parietal activity in some subjects but not all.

6.4.2 Position specific effects

6.4.2.1 Stimulus specific effects

Position specific localiser:

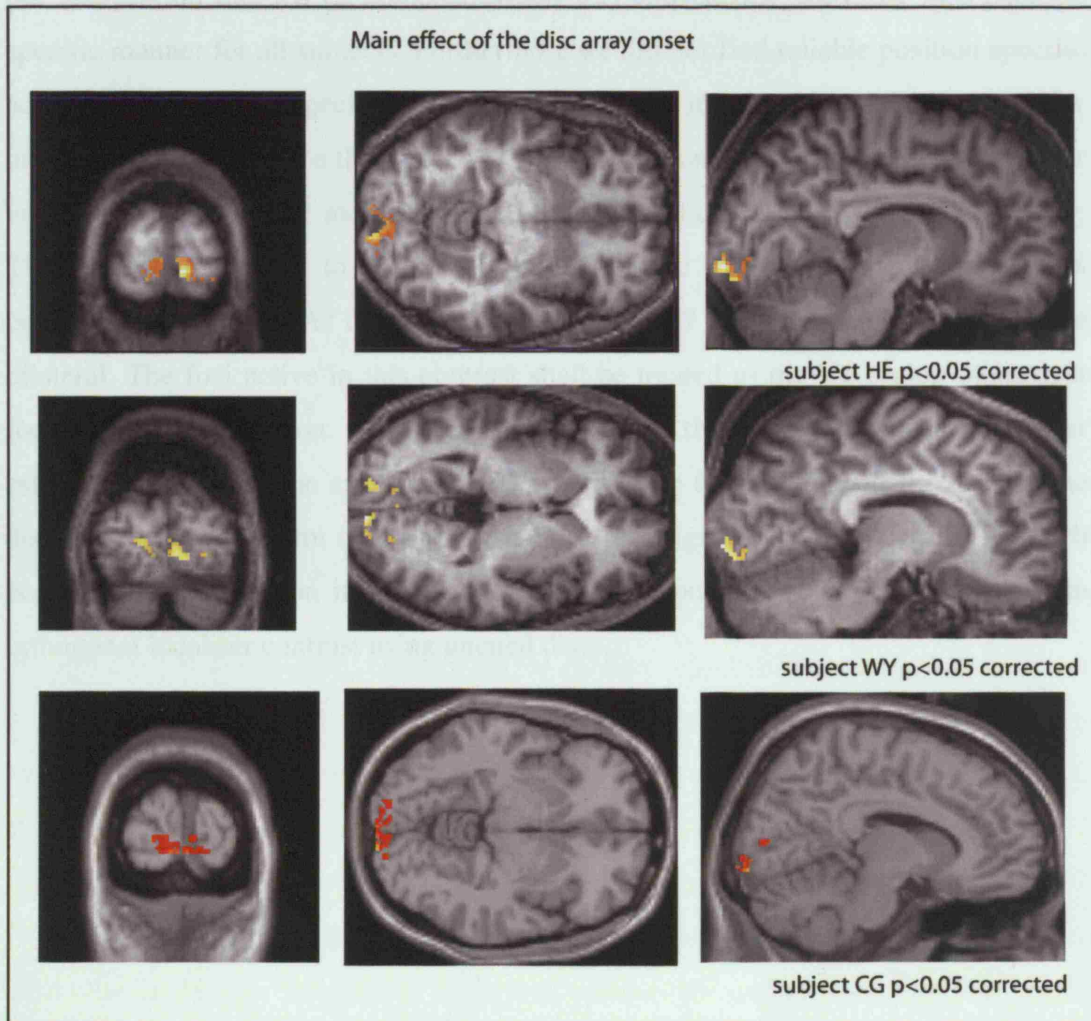


Figure 47. Shows main effect of disc onset for all unattended locations (masked areas outside V1 for display) rendered onto subjects own structural.

The main effect of disc array onset is presented above (masked areas outside dilated V1 for display) as voxels that show a significant increase in activity to the onset of the disc array. This was achieved with an f-contrast for all unattended location regressors. The figure should not be interpreted as the only areas of the brain activated by

stimulus onset of the disc array because the activations have been masked by a dilated template of V1 (BA17) from the SPM anatomy toolbox. The masking was performed only for inspection and illustration of V1 activity. Because this template cannot rule out V2 activity, the activity within should be considered as either V1 or V2.

Prestriate visual areas were activated inconsistently across subjects, although without retinotopy we cannot be sure which area is which. For this purpose we have focused our analysis on the V1/V2 complex which was consistently activated in a position specific manner for all subjects. Furthermore we did not find reliable position specific activity at the position specific or octant specific level in sites beyond the V1 mask for unattended locations. We therefore have no grounds on which to assume a particular voxel outside of V1/V2 as being selective for a particular subset of the disc array. Thus it is not possible to perform position specific analysis in the same way as performed for V1/V2. As can be seen from figure 47 the activations within V1 are bilateral. The foci active in this contrast shall be treated as the V1 region of interest for stimulus presentation. It is necessary to extract the location specific parameter estimates to evaluate the specificity of V1 activity to the location of the discs in the disc array. Since the aim of the analysis is to plot signal detection profiles for each position selective region it is necessary to localise position selective regions by an orthogonal localiser contrast using uncued discs.

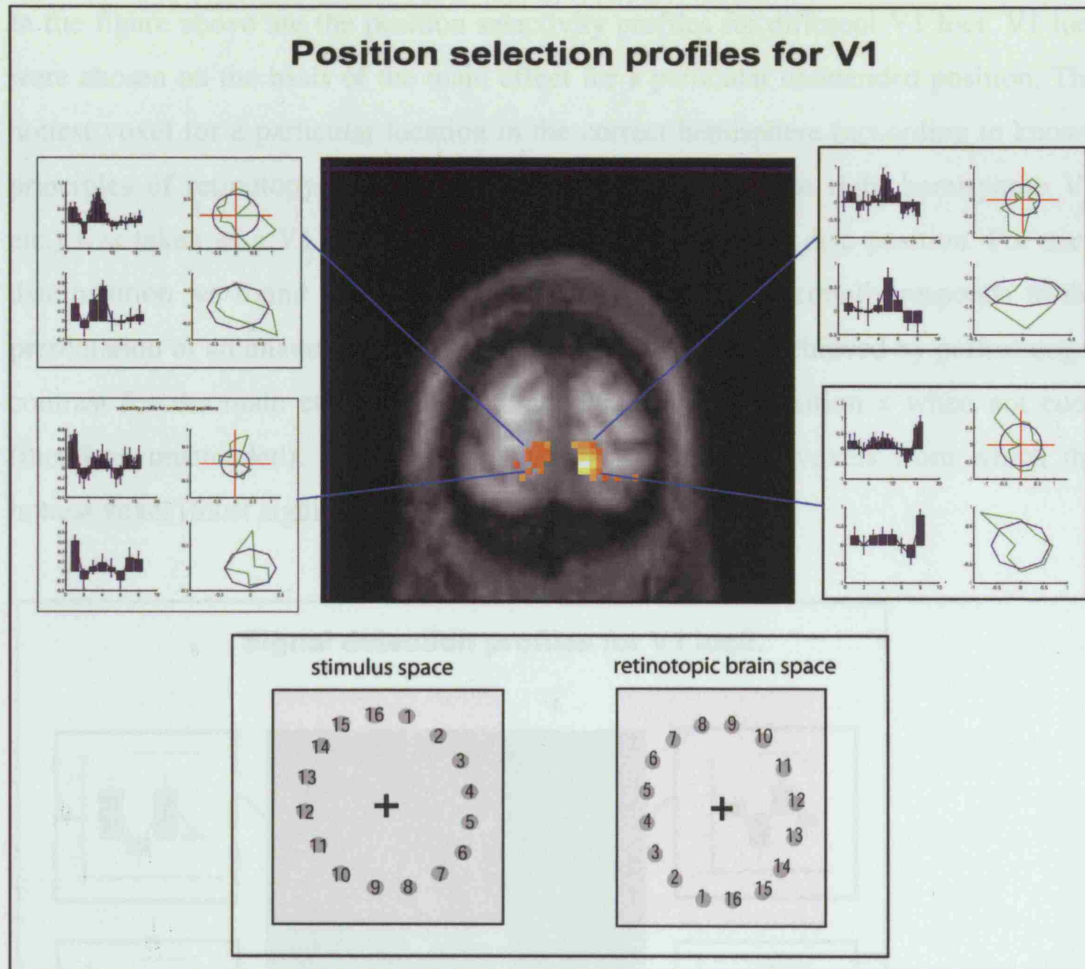


Figure 48. Shows position selectivity for different loci in V1. Criteria for V1 loci are the hottest voxel for the main effect for a single unattended stimulus position, which is in the correct quadrant of retinotopic space. The figure depicts 4 voxels for illustration, one taken from each quadrant of retinotopic space (upper bank calcarine left hemisphere, lower bank left etc.). For each voxel there are 4 plots. The upper two plots are based on plotting parameter estimates for all 16 stimulus positions. The lower two plot parameter estimates averaging across octants (8 stimulus positions). The left-hand column shows standard parameter estimate plots whereas the right hand column plots the same data in terms of a polar plot. Polar plots are used to give an intuitive grasp of the position selectivity in stimulus space (see inset labelled stimulus space). Green plots the magnitude of the parameter estimate and the polar angle is in line with stimulus space. Blue indicates an isomagnitude contour that is what would be expected if the parameter estimates all conformed to their average value.

In the figure above are the position selectivity profiles for different V1 loci. V1 loci were chosen on the basis of the main effect for a particular unattended position. The hottest voxel for a particular location in the correct hemisphere (according to known principles of retinotopy left hemifield stimuli should activate right hemisphere V1 etc.) was taken as a V1 response selective for that particular disc position. For each disc position we found the 'hottest' voxel in V1, which maximally responds to the presentation of an unattended disc in that position. This was achieved by performing a contrast for the main effect of the presence of a disc at position x when not cued (therefore unattended). Such contrasts reveal a number of voxels from which the hottest voxel (most significant) is selected.

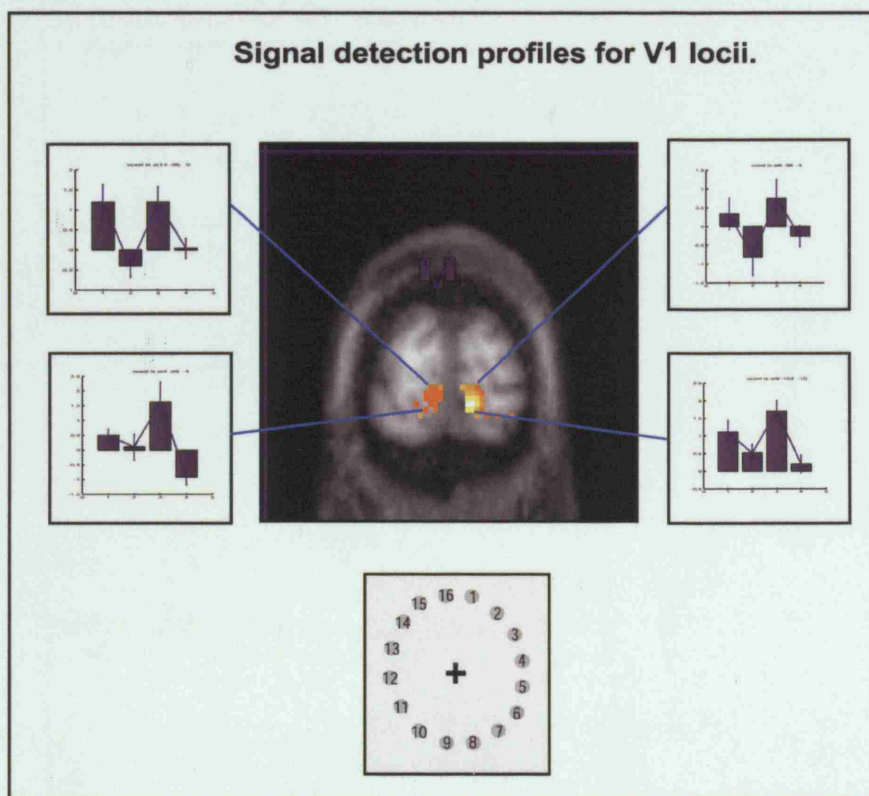


Figure 49. Signal detection profiles plotted for each of the V1 loci in 48. Parameter estimates plotted for each signal detection category (averaging across report modes) only for trials in which the position which the V1 locus is assigned to (by virtue of it being the most significantly active voxel for that position as above) is cued. Example data from subject HE.

For each of the V1 loci we plotted its signal detection profile to its preferred position. This is best explained with an example. Having identified that 'V1 locus # n ' responds more to a disc presented at position # n than any other disc position. We then looked at its different responses to each signal detection category when position n was cued. This is plotted as a position selective signal detection profile, four of which are illustrated in the above figure.

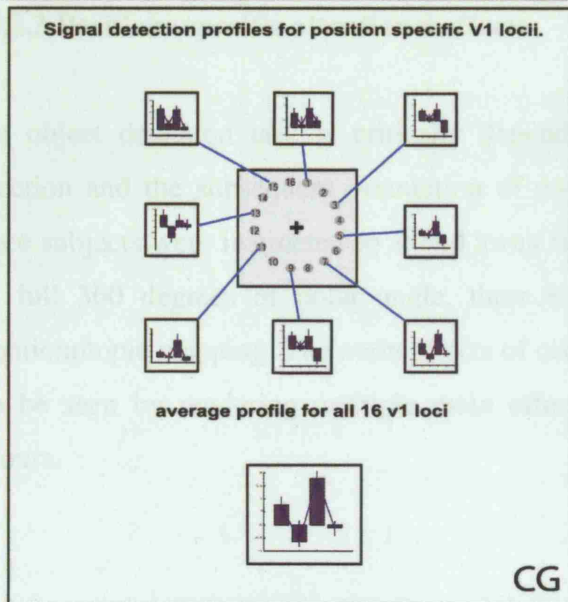
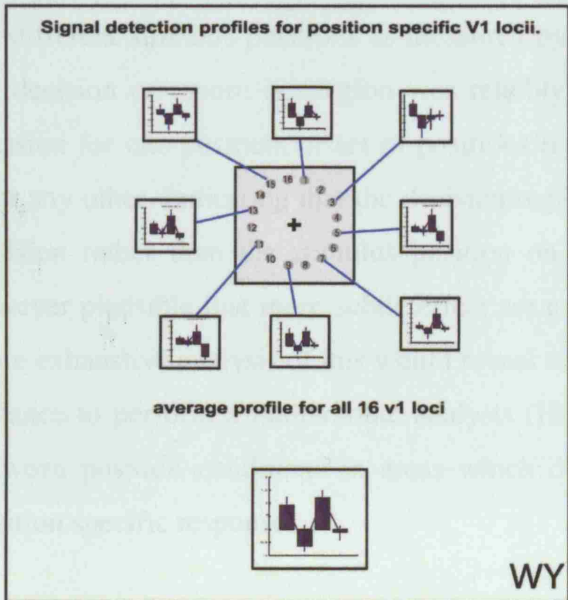
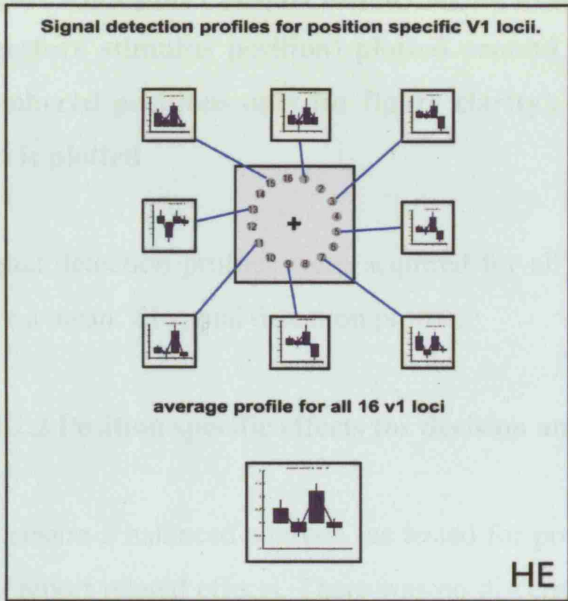


Figure 51. Upper: graphs display signal detection profiles for each V1 loci (and therefore stimulus position) plotted around a depiction of stimulus space (odd numbered positions only for figure clarity). Lower: The average of all of these loci is plotted

Signal detection profiles were acquired for all 16 V1 loci and averaged together to give a mean V1 signal detection profile.

6.4.2.2 Position specific effects for decision and report

To ensure a balanced analysis we tested for position specific effects for the decision and report related effects. There was no discernable difference between yes decisions for different stimulus positions as measured by interposition subtractions controlling for decision or report. No region was reliably more active for yes decisions or no decision for one position, or set of positions (collapsing across octants or quadrants) over any other, indicating that the decision regions are abstractly related to the actual decision rather than the stimulus position on which the decision was based. It is however plausible that more subtle effects are evident within regions of interest and a more exhaustive analysis of this would reveal some effects. It would be interesting for instance to perform a multivariate analysis (Haynes and Rees 2005) to discriminate between position conditions in areas which do not show univariate differences in position specific responses.

6.4.2.3 Positions specific attentional effects

The object detection task is critically dependent on the recognition of the cue's direction and the subsequent orientation of covert attention to the correct position. Since subjects were instructed to attend away from fixation in 16 directions covering the full 360 degrees of polar angle, there is a rich data set for looking at the attentionotopic mapping. The main effects of orienting attention in different directions can be seen by rendering multiple main effects on to the same brain in different colours.

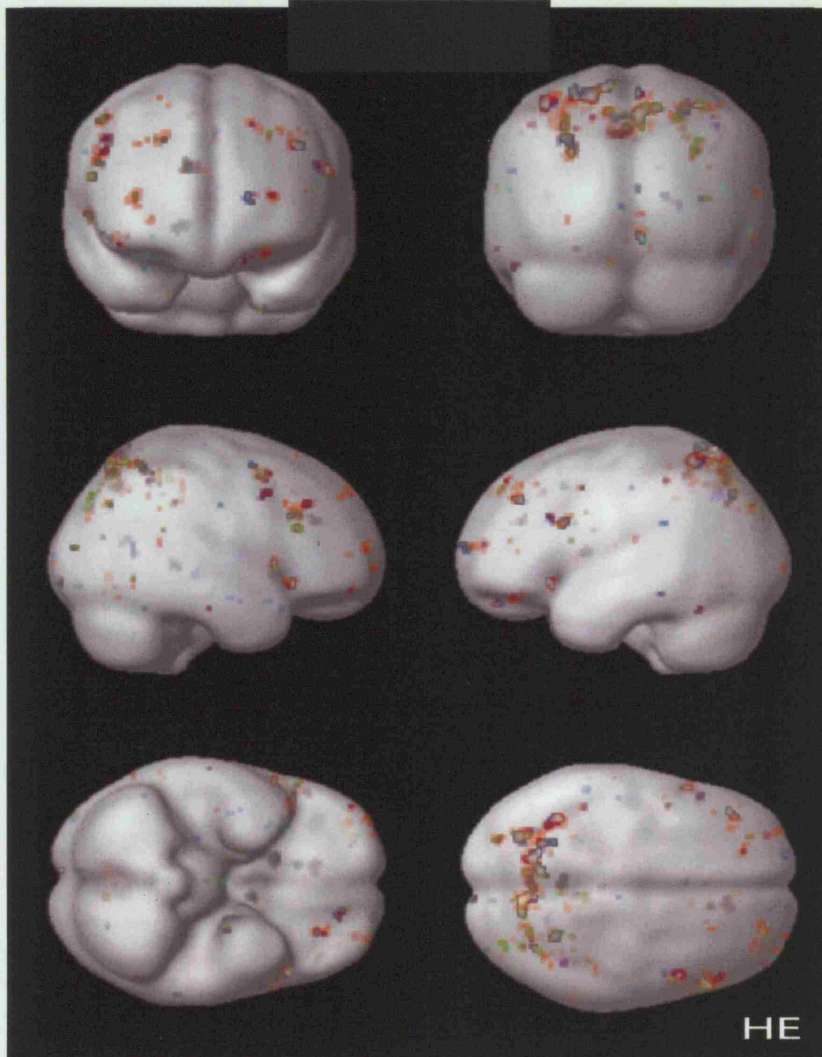
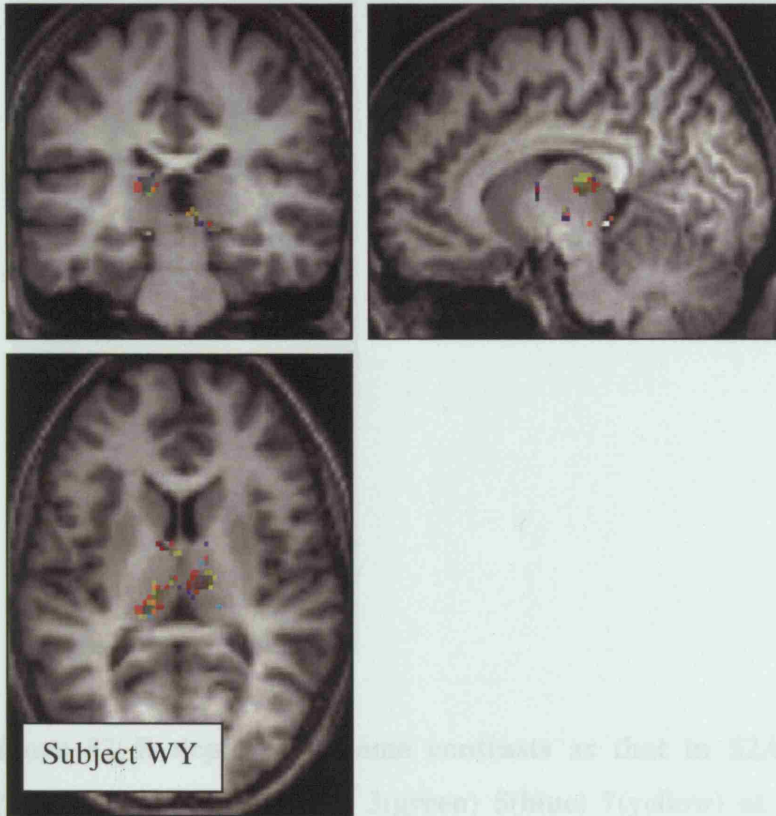


Figure 52A. Different main effects for different directions of attentional orientation (grouped into octants) are plotted in different colours. There is no clear lateralisation. The main effect of attending to octants 1 (red) 3 (green) 5 (blue) and 7 (yellow) are rendered above at a threshold of $p < 0.05$ corrected. Canonical frontoparietal activity can be observed similar to many other previous studies (for review see Raz and Buhle 2006)

Cortical mapping of attention has been performed previously using phase-mapping techniques adapted from the retinotopic studies (for example see Silver et al 2005). Since this has been performed before, it was not our intention to investigate it further.

An unexpected finding (unexpected due to resolution issues) however was the attentionotopic response of the thalamus (see below figure 52B).



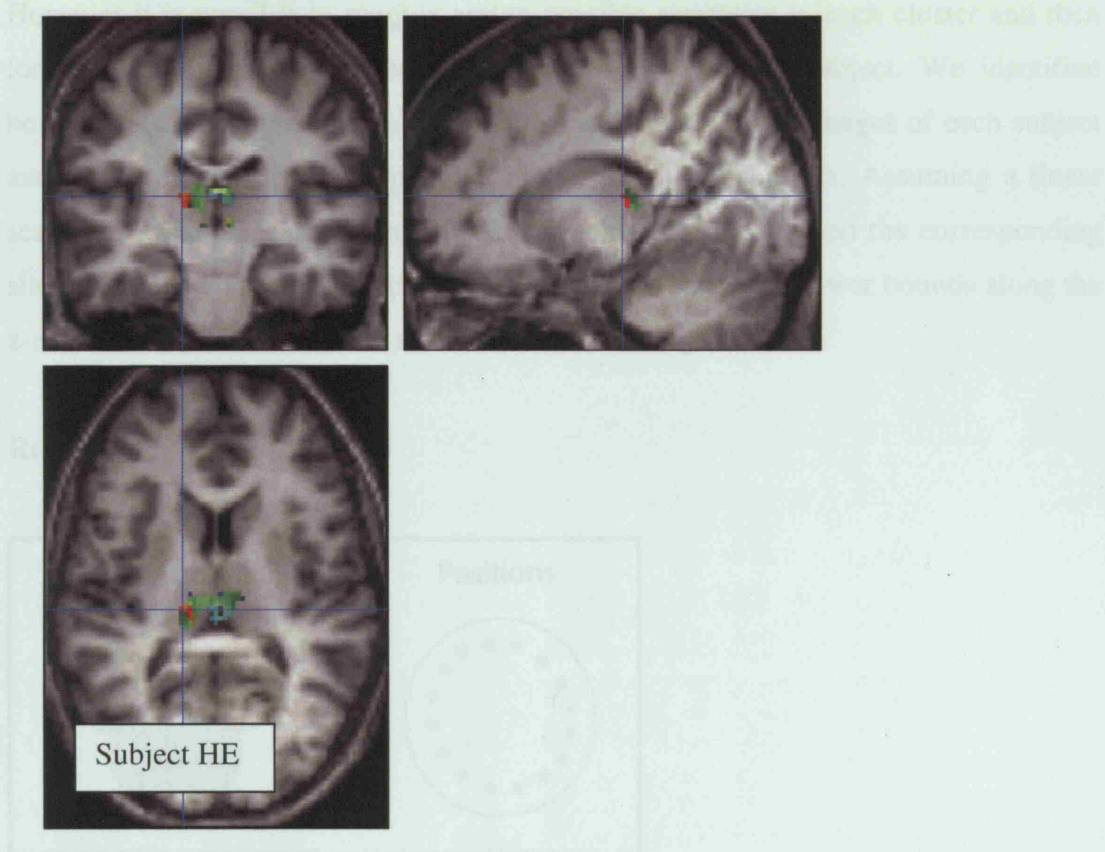


Figure 52 B depicts the same contrasts as that in 52A, the main effect of attending to octants 1(red) 3(green) 5(blue) 7(yellow) at $p < 0.001$ uncorrected. For clarity an anatomical mask was used to isolate the thalamus.

The finding that there are attentionotopic maps in the posterior thalamus is important since this has never been shown before with either electrophysiology or neuroimaging. The location of the attentionotopic clusters appears to be the pulvinar nucleus. This unexpected finding motivates a brief digression into a more detailed investigation of the underlying anatomy.

6.5 Anatomical analysis

In the absence of any probabilistic atlas of thalamic subnuclei we resort to a visual comparison of our results against a stereotactic atlas of the human thalamus (Morel et al 1997). Given the anatomical variability between single subjects, and indeed the 2 single subjects used in the atlas, any anatomical localisation is going to be crude.

However it is possible to roughly assign possible structures to each cluster and then look for commonalities in the structures assigned to each subject. We identified homologous landmarks visible in both atlas and the structural images of each subject and using this defined corresponding boundaries in all subjects. Assuming a linear scaling in brain space between subjects and the atlas we defined the corresponding slices between atlas and subject. E.g. 50% between upper and lower bounds along the z-axis for the atlas is the same slice as that for the subject.

Results

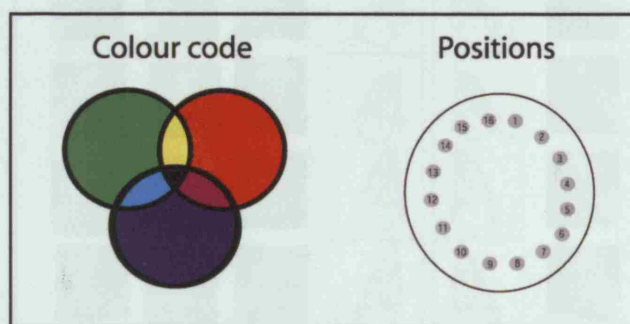


Figure 52C. Colour code depicts the colour coding rules for the following anatomical and functional images. Red, Green and Blue displays voxels significant for the main effect of only one attentional direction (colour – position correspondence is indicated at the top of each anatomical figure) at 0.001 corrected for multiple comparisons within a thalamic search volume (anatomically defined using SPM anatomical toolbox). Yellow, Cyan and Magenta, indicate voxels in which there are two significant effects as indicated by the overlap of the two colours in the diagram, e.g. Red + Green= Yellow. Black indicates voxels where all 3 effects are significant.

To visualise attentionotopic maps within the thalamus we rendered the main effects of different attentional shift directions onto the same structural image for each subject. For an intuitive colour mixing rule (see figure 1) we rendered 3 attentional conditions at a time, as far apart around the circle as possible (e.g. positions 2, 8 and 13), individually for each subject. This was repeated twice for positions 3, 9 and 14 and 4,10 and 15. The results for positions 2, 8, 13 are plotted axially in figures 2 sagittally in figure 3.

Figure 224. Super axial slices of all 3 subjects showing thalamus. Ref. voxel

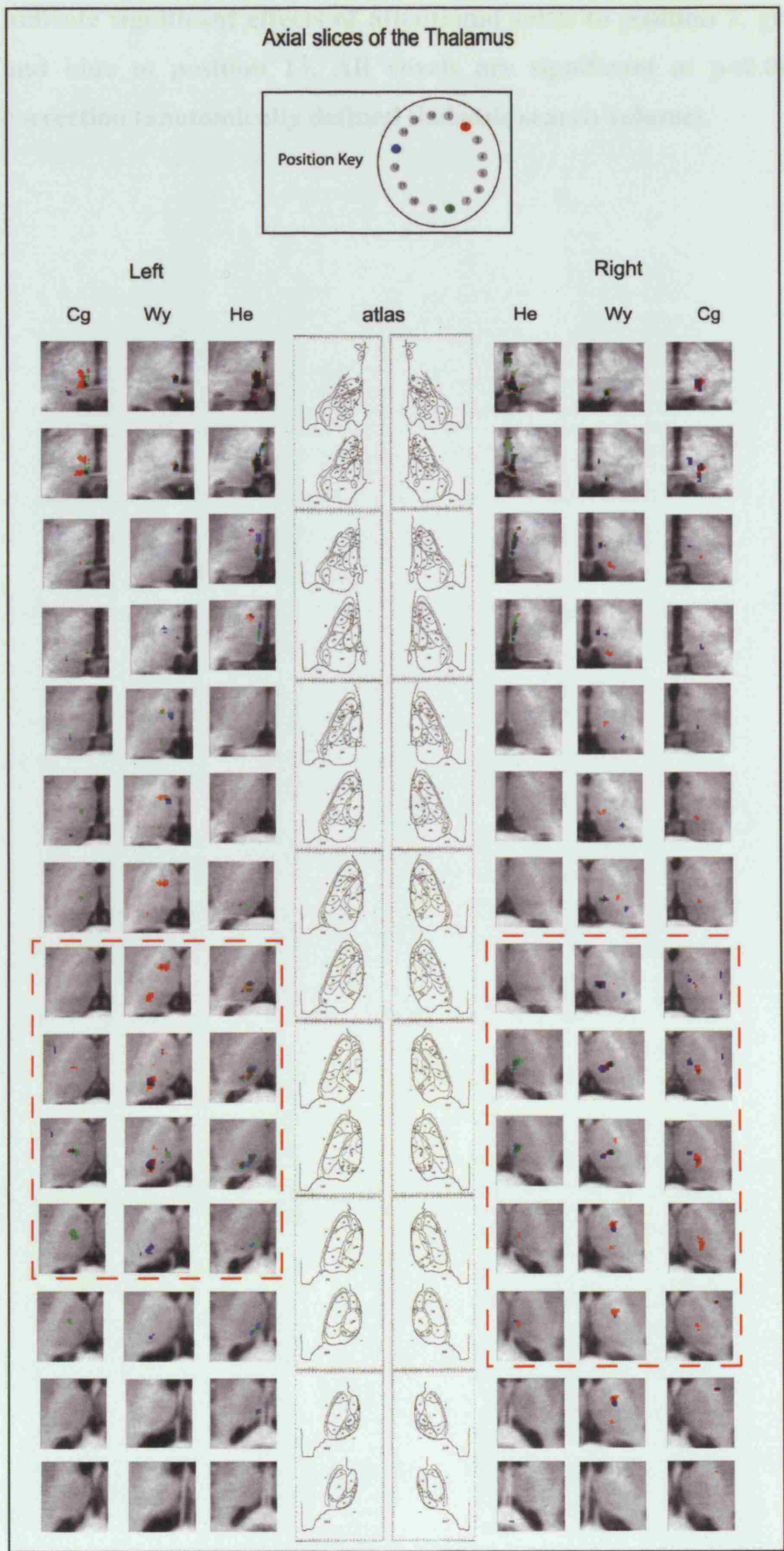


Figure 52D. Shows axial slices of all 3 subjects complete thalamus. Red voxels indicate significant effects of attentional shifts to position 2, green to position 8 and blue to position 13. All voxels are significant at $p < 0.001$ small volume correction (anatomically defined thalamic search volume).

Saggital slices of the Thalamus

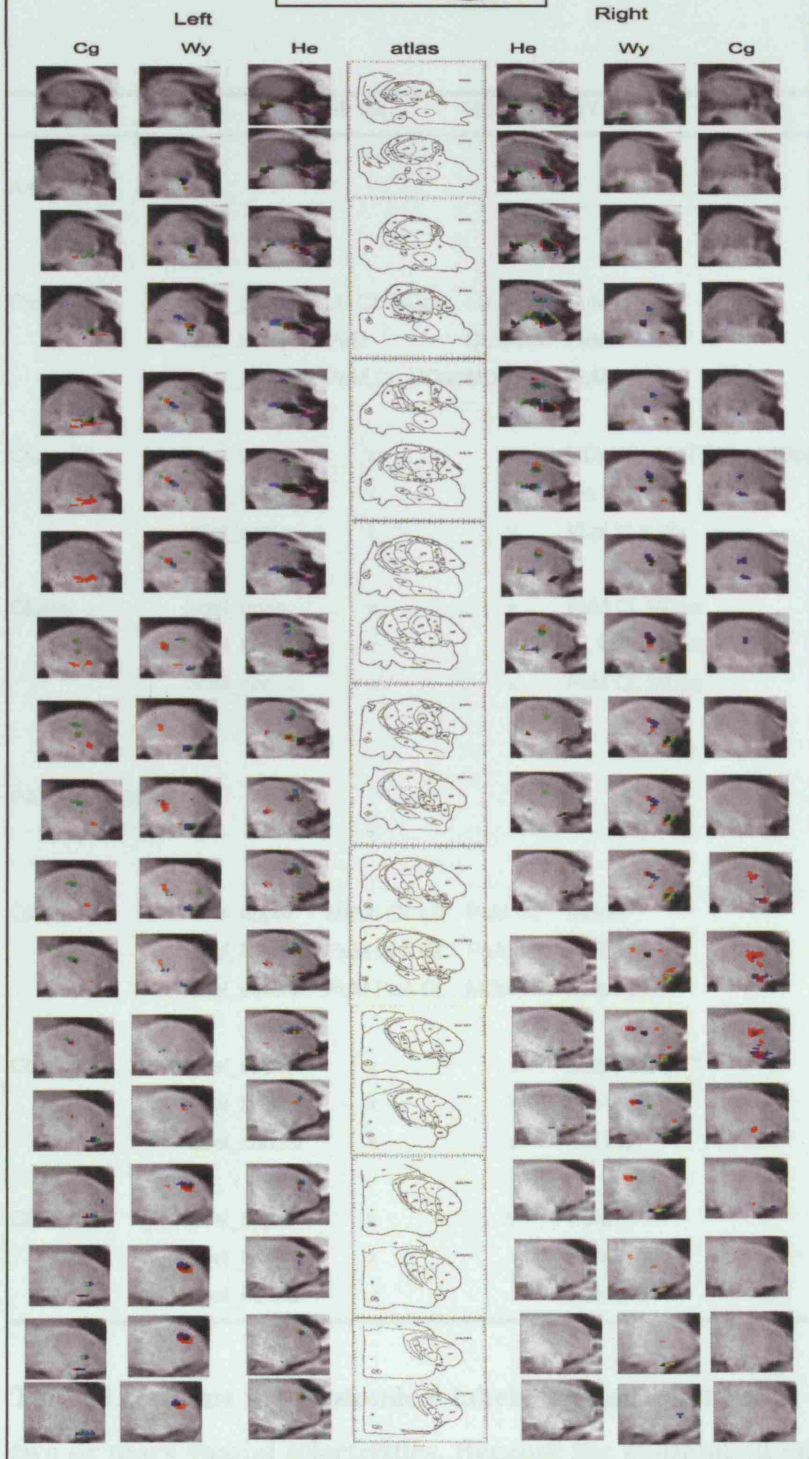
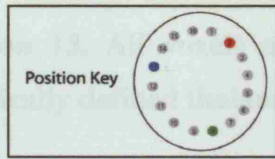


Figure 52E. Shows axial slices of all 3 subjects complete thalamus. Red voxels indicate significant effects of attentional shifts to position 2, green to position 8 and blue to position 13. All voxels are significant at $p < 0.001$ small volume correction (anatomically defined thalamic search volume).

		HE left	HE right	WY left		WY right	CG left	CG right
AXIAL								
Cluster 1	label_upper	LD CL MDpc	LD CL	PuM		MDpc MDpl	MDpc MDpl VLp	PuM LP CL
	label_lower	PuM	CL MDpc	PuM		MDpc CL	VLp	PuM LP CL
	label_central	PuM,CL,MDpc	MDpc CL	PuM LP		MDpc CL PuM MDpl	VLp	PuM LP CL
Cluster 2	label_upper	x	x	MDpl VLpl MTT	VLp MDpc	PuM LP CL	x	x
	label_lower	x	x	Vla VLp VA		PuM LP CL	x	x
	label_central	x	x	VLpl VLp Vla		PuM LP CL	x	x
Cluster 3	label_upper	x	x	PuM CL MDpc		x	x	x
	label_lower	x	x	CL MDpc MDmc		x	x	x
	label_central	x	x	PuM CL MDpc		x	x	x
PARA-SAGGITAL								
Cluster 1	label_upper	MDpc CL LD	PuM CL	MDpc		MDpc Pf	MDpc CL MTT	PuM PuA VPLp
	label_lower	PuM PuA CL	PuA VPLp	CL MDpl		MDpc CM	VLpv Vla	PuA CL VPLp
	label_central	PuM PuA CL	MDpl CM	MDpc CL		MDpc CM	VLpv Vlap	PuA PuM VPLp
Cluster 2	label_upper	x	x	MDpc CL MTT		CL MDpl	x	x
	label_lower	x	x	Vla VA		LP PuM	x	x
	label_central	x	x	VA Vla		PuM CL	x	x
Cluster 3	label_upper	x	x	PuM		x	x	x
	label_lower	x	x	PuM		x	x	x
	label_central	x	x	PuM		x	x	x

Table 4 contains the anatomical labels for each significant cluster identified with two or more spatial selectivities. Because the anatomical labelling by comparison with single subject atlases is particularly crude, we are careful not to be overly specific in the anatomical labelling. Between 1-3 clusters are found for each

hemisphere. For each cluster we attach 3 labels, one for the each boundary limit (defined in x or z according to whether slice is sagittal or axial) and one for the centre of gravity.

Abbreviations			
AC	anterior commissure	Po	posterior nucleus
AD	anterodorsal nucleus	PuA	anterior pulvinar
AM	anteromedial nucleus	PuI	inferior pulvinar
AV	anteroventral nucleus	PuL	lateral pulvinar
Cd	caudate nucleus	PuM	medial pulvinar
CeM	central medial nucleus	PUT	putamen
CG	central gray	Pv	paraventricular nuclei
CL	central lateral nucleus	R	reticular thalamic nucleus
CM	centre median nucleus	RN	red nucleus
Ex	fornix	Sg	suprageniculate nucleus
GPI.e	globus pallidus, internal and external segments	SN	substantia nigra
Hb	habenular nucleus	sPF	subparafascicular nucleus
IC	internal capsule	STh	subthalamic nucleus
LD	lateral dorsal nucleus	VA	ventral anterior nucleus
LCN	lateral geniculate nucleus	VAmc	ventral anterior nucleus, magnocellular division
Li	limitans nucleus	VLa	ventral lateral anterior nucleus
LP	lateral posterior nucleus	VLP	ventral lateral posterior nucleus
MB	mammillary body	VLpl	ventral lateral posterior nucleus, paralamellar division
MD	mediodorsal nucleus	VLpa	ventral lateral posterior nucleus, dorsal division
MDmc	mediodorsal nucleus, magnocellular division	VLp	ventral lateral posterior nucleus, ventral division
MDpc	mediodorsal nucleus, parvocellular division	VM	ventral medial nucleus
MDpl	mediodorsal nucleus, paralamellar division	VPI	ventral posterior inferior nucleus
MGN	medial geniculate nucleus	VPI	ventral posterior lateral nucleus
MGM	medial geniculate nucleus, medial division	VPI.a	ventral posterior lateral nucleus, anterior division
MGe	medial geniculate nucleus, ventral division	VPLp	ventral posterior lateral nucleus, posterior division
MTT	mammillothalamic tract	VPM	ventral posterior medial nucleus
MV	medioventral nucleus	VPMpa	ventral posterior medial nucleus, parvocellular division
OT	optic tract	ZI	zona incerta
PC	posterior commissure		
PF	parafascicular nucleus		

Figure 52G. Table of abbreviations for anatomical locations within the thalamus

In contrast to the strong attentionotopic effects in the thalamus, there was no such effect for the location of stimulus (either cued or uncued). Even at liberal thresholds of $p < 0.05$ uncorrected there was no meaningful activity for position specific stimulus onsets.

7.1 Discussion Experiment 1

Experiment 1 shows that areas of the visual brain that are selective for the processing of faces and houses respond with the same magnitude and follow the same timecourse whether a stimulus is visible or occluded (figure 31). This result is partly surprising and partly not. To a certain degree, one would expect that the same set of cortical areas would be involved in the processing these stimuli, regardless of whether they are directly perceived or occluded. The surprise lies rather in the fact that the areas were activated with the same magnitude and had the same timecourse, whether perceived or not. The results imply that the activity that is specific to a visual stimulus is maintained in these areas even after the stimulus has disappeared from view. That there is a difference in cortical response to directly perceived and occluded visual stimuli is shown by our other result (Figure 32), namely that occluded stimuli selectively recruit different parts of prefrontal cortex according to the type of stimulus (face or house) and that premotor cortex is commonly recruited by the occlusion of either stimulus type. Finally we show that the premotor cortex is the only region to significantly increase its covariance with the object selective regions (FFA and LOC) under conditions of occlusion (Figure 33). We therefore speculate that the premotor cortex may exert a controlling influence over the retention of a previously seen visual stimulus that is subsequently occluded. Before evaluating our results in terms of perception and awareness it is necessary to consider the experiment in light of other cognitive processes that might influence our interpretation.

Working memory and Imagery

Although this experiment did not explicitly require the memorisation or imagery of objects, it is at least plausible that viewing a stimulus that becomes occluded automatically recruits mechanisms in common with working memory or visual imagery. Object working memory tasks, such as the delayed match to sample task, have been shown to recruit prefrontal regions (different to those in this study) and object-selective regions of inferotemporal cortex in humans and in non-human primates (Druzgal and D'Esposito 2003, Druzgal and D'Esposito 2001). Similarly

visual imagery tasks have also been shown to activate fronto-parietal and object-selective regions (O'Craven and Kanwisher 2000).

Despite the similarities between the activities reported here and in previous working memory and imagery studies, there are important differences. During the delay period of working memory tasks, in which the face is memorised over a period of time during which it is not seen, FFA activity drops significantly (Druzgal and D'Esposito 2003, Druzgal and D'Esposito 2001). The working memory of a face therefore results in activity that is significantly lower in magnitude than the response to the perception of a face. In the same way, visual imagery of faces has been shown to activate the FFA above baseline but to a level significantly less than during perceptual conditions (O'Craven and Kanwisher 2000). In this study we show that the timecourse of the FFA and LOC during the occluded (OC) conditions matches that of the visible (VS) conditions and therefore is significantly different to that reported previously under conditions of working memory and imagery.

The difference in response between occluded stimuli and those that are either imaged or held in working memory mirrors the phenomenological differences between the tasks; looking at a face that becomes occluded is very different from holding a face in working memory or imagining it in your mind's eye. The occluded face is known to be actually present, whereas the imaged face, or the working memory of a face, is known to be not actually present. One would expect such an important distinction to be reflected robustly in the response dynamic of a visual area (Pollen 1999) which it indeed seems to be.

One possibility is that activity common to VS and OC conditions is due to the appearance or recognition of the very top portion (2.5% of surface area) of the object that remained visible during OC conditions. This is controlled for with the segment (SG) conditions. The timecourse data for condition SG (Figure 31) show that the FFA or the LOC does not respond to the appearance of the segment. This effectively rules out any explanation of the occluded stimulus activity in terms of spatial context or low-level visual properties.

It is of interest that we did not observe significant Parahippocampal Place Area (PPA) activity for the visual presentation of houses that had been previously observed (Tong et al 1998). Note that the PPA responds not to houses *per se* but to the spatially extended geometry of the local environment (Epstein and Kanwisher 1999). Since all our stimuli were depicted in the context of 3-dimensional space there should have been no systematic difference in PPA activity between house conditions and face conditions. The only difference is in the visual form of the house as an object, and thus the difference should be reflected in regions selective for object form such as LOC, which we found.

The specificity of premotor cortex for occluded stimuli

The most significant cluster of voxels for occluded stimuli was located in ventral premotor cortex. Given that premotor cortex has classically been implicated in the planning and control of movements (Kurata and Tanji 1986; Rizzolatti et al 1981, 1988) its response to occluded stimuli may seem surprising and counterintuitive. An explanation in motoric terms must therefore be considered before alternate hypotheses are entertained.

Premotor activity could plausibly indicate systematic differences in motor planning or motor actions, but we consider this to be unlikely due to the design of the paradigm; subjects were only instructed to press a button in the 15 % of all trials in which a 'hole' appeared at a randomly selected time. The onsets of the 'holes' were randomly distributed and unpredictable and there could therefore be no systematic difference in motor preparation that could account for the specificity of the premotor response. Furthermore, the trials in which holes appeared (and therefore where motor responses were required) were independently modelled as effects of no interest.

The premotor activity observed in this study is perhaps better explained in terms of visual processing. Ventral premotor neurons show selectivity for stimuli independently of any motor response (Murata et al 1997) and of actions directed towards objects (Umiltà et al 2001). Additionally, the premotor cortex response to an observed object-directed action persists even when the object is occluded (Umiltà et al 2001). Such a response is contingent on the monkey knowing that an action is

directed toward an occluded object and cannot be induced if the monkey knows there is no occluded object. This should stand as positive evidence that the response of the premotor cortex can be contingent on the existence of object permanence. Finally, in monkeys, the premotor response to a visible object can persist in time even when the lights are extinguished and it is rendered invisible (Graziano et al 1997). This has been interpreted as evidence for the role of premotor cortex in object permanence (Graziano et al 1997), and is therefore directly relevant to the occlusion specific premotor response we observed. In both the monkey and the human, ventral premotor cortex responds when objects are known to be present but not seen.

Our result, in light of convergent evidence in the monkey and human, prompts us to support a functional role for premotor cortex in object permanence. We speculate that for the brain to sustain knowledge of the presence of objects when deprived of local sensory information, premotor cortex forms a functional coupling with object-selective regions. We therefore predict that disruption to the premotor cortex or its connection to object-selective regions would disrupt the ability of object-selective regions to maintain heightened activity during occlusion and thus would disrupt object permanence.

FFA does not correlate with perception

Although interesting in its own right, the principal objective of this experiment was not to study occlusion *per se*, but to study occlusion because of the perceptual dissociation it provides.

To investigate the relationship between neural activity and perceptual awareness it is necessary to go beyond correlations and examine what neural activity is necessary and/or sufficient for perceptual awareness. One approach that follows from this strategy is to evaluate the relationship between neural activity in the FFA and the corresponding state of perceptual experience of a face by considering each of the four possible combinations of face perception and FFA activity (Kanwisher 2001). For this purpose alone we will consider face perception as being either present or absent and FFA activity as being either high or low with high activity being that observed in the

FFA during direct face perception and low activity being any level significantly lower than this.

The evidence that (1.) *high FFA activity correlates with face perception* (Moutoussis and Zeki 2002) and (2.) *low FFA activity correlates with unconscious face processing* (Moutoussis and Zeki 2002) is the basis for the basic correlation between the level of FFA activity and face perception. There are two remaining possibilities that could elucidate the relationship further. If (3) *Low FFA activity correlates with face perception* then the necessity of high FFA for face perception could be ruled out. There is however, no evidence to date to support this. The results of this experiment are interesting because they satisfy the final possibility of (4) *High FFA activity in the absence of face perception*. High FFA activity alone, as our evidence suggests, cannot be sufficient for face perception. This is an important point because it forces a re-evaluation of the contribution of the FFA to visual awareness.

A theory that accounts for the FFA's contribution to visual awareness must extend beyond just the perception of faces. Whenever a face is visible, the perception of its attributes leads, under normal conditions, to the awareness that a face is present. When a face is occluded, only awareness of its presence remains and therefore we argue that the most parsimonious explanation is that the common FFA activity recorded in both cases correlates not with face perception but with awareness of the presence of a face. Our findings appear to confirm the theoretical speculations of Pollen whose model states that neural representations of an object 'can achieve an independent but non-phenomenal existence apart from the sensory data from which it was derived' (Pollen 1999). All of the above arguments are equally valid for LOC and its response to houses.

7.2 Discussion Experiment 2

7.2.1 Psychophysics

The psychophysical results show that there is a significant decay in performance with increasing cue delays for all positions for all subjects. In contrast to previous findings, we estimate that iconic memory decays slowly taking up to ~1000ms to reach threshold rather than previous estimates of 300-500ms (Lu et al 2005, Sperling et al 1960). This is likely to be due to the differences in experimental design. The task in experiment 2 was a simple luminance based object detection task whereas most other tasks have involved letter or number reporting, in which case there are either 26 (letters A-Z) or 10 (numbers 0-9) possible choices to report. With the object-detection task there are only two possible choices present or absent and thus this could plausibly account for a difference in the length of iconic memory traces. In support of this, change blindness studies using very simple oriented bars have recorded that post-offset cueing can protect against change blindness for up to 1200ms after offset (Landmann and Lamme 2003).

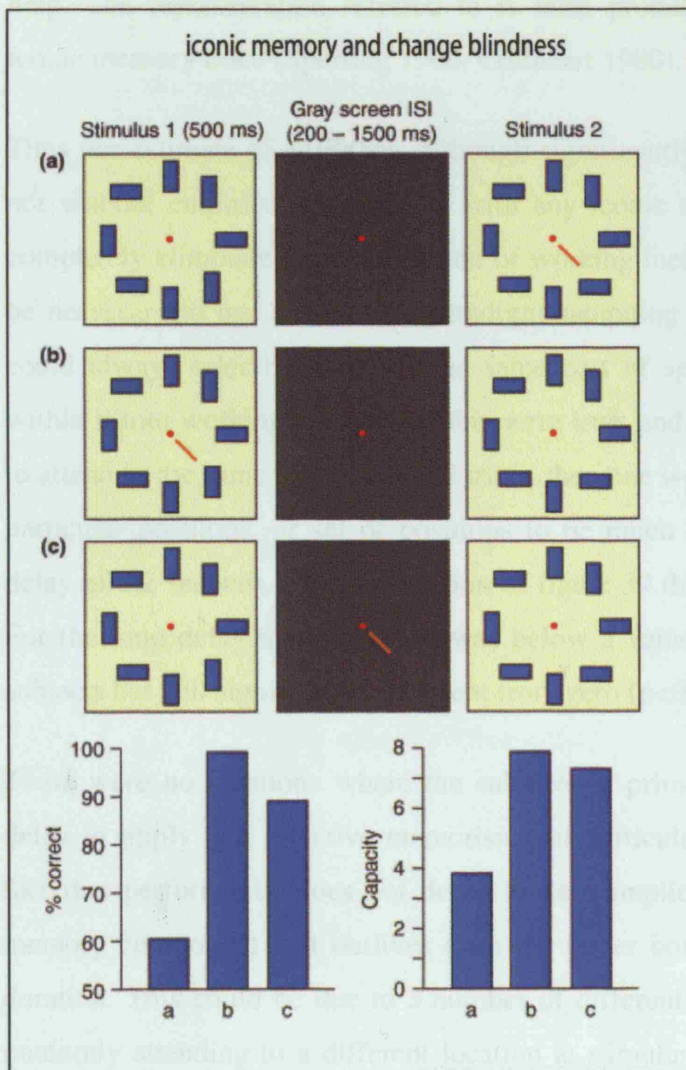


Figure 53 shows results from Landmann and Lamme 2003 which imply iconic memory decays slowly. The task is to detect a change in orientation at the cued location. a) Shows that cueing the location after the onset of the second stimulus results in low performance (just above chance). b) Shows that cueing at the time of the first stimulus onset results in perfect performance. c) Shows that cueing up to 1200ms after the offset of stimulus 1 can still result in high performance. [figure taken from Lamme and Landmann 2003]

The interpretation of Landmann and Lamme is that this indicates that after the first display has disappeared, a neural representation of almost the whole scene is still present and attention can select from this representation to store the relevant item in working memory. After the onset of stimulus 2, this representation has vanished (in that it has presumably been overwritten by stimulus 2), as cueing at that time does not

help. The representation referred to is most probably, in psychological terms, an iconic memory trace (Sperling 1960, Coltheart 1980).

Thus our estimate of ~1000ms, although significantly higher than some estimates, is not without empirical support. As with any iconic memory experiment one cannot completely eliminate the contribution of working memory (to do so it would perhaps be necessary to use a dual task paradigm saturating working memory). The subject could always selectively attend the same part of space and encode whatever falls within it into working memory. If this were true, and a subject persisted in a strategy to attend to the same location on all trials, then one would expect that d' for one particular positions, or set of positions to be much higher than zero at the longest delay of 2.2 seconds. From inspection of figure 39 this was found not to be the case. For the long delay mean d' was below a value of 0.5 for all sessions and all subjects but still significantly different from zero ($p < 0.05$).

There were no locations where the subjects' d' was high enough at the long delay to imply that selective memorising of particular positions was occurring. The fact that performance does not decay to zero implies that there is some additional memory component that outlives even the upper bound estimate of iconic memory duration. This could be due to a number of different sources. The subjects could be randomly attending to a different location at stimulus onset during the delay period, which would result in all positions receiving a small but significant working memory-related advantage. It is also possible that subjects were using a shape related strategy focusing on the form of the stimulus configuration, which could then be encoded into working memory and then later recalled. This is certainly a strategy that could easily be taken if the spatial distribution of discs was unconstrained since discs would group together on a random basis. This strategy was made more difficult however by the fact that discs were randomly assigned to one of two locations in each octant. That we did not detect any activity related to decisions in shape-specific cortices such as LOC or V3 may suggest that a shape related strategy was not being implemented. Other sources to consider would be a weak effect of working memory, equally distributed between stimulus positions or alternatively some priming effect. However these alternatives cannot be deciphered on the basis of this data alone.

It is unlikely that the subject is accessing a lasting afterimage of the disc array since the sinusoidal luminance sequence of the disc array was parameterised to prevent any discernable retinal afterimage. This was tested by asking each subject if they could see an afterimage at the time of the shortest cue for each subject-specific contrast level (200ms). Subjects reported that there was an afterimage with a non-sinusoidal luminance defined disc but not with the sinusoidal version.

The implications for research into the psychophysical characterisation of iconic memory are however not our principle concern. Rather we are interested only enough to establish that the task performed dissociates stimulus processing from report at the medium delay condition.

Why does the delay dissociate stimulus processing from perceptual decision?

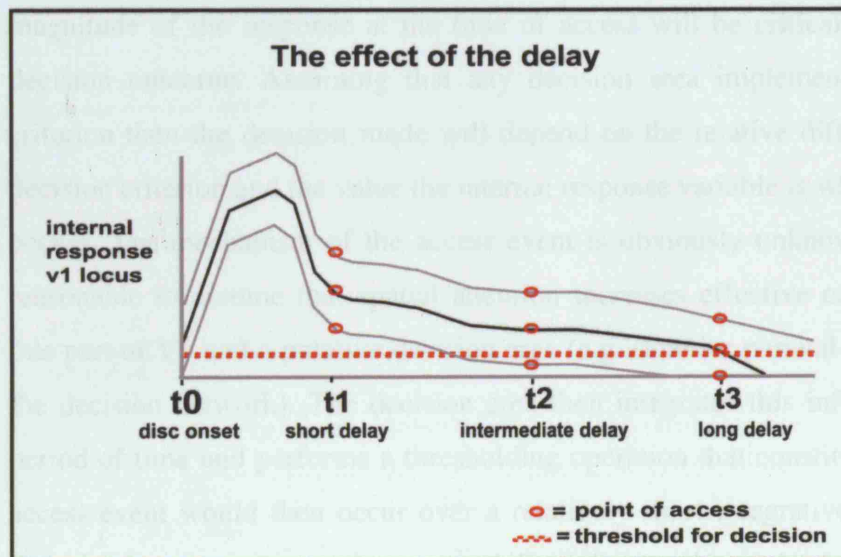


Figure 54 depicts a simple hypothesis for how the time of the cue could affect the accuracy of the detection decision. Because of the variability in the internal response at longer delays certain proportions of responses will fall below the threshold resulting in a change in the proportion of hits and misses.

In this experiment the delay between the stimulus and the cue is experimentally manipulated to decouple stimulus processing from report. One would assume that the delay makes the decision particularly susceptible to noise, as would be predicted by any signal that decays back to baseline noise. So what causes the dissociation between

stimulus processing and report in a partial report task? This cannot be known for sure but we can theorise about the most likely sources of the dissociation

A simple model is presented in the cartoon in figure 54. The cartoon depicts a measure of neural activity (single unit recording or LFP) of a V1 locus (cell or group of cells) in response to the onset of a disc in a position that falls within its receptive field/area. The black line represents the mean response. Naturally because of omnipresent noise in the system (presumably at all levels of visual system, inherited from the retina, and accumulated at each subsequent stage in the system) there is a variable magnitude of response as represented by the upper and lower grey lines. Assuming that the brain accesses this information in order to decide whether the stimulus was present or absent and that the access event is always within a fixed time from the cue onset (for the sake of argument, but see the next paragraph), then the magnitude of the response at the time of access will be critical in determining the decision outcome. Assuming that any decision area implements a fixed decision criterion then the decision made will depend on the relative difference between the decision criterion and the value the internal response variable is when the access event occurs. The mechanism of the access event is obviously unknown, but it would be reasonable to assume that spatial attention increases effective connectivity between this part of V1 and a putative decision area (e.g. superior parietal or any other part of the decision network). The decision area then integrates this information over a set period of time and performs a thresholding operation that constitutes a decision. The access event would then occur over a relatively short integrative temporal window. This is the sort of model supported by primate electrophysiology of perceptual decisions (e.g. Gold and Shadlen 2001, Palmer et al 2005). Regardless of the mechanism of the access event one can reasonably assume that a secondary area has to integrate information from a locus of visual cortex encoding the presence or absence of the disc and that the time at which the integration period begins (time at which it starts with respect to disc offset) determines the strength of that sensory evidence and therefore determines its susceptibility to noise.

Having considered the effect of response magnitude on the decision outcome we can consider the effect that the time of the access event has on performance. Consider the short delay condition (see figure 54). Regardless of the variation in the response, the

majority of access events will access the variable encoded by V1 when it is above its threshold and therefore the report will be highly accurate across repetition. At the medium delay the mean response is still well above the threshold resulting in performance well above chance but the lower bound of the response falls below threshold. Therefore on such trials the report will be 'no' even though the disc was there (miss). This exemplifies the point that one of the potential causes of a non-veridical report (i.e. false alarm or miss) is the variability of the internal response at the time of the access event. This can be due to variability in the magnitude of the response or could be due to the variability of the decay of the response or most probably both. There is another potential source of variability that can equally well account for the variability in response to a constant stimulus and that is variability in the access event itself. One example would be variability in the latency with which the access event arrives after the cue.

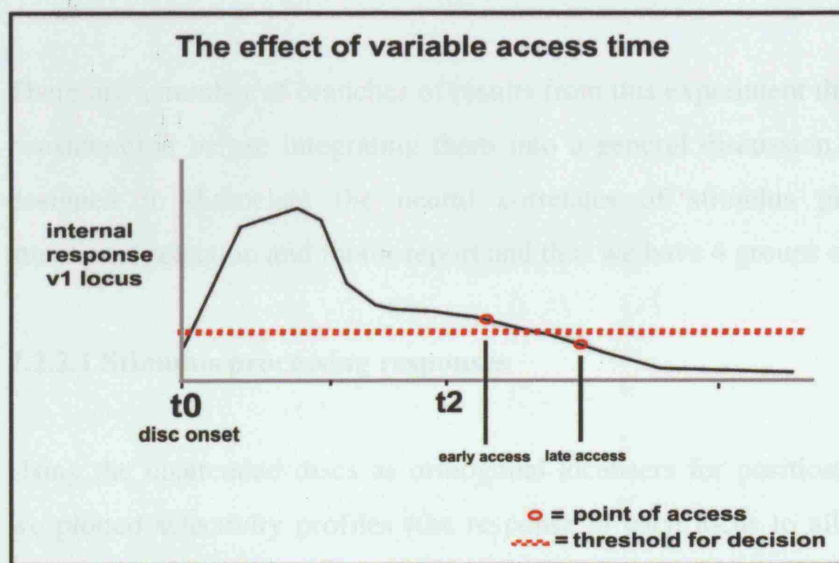


Figure 55 shows a schematic of V1 activity in response to stimulus onset. Depending on the point in time in which the variable represented by V1 is accessed can have different effects in terms of whether it was above or below the set threshold for the detection decision. As can be seen an early access event exceeds threshold and would result in a 'hit' but a late one does not resulting in a 'miss'.

Using the medium delay as an example we can imagine how the variability in the time of access after the delay can affect the outcome of the decision. After the cue occurs

at T2 there will plausibly be some early access events and late access events. If the decision is simply based on the magnitude of the internal variable at the time of access then the difference in time of access could be the difference between a suprathreshold variable (for the early access) and a subthreshold magnitude (for the late access). Another plausible source of variability may include the loss of spatial localisation information over time, such that the spatial information associated with the disc presence is subject to noise which grows over time resulting in less accurate reporting of disc detection.

In summary, one can identify plausible sources of noise that could result in the variability in report that is observed at the different delay conditions. There are of course likely to be many more, but this was an exercise in identifying a simple model that could exhibit the behaviour that we observe from the subjects.

7.2.2 Scanning Results

There are a number of branches of results from this experiment that require individual consideration before integrating them into a general discussion. The paradigm was designed to dissociate the neural correlates of stimulus processing, decision, attentional selection and motor report and thus we have 4 groups of results to discuss.

7.2.2.1 Stimulus processing responses

Using the unattended discs as orthogonal localisers for position selective responses we plotted selectivity profiles (the response of each locus to all stimulus positions) that show the peaked responses to individual positions of the 16-position disc array. This confirmed that our V1 loci had meaningful position selectivities. This was further supported by the location of the V1 loci being concordant with the known principles of retinotopy. Operationally defining these voxels as the V1 loci for each disc position, we extracted the signal detection profiles for each of 16 different cued loci independently. The signal detection profiles showed variation between loci but in general were consistent (see figure 51). The average response of all 16 loci demonstrates a distinct signal detection profile that is characteristic of all 16 loci.

The most important result is that the average signal detection profile for V1 does not correlate with decision or report. A region correlating with decision would show a high response to hits and false alarms and a low response to misses and correct rejections (see figure 40). We found however the average signal detection of all V1 loci the response to both hits and misses is significantly larger than the response to correct rejections and false alarms (figure 51). This indicates that V1 correlates with stimulus presence not decision or report.

Does the data fit with the existing literature?

V1 and signal detection analyses

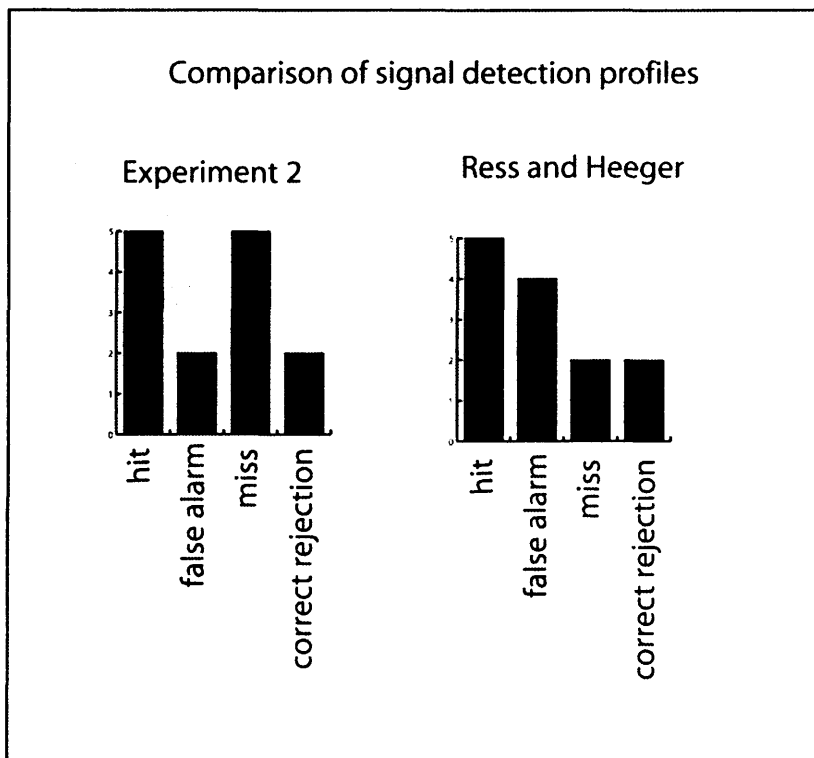


Figure 56 depicts the difference in signal detection profiles between experiment 2 and Ress and Heeger 2004 (R+H from herein). The parameter estimates for R+H are reconstructed for illustration purposes from the peak magnitude of the PSTH's reported in the original R+H paper. If convolved with a canonical HRF these parameter estimates should yield qualitatively similar PSTH's as reported in the original paper.

As discussed in the introduction (page 64) Ress and Heeger (R+H) applied a signal detection theoretic analysis to an fMRI study of contrast detection. The signal detection profile obtained for experiment 2 does not closely resemble that which R+H report (see figure 56). However, due to the differences in experimental design, the data most consistent with R+H's would not predict a replication of their signal detection profile. This paradigm (experiment 2) was designed to dissociate stimulus processing from decision and report, whereas R+H assumed a full association between perception, decision and report. Because of this fundamental difference, the meaning of the signal detection profile in the two different contexts is different.

This difference between the two paradigms is determined by the assumed congruence between perception and report – i.e. that there is a reliable high fidelity mapping between the contents of each. Most NCC experiments, including R+H's, implicitly assume a full congruity between what is perceived and what is reported. This is a deeply pervasive, and often dogmatic assumption, which I will argue is causing much of the disagreement within the field.

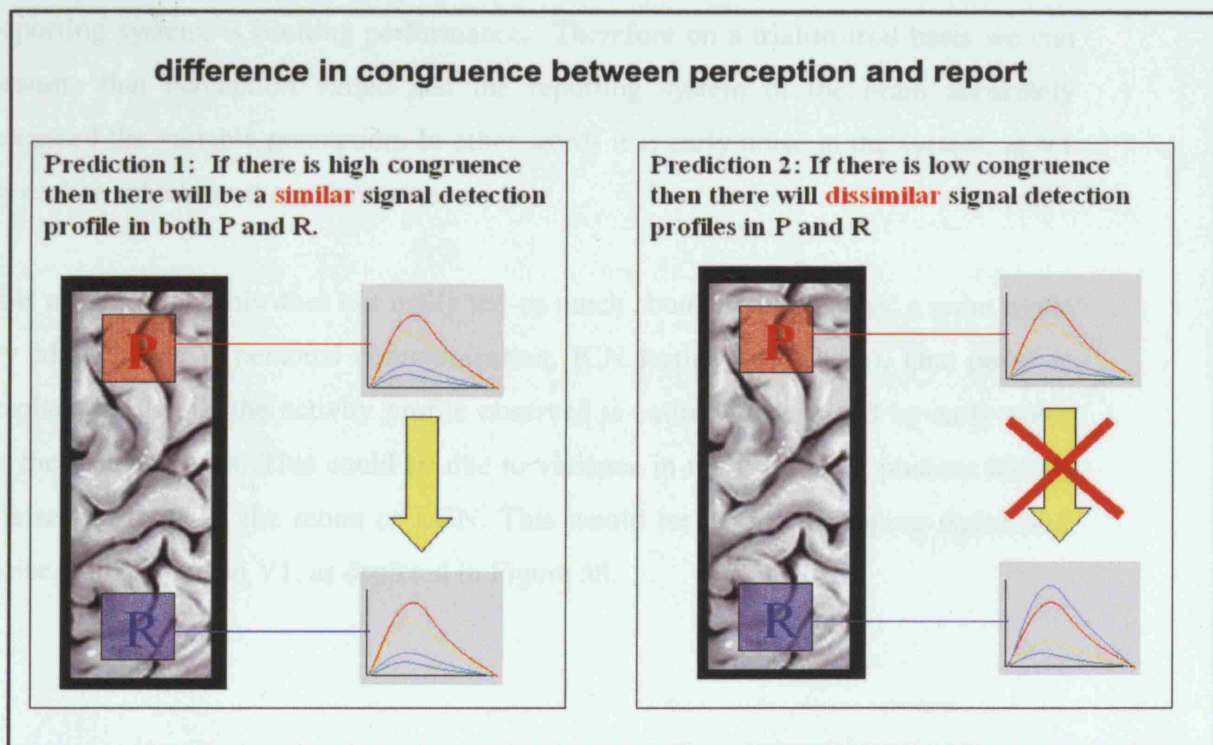


Figure 57 depicts the effect that the congruence between percept and report has on the expected signal detection profiles in the perceptual and report areas. P = an area correlating with perception. R = an area correlating with report.

Prediction 1 characterises the R+H experiment in which perception is assumed to be congruent with report. Prediction 2 characterises Experiment 2 of this thesis where there is a demonstrable (partial) incongruence between perception and report. This has important consequences for the interpretation of the signal detection profiles in each of the two different contexts.

Ress and Heeger (2003) implicitly assumed a high congruence between perception and report, which is reasonable since there was no obvious limitation to the subjects' ability to report the stimulus. Only one stimulus was to be detected and reported, in only one position. The subjects' full attention was directed at the location of the detection judgement and there was only one decision to be remembered and reported. Thus one can reasonably assume that under such conditions the reporting system is operating to its 'full capability'. Therefore there are reasonable grounds on which to assume that whatever was perceived was congruent with what was reported. The variability in report in their experiment is more likely to be attributable to variation in neural activity underlying stimulus processing and perception, since the stimulus was of the threshold contrast. This perhaps means that the perceptual system, not the reporting system, is limiting performance. Therefore on a trial-to-trial basis we can assume that perception varied and the reporting system of the brain accurately accessed the variable perception. In other words it is early noise in the system, at V1 or earlier, which limits performance.

It is arguable that this does not really tell us much about consciousness; a point made by Mike Morgan (personal communication, ICN lecture June 2005). One possible explanation is that the activity profile observed is entirely determined by early noise in the visual system. This could be due to variance in the number of photons hitting the retina, noise in the retina or LGN. This would result in overlapping signal and noise distributions in V1, as depicted in Figure 58.

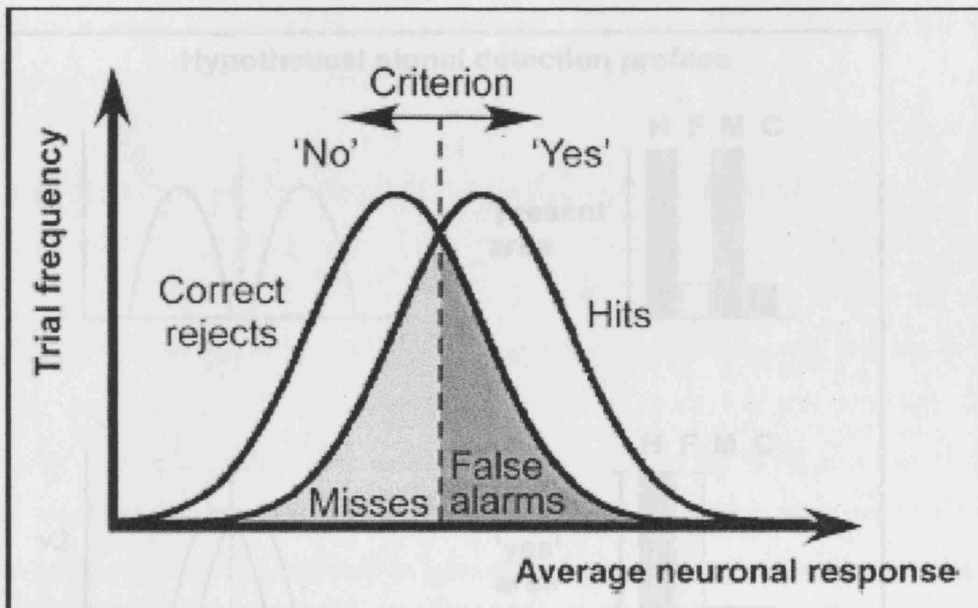


Figure 58 from Ress and Heeger 2004 depicting overlapping noise and signal (plus noise) distributions.

Applying a particular decision criteria will inevitably give the signal detection profile observed of Hits > False alarms > Misses > Correct rejects. Any region that is downstream of V1 will inevitably also have the same profile. Thus the result is trivial because it does not isolate any correlate independently of early noise (upstream) or inherited noise (downstream) in higher visual areas such as V2.

However, I would argue that although apparently trivial, in a sense it is a very important contribution with regard to what it did *not* show. Imagine an alternative scenario in which early noise did not determine the perception and V2 was a 'true' correlate of contrast detection but V1 wasn't, then one may expect a profile of responses such as this:

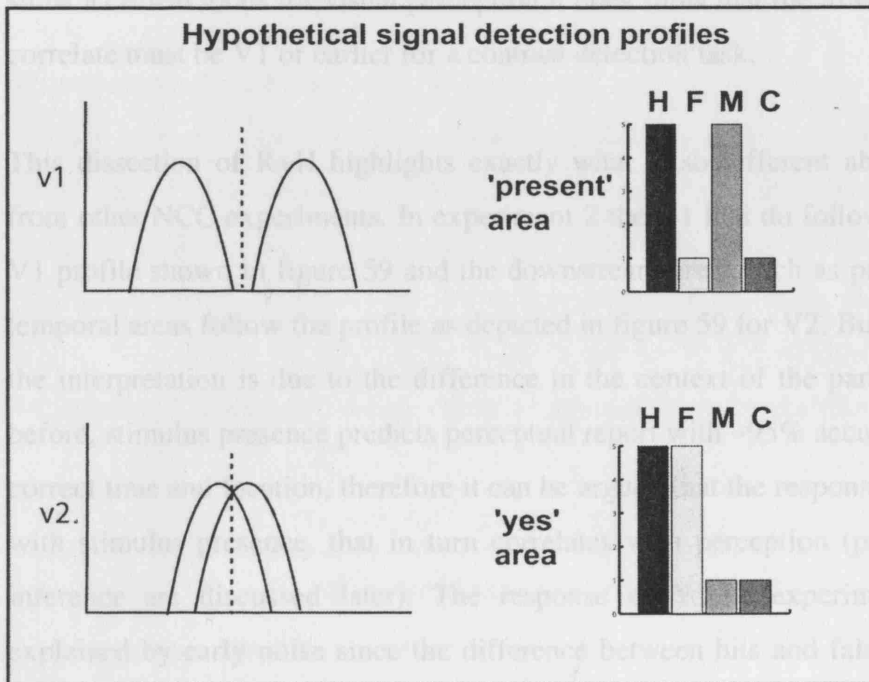


Figure 59 depicts alternative possible results for the Ress and Heeger study if early noise (defined here as prior to V2) did not determine perception. The graphs on the left represent the two distributions of noise and signal + noise, with frequency along the horizontal axis, and internal response magnitude along the vertical. Graphs on the right depict the signal detection profiles associated with each.

In the example above (Figure 59) the early noise in the system does not result in significant overlap between the two distributions. Therefore a decision network sampling only V1 with a fixed criterion would not necessarily lead to mistaken reports (misses and correct rejections). By setting the criteria at the right place the subject could achieve perfect performance. If one found that V2 was the first area in the visual stream to exhibit the sort of distribution overlap which could account for the subjects' behaviour, then it could be argued that the lower envelope of the neural correlate of consciousness (for contrast perception at least) had been found. Of course this would never circumvent the problem that all downstream areas would inherit the same distribution. Solving the downstream inheritance problem would require the sort of experimental manipulations performed in Experiment 2, dissociating percept from report by introducing late noise susceptibility. Although the R+ H experiment fails to

show a critical locus for visual perception it does show that the lower envelope of the correlate must be V1 or earlier for a contrast detection task.

This dissection of R+H highlights exactly what is so different about experiment 2 from other NCC experiments. In experiment 2 the V1 loci do follow the hypothetical V1 profile shown in figure 59 and the downstream areas such as parietal and middle temporal areas follow the profile as depicted in figure 59 for V2. But the difference in the interpretation is due to the difference in the context of the paradigm. As argued before, stimulus presence predicts perceptual report with ~95% accuracy if cued at the correct time and location, therefore it can be argued that the response of V1 correlates with stimulus presence, that in turn correlates with perception (problems with this inference are discussed later). The response of V1 in experiment 2 cannot be explained by early noise since the difference between hits and false alarms is large, whereas an early noise explanation would predict a small difference between the two as seen in R + H. We conclude that this indicates that performance is therefore limited due to late noise beyond V1. The late noise would presumably sit between the V1 and regions comprising the yes and no decision networks.

Because stimulus presence determines perception it confounds it.

A potential problem for the interpretation of experiment 2 is that stimulus presence is always coupled to perception and therefore perception and stimulus are confounded. This raises the question of how to determine what V1 correlates with. V1 could either be entirely driven unconsciously by the stimulus, its activity could be a perceptual correlate or both. This is not too much of a concern given that the experiment was designed to move away from the common strategy of 'keep stimulus same change perception', rather to dissociate the system further down the stimulus-response chain. To make this distinction clearer I will compare the two in more detail.

Most other experiments on consciousness have utilised the following logic: *'keep stimulus same, vary perception and report together, assume report accurately indexes perception - anything that correlates with the change is the correlate of consciousness.'* This suffers from the fallacy of conflating perception and reportability.

Experiment 2 utilises: *'vary stimulus and perception together, change decision, change report - anything that correlates with the stimulus is a candidate* for a correlate of perception. Anything that correlates with decision is a correlate of decision, and likewise anything which correlates with report is a correlate of report.'*
 This suffers from the fallacy of conflating stimulus processing and perception.

*candidate is used here because stimulus processing and perception covary.

These two different experimental logics are compared in the diagram below.

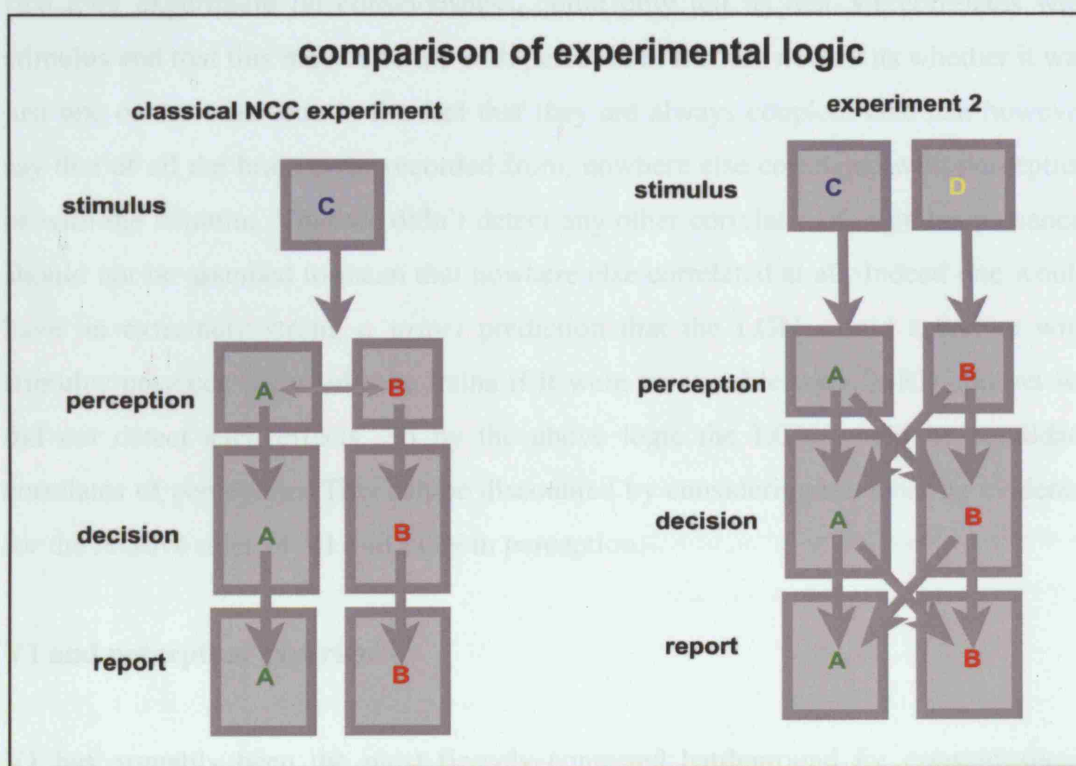


Figure 60 depicts the experimental logic of two types of paradigm. On the left is the canonical NCC experiment of which there are many examples. On the right is the paradigm logic employed by experiment 2 of this thesis.

From inspection of the classical NCC experiment diagram (see figure 60) one would expect that not only V1 would correlate with perception but anywhere up stream of it would also correlate including decision and motor regions. This is arguably why most experiments using this design show widespread neural correlates of consciousness

encompassing occipital, parietal and frontal regions. Through consideration of this schematic such claims are clearly suspect.

Now consider the right hand schematic of figure 60, where there is a coupling between stimulus and perception. Therefore, if one were to look for every brain event that correlated with perception then one would find neural correlates of perception wherever the stimulus was processed, i.e. from the retina up. Clearly this is a logical flaw which must be considered when interpreting the results.

So what do the results of experiment 2 tell us? The result on its own, if it were the first ever experiment on consciousness, could only tell us that V1 correlates with stimulus and that this may correlate with perception. It could not tell us whether it was just one or the other due to the fact that they are always coupled. One can however say that of all the brain areas recorded from, nowhere else correlated with perception or with the stimulus. That we didn't detect any other correlates of stimulus presence, should not be assumed to mean that nowhere else correlated at all. Indeed one would have an extremely strong *a priori* prediction that the LGN would correlate with stimulus presence (as would the retina if it were measurable with fMRI) and yet we did not detect such effects. So by the above logic the LGN would be candidate correlates of perception. This can be discounted by considering the existing evidence for the relative roles of V1 and LGN in perception.

V1 and perceptual experience

V1 has arguably been the most fiercely contested battleground for consciousness. Different models make different assumptions about the role of V1 in consciousness (for a longer discussion see Tong 2003). The **localizationist** camp approach proposes that the experience of specific visual dimensions such as luminance and basic contrast is critically dependent on V1 activity *alone* (necessary and sufficient) (Pollen 1999). The **hierarchical** camp claim that V1 is only critical in the sense that it is the key distributor of visual information to extrastriate areas. The canonical and arguably most controversial hierarchical theory is put forward in the frontal lobe hypothesis (Crick and Koch 1995), whereby in order for an area to be a true correlate of consciousness it must a) encode explicitly the dimension which it contributes and b)

project directly into the frontal lobe. This is motivated by the assertion that the function of consciousness is for acting and planning and thus if an area does not directly drive frontal cortex (assumed according to the literature to be responsible for action and planning) then it cannot contribute directly to consciousness. The **interactionist** camp claim that it is the interactions between V1 and other areas which is the key correlate of visual experience (a key example being Lamme as discussed previously).

Of all these positions, I would argue that it is the evidence from the localizationist camp which is the most convincing (for a defense of locationism against secretion theory see appendix 3). From the known properties of V1 it has been argued that if V1 were to make any contribution to conscious vision one of the strongest cases would be for static achromatic luminance based visual experience (Pollen 1999). V1 and V2 lesions cause retinotopically specific scotomas which match the location of retinotopic lesions (as revealed by bullet wounds during the 1st world war, Inouye 2001). A key issue is whether such lesions indicate that V1 is essential for visual experience or whether the scotomas caused are due to the fact that V1 is depriving down stream areas (cortical areas higher up in the hierarchy) of input, a key difference between the localizationist and the hierarchical camps. The strongest evidence against the hierarchical camp is the fact that lesions beyond V2 do not result in visual field defects whereby *distinct portions of space* cannot be perceived (Horton and Hoyt 1991). Given these restrictions, V2 may lay equal claim to such a role in subserving luminance-based experience. However V2 lesions cause no change in visual acuity or contrast sensitivity, whereas V1 lesions devastated vision (Miller 1980). The integrity of V1 therefore appears to be essential for perceptual experience of low-level fine-grained luminance. As argued by Pollen, this position is supported by the fact that static achromatic experience can remain intact despite a multiplicity of diverse lesions to parietal, temporal, frontal and LGN. In fact never has a visual field cut been reported from damage to anywhere beyond V1 (Pollen 1999).

There is further support for the link between V1 and perception (Pollen 1999) which derives from our phenomenology, in which we experience fine visual detail and retinotopic location in precise correspondence. This is surprising given that most of the brain decouples retinotopic location from spatial detail information. Pollen argues

that if one accepts the premise that an explicit representation is required for an area to contribute to consciousness (Crick and Koch 1995) then the only cells in the brain which code an explicit conjoint representation of fine detail and retinotopic location are phase-specific simple cells in V1 and V2. Beyond V1 and V2 location and luminance based information is decoupled, and even within V1 or V2 complex cells are unsuited to the lowest dimensions of visual experience since they do not signal polarity of brightness. Thus the finest conjoined representation of retinotopic location and spatial detail lies in the simple cells of V1. V1 and V2 are the only cortical regions to contain the a fine-grained representation of the visual panorama and thus serves as the only region of the brain which could satisfy the necessary processing requirements to explain the fact that we have such fine-grained visual experience at precise locations relative to the line of sight.

The opposing evidence against V1 correlating with phenomenal experience has gained a momentum of its own and principally comes in two flavours.

1. Evidence that V1 is activated by stimuli that are not consciously perceived

He et al. provide evidence that V1 can be activated in the absence of awareness, by showing that there can be orientation specific adaptation effects to gratings which are of such high spatial frequency that they cannot be perceived (He et al 1996). In a second experiment, He et al showed orientation specific adaptation for gratings, which were crowded out of perception indicating that V1 neurons were still active even in the absence of perception (He et al 1996). The fact that V1 or any other area can be demonstrated to be active in the absence of consciousness does not deny the role of such areas in consciousness, but rather suggests that more than just any activity is necessary. Caution should always be taken when it is stated that subjects are unconscious since there are many possible explanations for why reportability (the basis on which consciousness is judged) could fail without a concomitant failure in perception. For instance it could that memory has failed before there is chance to report, or that it is possible for perceptions to be reported through some motor modalities but not others. But even if we assume that in no way are subjects aware of these unconscious stimuli the evidence is not sufficient to undermine V1 as a minimally sufficient correlate of consciousness.

2. Evidence that damage outside of V1 can result in a lack of visual experience

The most obvious example is that of hemi-spatial neglect (Critchley 1953), whereby patients cannot attend to the left side of visual space and act as if it does not exist. This is commonly interpreted as a problem with the orientation of attentional resources to one side of space. However, this interpretation is complicated by the existence of extinction, where the patient can identify a test object in the left when it is presented in isolation, but not when presented simultaneously with another 'extinguishing' object on the right. This is used often to argue that early visual areas alone are not sufficient for visual experience (see Tog 2003). In light of the discussion of the subtleties of reportability, one has to question what it is that is lacking in these patients. How do we know that they lack visual experience? Experimenters infer a lack of awareness because they cannot report anything in their neglected field. But recall the fallacy of conflating reportability with perception. If the mechanisms of reportability are anatomically and functionally distinct from the mechanisms of phenomenal experience, as I propose, then neglect can easily be explained without undermining the hypothesis that V1 alone is necessary and sufficient for certain low-level dimensions of visual experience. The explanation (which has only previously been posed as a 'question' Lamme 2003, see also 'future questions' section) is that neglect patients in fact retain a phenomenal experience in their neglected field. In this sense they have a similar experience as normal subjects, but are unable to direct their attentional selection networks, decision networks and consequently motoric report networks to access the information encoded in sensory cortices within one hemisphere. Proving this interpretation runs into the same problems discussed previously when talking about the refrigerator light illusion (page 46). I would argue that since this is the case it should be taken equally seriously as the assumption that neglect patients *lack* perceptual experience in their neglected fields.

The LGN and perceptual experience

The LGN is a 1st order relay thalamic nucleus which receives 90% of retinal output (Sherman and Guillery 1998). Surprisingly for its widely known visual function only

5% of its inputs comes from the retina, the remainder arriving from the brainstem or from cortical feedback. Despite the small proportion of inputs from the retina, the fact that these are all driving inputs, which are relayed with little transformation in receptive field properties to the cortex highlights its importance in visual function. Despite its critical importance for visual function the LGN does not appear to be essential for perception since elemental visual experiences of white lights can be evoked in humans after severance of connections between V1 and LGN (Brindley and Lewin 1968, Dobelle and Miladejovsky 1974). When considering that the LGN does not display explicit representations for colour or form it then seems that there is little evidence for its direct contribution to perceptual experience of the agreed dimension. The results of experiment 2 suggest that the neural correlates of perceptual experience for luminance based discs lie in V1/V2 or earlier. Since it has been persuasively argued that LGN is not necessary for perceptual experience then one could argue that the critical locus is in V1/V2 or beyond. Since the only region which we recorded which showed a spatially selective stimulus induced response was V1/V2 we therefore propose that this is consistent with the localizationist model of visual consciousness.

7.2.2.2. Decision specific responses

The second branch of results pertains to the delineation of a network of decision specific responses that occur independently of stimulus processing and report mode. This network of areas was relatively consistent across all 3 subjects and included superior parietal, frontal, and temporal regions. There were differences within this network with a subset of regions expressing higher activity for yes responses and another subset of regions expressing higher activity for no responses.

Is this supported by the literature?

There has been a wealth of studies performed on perceptual decisions with electrophysiology and fMRI. There is strong experimental support for perceptual decisions to be computed by parietal and prefrontal regions (for review see Schall 2001). Although we report superior parietal activity in all subjects for both yes and no decisions, there was no decision related response in DLPFC, or any other frontal region (as had been reported previously e.g. Heekeren 2004) The fact that we do not find a prefrontal correlate of performance speaks against an interpretation of subjects using working memory to perform the task. DLPFC may be important for decisions which require, active comparison between two entities across space, (Heekeren et al 2004) or time (Fuster 2001)

In a contrast detection task similar to that of Ress and Heeger using whole brain rather than ROI (regions of interest) analysis, Pinns and Ffytche 2003 found a similar distribution of areas as we report in experiment 2. For the 'Yes' network they reported increased activity in thalamus, striate cortex, fusiform gyrus, medial occipital lobe, sensorimotor cortex, cerebellum, insula, and posterior supplementary motor area. For the 'No' network areas of suppression were found bilaterally in the supra-marginal gyrus (parietal), the posterior cingulate and parahippocampal gyrus (limbic) and in an extended region passing from the superior frontal sulcus through the medial frontal gyrus to the anterior cingulate gyrus (frontal). There are differences however between our paradigm and Pinns and Ffytche in that a) theirs was detection of a single event with no attentional orientation b) they only reported hits vs. misses whereas we were

looking at the full compliment of signal detection categories. Superficially the results are in partial agreement with the networks delineated here in that; there is an asymmetry of areas responding more to yes decisions than to no decisions and they are distributed in a similar network of areas.

In this experiment multiple distantly connected areas, resembling a neuronal workspace, as described in Deheane's theory, are activated together in response to different decisions. All such areas may not necessarily be critical for the task, but may, by virtue of being dynamically mobilised together, still have access to the decision related information. This makes sense if attention allows for subjects to act upon perceptual information in a way that is free for flexible use by multiple modalities including cognition, language and motor action. This result does therefore offer some support for the global neuronal workspace (Deheane and Naccache 2001) model which posits a network of distantly connected areas which have access to each other and share information. In the context of this experiment however, I would propose a different interpretation for the model, suggesting that it is a model of reportability rather than conscious perception *per se*. In summary, on the one hand this interpretation would be supportive of the global neuronal workspace model as a mechanism of sharing information, but on the other it would considerably change the interpretation of it being an all-encompassing theory of consciousness. Critically it would contradict the claims that the workspace is necessary for visibility, or perception.

7.2.2.3. Attentional regions

As discussed we reported standard fronto-parietal attentionotopic activity which was expected from previous studies. Proper attentionotopic mapping of cortex would require flat mapping techniques and would have been another thesis worth of work. Since this was not part of the experimental objectives we therefore chose not to pursue a full cortical attentionotopic mapping of cortex. Surprisingly however we observed consistent attentionotopic activity in the posterior thalamus, in a region most likely to be the dorsal pulvinar.

The pulvinar and attention

Theories of spatial attention (Shipp 2003) and salience (Robinson and Peterson 1992) place a central role on the pulvinar nucleus of the thalamus. The rationale is based largely on it being a higher order thalamic relay (which means it does not receive direct projections from the retina unlike the first order relay LGN) whose anatomical connections are well suited to exerting simultaneous and remote modulation of all visual cortical areas. This is potentially achieved by the bi-directional connectivity with visual areas which unlike the cortico-cortical feedback connections are thought to be capable of driving cortical regions (Sherman and Guillery 1998). This is particularly critical since it is known that attention can enhance the baseline firing of cortical visual neurons (Driver and Frith 2000). If cortico-cortical feedback from fronto-parietal can only modulate ongoing activity, then any baseline shift is likely to originate from the thalamus and if it is attention related then this is most likely to be from the pulvinar. Thus a thalamic region which has a topographic map of attentional space is of considerable theoretical interest for the neural basis of attentional effects. Direct evidence for the attentional role of the pulvinar has remained largely elusive. This may be due to the difficulty that experimenters have experienced in characterising the receptive field properties of pulvinar cells, which has resulted in relatively few electrophysiological studies on the pulvinar despite it being the largest of the thalamic nuclei. Likewise fMRI studies have only infrequently reported pulvinar activations for attentional effects, possibly due to the size of the nucleus in relation to voxel sizes. Despite this some fMRI studies have shown reliable pulvinar activation under conditions of fear (Morris et al 1998) attentional shifts (Haynes and Rees 2005, Kastner et al 2004), however there has yet to be a satisfactory functional mapping of the pulvinar's visuospatial topography. That this has not been shown previously is particularly surprising given the pulvinar's striking visuotopic anatomical organisation and its proposed role in spatial attentional shifts (Shipp 2003). It has been previously hypothesised, on the grounds of anatomical connectivity, that the pulvinar would encode spatial attentional shifts. The results of experiment 2 are congruent with such a hypothesis. The results show that for each subject there is attentionotopic activity in the posterior thalamus and that the anatomical locations is consistent with the hypothesis that these activations are in the medio-dorsal portion of the pulvinar. We found that each cluster had a full field representation of

attentionotopic shifts of different polar angles away from fixation. This is perhaps in contrast to the expectation from the primate anatomical literature which shows a lateralisation of each hemifield in the contralateral hemisphere (Shipp 2003). Whether this duplicate full field representation within the pulvinar holds only for parafoveal locations cannot be ascertained from this data and will be of interest for future pulvinar mapping studies. The attentionotopic maps that we report were not stimulus driven since the equivalent plots for the main effect of stimulus presentation and a control conditions in which cue alone was present in the absence of a task, did not show any activity in the pulvinar.

7.2.2.4. Report regions

Regions correlating with report should be predictable from the existing literature on correlates of button pressing. As would be predicted left and right hand reports revealed contralateral motor cortex and ipsilateral cerebellum.

Does this fit with the literature?

This is a well-studied area, so no discussion is required to discuss its agreement with previous data that motor actions correlate with primary motor, premotor and cerebellum activity.

7.3 Interpretational problems

Having discussed the experimental findings in the context of the existing scientific literature, there are still problems of interpretation which relate to enduring theoretical issues. The key point of contention is whether stimuli outside of attention or cued late are phenomenally *experienced*, or whether they are only experienced when attended and reported. As stated previously two important positions on this issue are exemplified by the work of Block and Dohaene. I will now give interpretations of my results based on these two opposing positions.

A Blockian interpretation:

The Blockian interpretation of this experimental design would be that the physical presence of a stimulus at a particular location is a better index of whether the subject perceived the stimuli or not than the subject's own report.

The psychophysical results showed that if the subject was cued to position x 200ms after the stimulus onset, they would be ~95% correct in reporting the presence or absence of the stimulus at that position. The disc location was thus highly reportable, and according to standard definitions therefore highly perceptible. One can however go beyond that. The subject does not know on each trial *which* position is going to be cued, or *when* the cue is going to appear, but subjects can still report with ~95% accuracy whether the stimulus was present at any location. We can therefore claim with ~95% confidence that they had a reportable perception of every position. This is the logic of the partial report superiority effect: one samples different parts of space and finds the probability of a perceptual event having occurred, even though only one position is sampled by attention, and thus reported, on each trial. This probability then refers to what *would have been reportable* had that perception been sampled, and allows one to make inferences about percepts that are not actually reported, or even reported mistakenly, on every trial. In my experiment this approach allows me to

conclude that every position in the stimulus display has the potential for report, therefore every position is reportable, and therefore every position is perceptible.

There are then two ways, in which this Blockian interpretation allows us to make inferences about perceptual experience from the presence of the stimulus, rather than from the subject's report,

The disc locations that are uncued are not reported. But had subjects been cued they would have been reported with ~95% accuracy. We can therefore infer with confidence what the perception was at each of the unattended locations just by knowing which stimuli were present at which location.

The disc location is cued, but cued late such that it is reported incorrectly. We can infer that had the cue arrived earlier, subjects could have reported accurately, therefore the perception can be inferred from stimulus presence better than it can by the actual report. This is clear with a more extreme example. If one cued the subject 10 minutes later (after the perception of the discs) then one would not expect the subject to remember. If they are forced to guess and they would likely get it wrong, one would not assume this necessarily indicated anything about their perception.

There are of course several problems with the Blockian interpretation, which shall be discussed shortly.

A Dehaenian interpretation

A Dehaenian interpretation would be that we do not in fact have a perception of every location, but that we only perceive the discs when they are cued or erroneously attended. The fact that the discs are demonstrably potentially accessible (referred to previously as the subset of I2), does *not* mean that they are consciously perceived. The state of such stimuli is referred to as 'pre-conscious' (Deheane et al 2006), a state which is processed sufficiently such that it is available to the global workspace if

selected by a top-down ignition process (in other words attention), but is not in fact phenomenally experienced.

Mediating between two opposing interpretations

The stimuli which would have been reportable with 95% accuracy *had they been cued* are 'perceptually experienced' under the Blockian interpretation, but are 'pre-conscious' under the Dehaeneian interpretation. As I have emphasised throughout, and discuss in the final section, many apparent contradictions in consciousness research can be resolved by aligning conflicting terminologies. However, this is not the case here. Although these two theorists may use similar terms in different ways, there is a fundamental difference between the two hypotheses, and problems with each.

Problems with the Blockian interpretation

There is a problem with a strict Blockian interpretation, since there is a possibility of a measurement effect, whereby the process one is measuring is contaminated by the act of measuring itself. In this context, one is trying to infer the existence of a perception, by attending to it and reporting it. 'Measuring' the putative perception in this way may change it, such that the measured perception is then *different* to perception at the unattended locations. The measurement effect is most problematic for the short cue condition, because of the temporal proximity between the cue and the discs array. It is possible that subjects do not in fact perceive all the discs at the time of onset (within a reasonable latency 100-500ms post stimulus onset). Rather, the arrival of the cue 200ms afterwards changes visual processing, facilitating perception to raise it above the limen (supposing there is such a thing). Under this scenario it may appear that in the short cue condition subjects are able to perceive every disc prior to report, but in reality they can only report because of the interaction between the cue and the stimulus changing perception itself.

In backward masking, a mask presented after a stimulus has appeared often affects the perception of the preceding stimulus. We might worry that this kind of effect is operating in our experiment, given that the cue in the short delay trials is presented

200ms after the offset of the disc array and therefore 400ms after the onset. This would then lead to a criticism of the Blockian interpretation of the kind just laid out. However, there are important differences between this paradigm and masking. Crucially, the effect of a mask is usually strongest within 100ms of the stimulus offset. Most visibility-SOA functions are U-shaped with the visibility returning to full strength just beyond 100-150ms (e.g. Haynes and Ress 2005). Therefore the 200ms SOA between the disc onset and the cue in our experiment should be beyond the reach of any backward masking effects.

In addition to the temporal distance between the cue and the stimulus, there is a spatial distance of 2 degrees of visual angle. Is it possible that the cue could effect the perception of the disc when the temporal distance is 200ms and the spatial distance is 2 degrees? There is no evidence that we know of that could support an interaction between cue onset and the perception of a stimulus 200ms and 2 degrees apart. Furthermore, any interaction between the two would be presumably be negative. Psychophysical testing (outside of the scanner) showed that accuracy is even higher (>97%) for cue delays of 100ms. If the arrival of the cue after 200ms does not change the perception of a disc then it is unlikely that the arrival of a cue after ~1000ms is going to effect perception. Evidence in support of this is that in the preliminary psychophysics, subjects rated the visibility of the disc array for all delays and there was no effect of delay on visibility. In fact, the visibility of the discs was rated as 10 for all delays for all subjects.

Problems with the Dehaenian interpretation

The Dehaenian interpretation of a pre-conscious stimulus being perceived only when accessed by the workspace is plausible when the time difference between the stimulus and the access event is short, however this plausibility is strained somewhat by considering the ~1000ms SOA condition. It should be noted here that the term 'pre-conscious' can cause confusion – it's suggested abbreviation to 'p-conscious' implies that it is a kind of consciousness, whilst in fact p-conscious stimuli are not consciously experienced at all. In addition, this is the same abbreviation used for years by others for phenomenal consciousness). In Dehaene's theory, only when pre-conscious V1 representations enter the workspace via top-down ignition by spatial

attention do the discs become conscious. For this argument to hold in our experiment one would have to argue that perception has not occurred by the time the cue arrives (1 second after stimulus offset). Since the discs cannot be accessed and reported for 1 second, this would imply that the discs remain preconscious for that duration and only become conscious after the cue. This is of course logically possible, but if it were the case then the role of consciousness would be reduced to a causally impotent epiphenomenon. If perception were to be of any use for guiding voluntary action it would need to complete within a shorter timescale - 1 second would be far too long for our volitional actions to lag behind the world. As it stands this a rather absurd scenario, but perhaps this is a straw man constructed through extrapolation of previous statements (see Dehaene and Naccache 2001). To avoid this accusation, we can reformulate the Dehaenian theory to make it better able to explain the results even at 1000ms cue delay. This requires us to suppose that the workspace will access a subset of the discs *during* the 1-second duration of the delay, since covert attention can move around in this time. With covert attention moving around, single locations or sets of locations will be accessed and therefore experienced, as well as the location which is accessed at the time of the cue. In summary, subsets of discs are accessed and experienced before the onset of the cue, when the cue arrives it redirects the workspace to access the target location. Even this is not particularly convincing since the experience of the discs is not of single discs lighting up in succession as ones attention moves about. Furthermore subjects perceive the discs as occurring long before the cue arrives. If the last explanation were correct, subjects would only experience discs as and when they attend which would be after the cued onset.

The argument still boils down to the refrigerator light illusion

So both positions can explain the results, and both have logical flaws which we can argue against from various pieces of evidence and theoretical speculations. However, neither appears to be ruled out conclusively by the data, so how can we tell between them?



Figure 55B. The refrigerator light. Is it always on or does it only turn on when one opens the door to check? Analogously do we experience objects outside the focus of attention or does that experience only begin when it is checked by attention.

In both theories there are 3 states stimuli can be in; inaccessible to report (I1), potentially accessible (I2) and accessed (I3). Both sides also make roughly the same physiological and psychological predictions about these states. Both assert that the difference between I1 and I2 is some essential dynamic within sensory areas, and that the difference between I2 and I3 is due to recurrent processing between frontoparietal areas engaging with those recurrently processed representations in sensory areas. Both theories would place an uncued disc as being part of I2 and a cued disc as I3. The only difference in interpretation is whether I2 is phenomenally conscious (under the Blockian scheme) or whether it is Preconscious (under the Dehaenian scheme). I know of no way to test between these two opposing interpretations. The 'refrigerator light' thought experiment (see Figure 55B) has been used to illustrate this problem – if we didn't understand the mechanism, we might wonder whether a refrigerator light is always on, or does it only turn on when one opens the door to check? Analogously, Block and Dehaene give opposing answers to the question: Do we experience objects outside the focus of attention, or does that experience only begin when it is checked by attention?

A new perspective

Unfortunately this issue cannot be resolved by experiment 2 nor any other experiment to date. It seems to remain to be a philosophical argument for the time being at least, since it is hard to imagine an experiment which would prove one way or another whether the potentially reportable stimuli are phenomenally experienced or not. One could choose to take a position, based partly on philosophical preference, and partly on deciding which methodological objections are most likely to be true. Alternatively, one may choose to be agnostic about the issue, which is probably the most scientifically defensible position. In this sense it might seem that the initial aim of experiment 2 was somewhat misdirected. Originally, the aim was to use the Sperling paradigm as a way of dissociating phenomenal consciousness from access consciousness and thus dissociate their neural correlates. However in light of this discussion, it is currently impossible for anyone to scientifically prove one position over another.

I will now take the scientifically agnostic view to summarise the results of Experiment 2. This Experiment was intended to dissect the component processes of reportability, and better understand their relationship to the distinction between phenomenal and access consciousness. The process of dissection is novel and important in its own right, and marks a first step towards a neuroscientific understanding of the neural components underlying stimulus reportability. However, at this point, I will not make any strong claims about the relationship to phenomenal and access consciousness, due to the problem of interpretation I have discussed here.

Experimental conclusions of experiment 2

In summary the results of experiment 2 show that it is possible to decouple the neural correlates of 4 component operations underlying the reportability of a stimulus. We found that retinotopically specific regions of V1/V2 correlated with stimulus processing, a frontoparietal network including the pulvinar correlated with attentional selection of spatial locations, a network of parietal, frontal and temporal regions correlated with perceptual decision and finally the motor system including PMC and cerebellum correlated with motoric report.

8. Theoretical speculations

Disclaimer

In the next section I will expand on the theoretical speculations pertaining to experiments 1 and 2. The data in this thesis does not in any way show either of these theoretical proposals to be the case, but is consistent with such accounts, and the delineation of such speculative proposals can serve as predictions for future studies.

8.1 Theoretical Proposal for Experiment 1.

Having established that there is no measurable difference (not measurable with the techniques employed) in object selective regions whether the object is visible or hidden, is there a theory that can plausibly account for the difference between hidden and visible objects?

As discussed earlier a popular theory of visibility invokes the notion of feedback to V1 as being necessary to 'see'. This stems from the late modulation effects recorded by Lamme and colleagues. It is of interest here that a valid interpretation of the feedback modulatory effects Lamme observes with figure ground segregation could be entirely attributable to binding. Lamme is recording from V1 cells that optimally respond to the orientation of lines presented to their receptive field. If the line is embedded on a figure then there is late rate modulation after approximately 100ms. Otherwise there is just the normal response to an oriented line. This is known as an extra-classical receptive field property because it cannot be explained in terms of its normal receptive field characteristics. The interesting thing is that if the primate fails to recognise that it is part of a figure then no rate modulation is observed. As the theory goes, the late modulation responses in V1 is a correlate of the monkey's conscious experience (Lamme 2003). There is however another interpretation however which is depicted in the following figure.

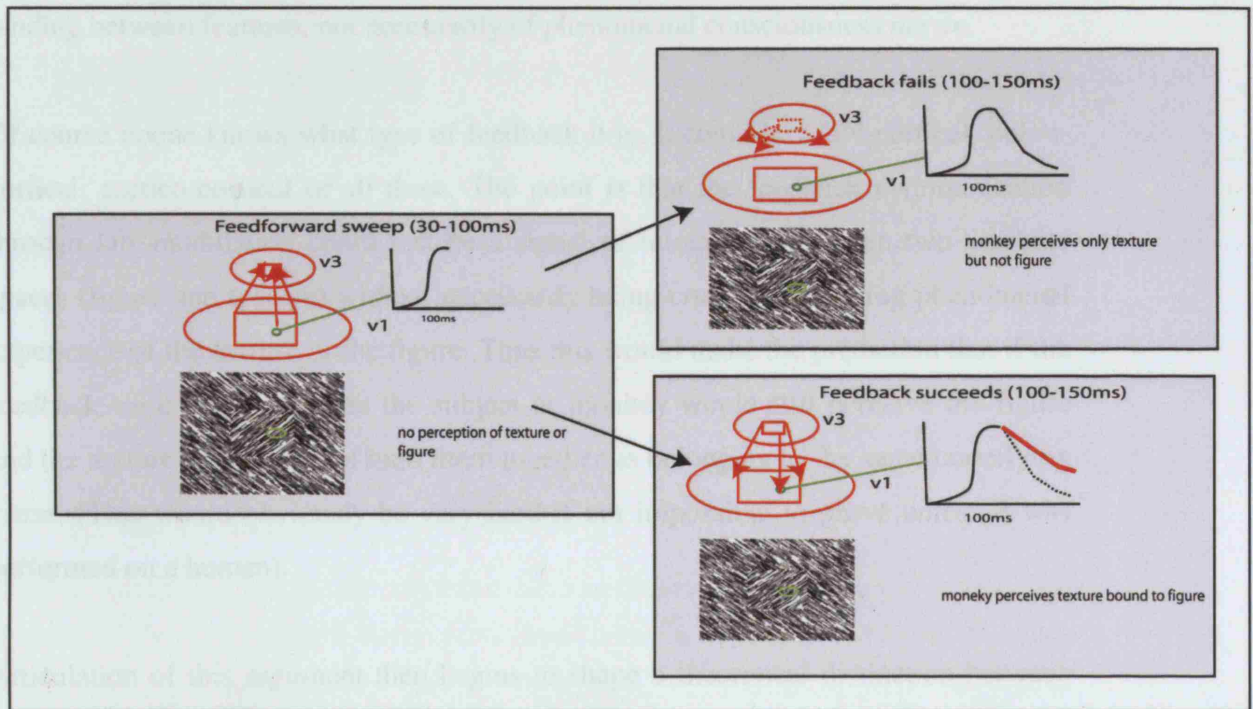


Figure 61 shows hypothetical neural events underlying successful or unsuccessful figure ground segregation.

Supposing that the V1 cell is part of a population of cells in V1 representing the retinotopic space subtended by the figure. The cell recorded from has orientation selectivity that matches that of the portion of texture in its receptive field. Thus the presentation of the stimulus causes the V1 cell to fire at an increasing rate up to 100ms. Meanwhile the information pertaining to the texture has been forwarded to a form responsive area such as V3. V3 integrates information from a larger population field of V1 cells and thus can compute the presence or absence of the figure. If the detection of the figure fails for whatever reason (does not reach threshold) then no feedback is sent back to V1. The lack of feedback does not necessarily mean however that the monkey did not perceive anything; rather one would assume it just perceived the texture on its own and therefore the correlate of that perception is that depicted in the upper right graph of the figure. If the computation of the figure succeeds, then a feedback signal is generated such that it causes a late modulation after 120ms. If the state of V3 is the explicit representation of the figure, and V1 is the explicit representation of the components of the texture, then for the two to be perceptually bound together it makes sense that the binding signal is indexed by the late

modulation. Late modulation through feedback then is plausibly a mechanism of binding between features, not necessarily of phenomenal consciousness *per se*.

Of course no one knows what type of feedback it is. It could be LGN-cortical, pulvinar-cortical, cortico-cortical or all three. The point is that the feedback communication through late-modulation could just be a signal of integration between two attribute spaces (figure and texture) without necessarily being crucial for causing phenomenal experience of the texture or the figure. Thus this would make the prediction that if the feedback were prevented then the subject or monkey would still perceive the figure and the texture but would not bind them together as belonging to the same underlying cause. (This would obviously be very hard if not impossible to prove unless it was performed on a human).

Articulation of this argument then begins to shape a theoretical distinction between modal and amodal perception. The failure or success of feedback between higher visual areas beyond V1 to establish an integrative state with V1 could plausibly be the difference between modal and amodal perception. Modal perception could occur when there is binding between V1 and extrastriate regions such that whatever is represented in extrastriate regions is integrated with the V1 representation of the surface which caused the extrastriate representation. This causes the modal perception of the object or the shape (which is therefore visible). If however the feedback is prevented then the object is amodally perceived (and therefore invisible).

This brings us to the theoretical framework for occlusion.

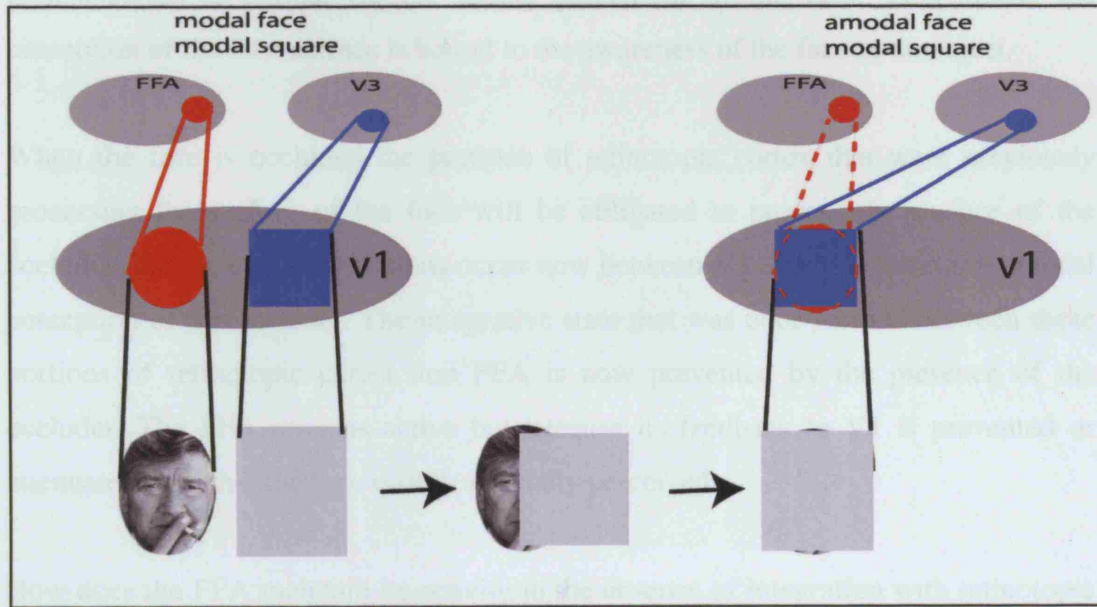


Figure 62 depicts a simple theoretical distinction in neural terms between modal and amodal perception of the face (of David Lynch). The bottom row depicts the outside world as input to a simplified brain. Middle row depicts V1 and upper rows depict object and form selective regions. Solid lines represent recurrent interactions between cortical regions, either directly or indirectly. Dashed lines represent attenuated or ceased feedback. The model on the left depicts state of brain before occlusion and on right after occlusion.

The model depicted above is undoubtedly an over-simplification but hopes to capture the core principles of the theory. For simplicity's sake I assume a two stage model for object processing consisting of luminance based contrast processing in V1 and object selective processing in one of two selective regions for faces and objects. Obviously in reality there could be many more stages analysing intermediate levels of form complexity etc.

When a face is presented (unoccluded) it results in an appropriate topographic representation of luminance-based contrast in a subset of V1 cells. Such cells represent the contrast-based surface of the face without representing it as a 'face' explicitly (Metzinger 2000). These cells project to the FFA (or the Occipital Face Area) which then feedback to the same set of cells that represent the surface of the face. The dynamic interaction between the surface representation of V1 and the object

representation is critical for the modal perception of the face. This means the perception of the face surface is bound to the awareness of the face as an object.

When the face is occluded the portions of retinotopic cortex that were previously processing the surface of the face will be obligated to process the surface of the occluder, the recurrent interactions occur now between V1 and V3 causing the modal perception of the occluder. The integrative state that was once formed between these portions of retinotopic cortex and FFA is now prevented by the presence of the occluder. The FFA remains active but because its feedback to V1 is prevented or attenuated such that the face is only amodally perceived.

How does the FFA maintain its activity in the absence of integration with retinotopic cortices? Given the evidence discussed for experiment 1 one could argue that in the case of occlusion it is the ventral premotor cortex that is integral to this maintenance function. The PPI data shows that when the face is hidden the only two areas to increase in functional connectivity were the FFA and premotor. We propose that this is evidence for multi-stage integration between ventral premotor and the FFA which, by forming a functional unit, acts to maintain the FFA in the active state required to process the presence of a hidden face in the absence of local sensory input.

The difference between a visible face and a hidden face, then, is not due to the level of FFA activity but due to a difference in the areas with which it forms multistage integration (Bartels and Zeki 1999). When a face is visible the FFA integrates with retinotopic cortices (representing surface attributes), and when the face is hidden the FFA integrates with ventral premotor. This model makes clear predictions for the functional connectivity differences supporting modal and amodal perception.

8.2 Theoretical Proposal for Experiment 2:

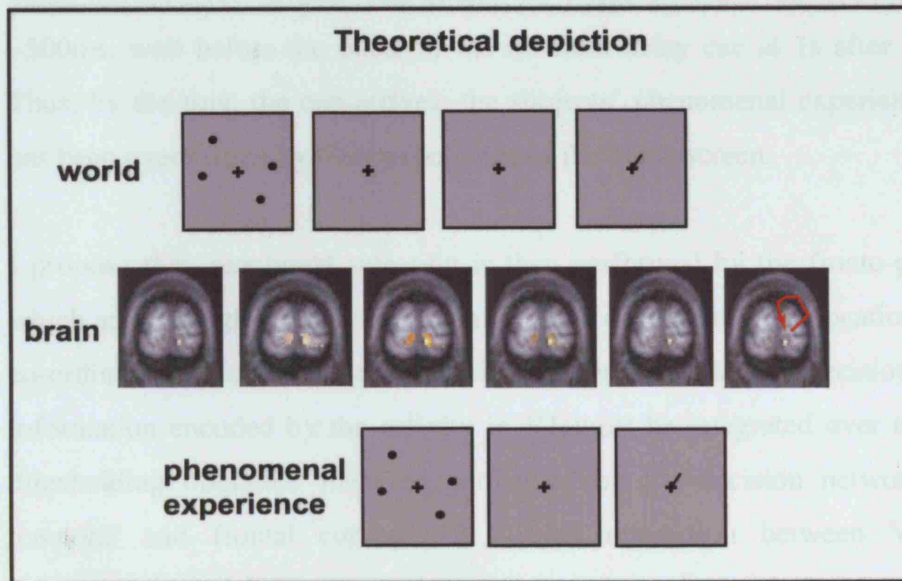


Figure 63 depicts a simplified paradigm consisting of only 4 discs.

Bringing the component stages of stimulus reportability together into one framework I consider a speculative hypothesis for their relation to phenomenal experience and access consciousness, considered in the context of the paradigm used in Experiment 2. Figure 63 depicts the events in a single trial, my hypothesis for which neural areas are active at various points, and how and when these correspond to phenomenal experience.

Note that for this speculative proposal, I make the default assumption that we do have phenomenal experience outside attention in the Blockian sense. If this turns out to be false, then it would be necessary to replace the word 'phenomenal' with 'pre-conscious' in what follows.

I propose that stimulus onset leads to processing of the discs at spatially distinct retinotopic locations in both LGN and V1/V2. After some critical dynamic event, which I shall agnostically refer to the 'essential dynamic', the subject has a phenomenal experience of all discs. The essential dynamic could be due to recurrent activity between the V1/V2 or intra-areal synchrony within V1/V2 alone. However, the claim from the localizationist account is that the correct activity within V1/V2

(however achieved), is minimally sufficient for the basic luminance-based experience of the disc-array. The timescale of this process cannot be ascertained from this data alone, although it would be reasonable to assume that phenomenal experience of the discs has evolved within 100-200ms of stimulus onset and has finished within ~500ms, well before the onset of the medium delay cue at 1s after stimulus onset. Thus, by the time the cue arrives, the subjects' phenomenal experience of the discs has been overwritten by their experience of the blank screen.

I propose that attentional selection is then performed by the fronto-parietal system, which acts through the pulvinar to amplify the cued retinotopic location in V1/V2 and co-ordinate the functional connectivity between V1 and the decision network. The information encoded by the activity in V1 must be integrated over time and then a thresholding operation must be performed by the decision networks in parietal, temporal and frontal cortex. It is this interaction between V1/V2 activity, frontoparietal-pulvinar attentional selection networks, and decision networks that supports access consciousness for this paradigm.

Because of the temporal delay between attentional selection and the stimulus bound activity in V1/V2, there is sufficient late noise between the V1/V2 response and the decision network to result in the erroneous decisions that afford the observed dissociation between stimulus processing and perceptual decision. Because phenomenal experience of the discs is by now over, the attentional system must be accessing something other than immediate phenomenal experience. I propose that the attentional system is accessing the iconic memory trace that the stimulus has left at the V1/V2 locus of the cued location. Due to the late noise the overlap between signal and noise distributions makes it impossible for even a perfect accessing system not to make mistakes. The key claim is that although these late noise effects can affect the perceptual decision they have no influence on the phenomenal experience, since it predated the access event by approximately 1 second. Thus the neural activity (real neural activity not BOLD activity) determining phenomenal experience is temporally divorced from that determining its access.

In essence the hypothesis is as follows. The V1/V2 activity is the minimally sufficient and necessary for phenomenal experience of luminance in retinotopic co-

ordinates, the attentional network co-ordinates the functional connectivity between the decision networks and specific locations within V1/V2 maps, and the mobilisation of the V1/V2 subregions into the decision workspace mediates the access consciousness of the discs. The corresponding mobilisation of motor networks, bringing about motoric reports in line with the binary outcome of the decision network, mediates the final behaviourally measurable report. The trajectory between all four components is the neural basis of stimulus reportability.

9 A general framework for consciousness

Having discussed each experiment individually, I will now discuss how they relate to each other within a common framework – a unified theory of visibility, invisibility and reportability. I will also consider how attempts to delineate such a theory fit into the context of contemporary consciousness research and visual science.

Throughout history, science has waxed and waned over whether consciousness is a suitable topic for scientific research. For most of the twentieth century it was virtually taboo, but in recent years the study of consciousness has become popular. The distaste for examining consciousness in the twentieth century partly arose from the nature of the prevailing scientific paradigm. Behaviourism held that all we should take as our data for understanding the mind is stimulus-response contingencies. The use of introspection by the pioneers of psychology had not led to the answers the Behaviourists sought, and was seen as deeply unscientific. However, with the advent of a more cognitive approach, and the availability of fMRI for relating subjective experience to the brain in a more concrete way, consciousness was again considered a tenable topic for investigation. The pioneering work of scientists like Crick and Koch in the 1990s brought a new found respectability to the study of consciousness. These scientists were concerned to make the definition of consciousness as robust and tangible as possible, and thus it's no surprise that an objective index of consciousness was sought. 'Reportability' was used to give an operational definition of consciousness – the motor act of reporting gives an objectively measurable index of the essentially subjective phenomena we are interested in studying. The quantification of reportability, and consideration of the sensory, cognitive, and motor processes involved, has given a highly successful measure of consciousness. Ingenious stimulus manipulations have been employed, exotic brain damaged patients recruited and technically sophisticated techniques have all been thrown at the problem in what now is a fashionable field of study.

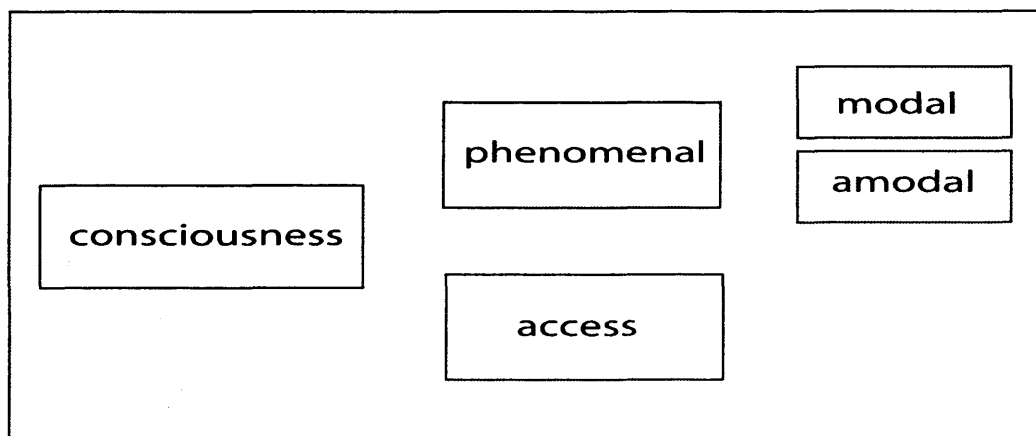
The corollary of this success is that 'consciousness' is again considered by some to be a dirty word in neuroscience research. Its use often leads to confusion, can make good experiments bad and rarely improves an experiment. However, the reason this time round is slightly different – rather than complaining that there is nothing that can be

measured, most objectors now complain that we are not clear about what it is that we are measuring. The field has grown quickly, without any consensus about how its concepts are meant to map onto neural events. I would claim that the experimenter often fails to identify what they mean by ‘consciousness’ and ‘perception’, let alone define these terms in the context of a paper. Indeed, ‘neural correlates of consciousness’ seem to have spread to every possible corner of the brain. This has led to considerable and frustrating contradictions. As discussed earlier, different groups of researchers believe that consciousness is located at the back, the front, or in both ends of the brain. Assuming, for arguments sake, that the experiments themselves are not at fault, there is a real contradiction to be explained. There are therefore strong theoretical and empirical reasons for reconciling these confused definitions and descriptions into a coherent theoretical framework.

I would argue that the contradictions stem from the verificationist approach of operationalising consciousness according to an overly simple notion of report, which is adhered to without consideration of exactly which processes involved in consciousness this measures, and crucially, which processes it does not. Restricting consciousness to those processes which report does directly measure can restrict one to studying only what is in the focus of attention, and thus what is in working memory and available to be reported. This assumes there is only one type of consciousness to be explained – that which allows people to make a motor act to report its contents. I believe that it is this assumption that has led to the current plateau in studies on consciousness.

In challenging this assumption, one is led to consider the possibility of *multiple types* of consciousness, for which there is already some experimental and anecdotal support. The hypothesis that there might be multiple types of consciousness naturally leads to the question of whether a single measure – the notion of reportability – is sufficient to index everything we mean by consciousness. It is in fact deeply insufficient, and this new approach requires a new operationalisation of consciousness to allow the measurement of hypothesised component processes and phenomena. This project has the potential not only to push the field forward in exciting ways, but also to reconcile the apparently contradictory results in the literature.

As its title suggests, this thesis is concerned with 3 aspects of visual consciousness; **visibility**, **invisibility** and **reportability**. In the main body of the thesis, these somewhat colloquial terms were replaced to aid integration with current terminology. In my taxonomy, visibility and invisibility both refer to types of **phenomenal consciousness** – **modal** and **amodal** respectively. The kind of consciousness indexed by traditional report is referred to as **access consciousness**. With this new taxonomy I make the following proposal, which is a hybrid of several previous models of consciousness.



9.1 A general theory of phenomenal consciousness

Phenomenal consciousness is what we might think of as the *qualitative* aspects of consciousness (Block, 1990) - ‘simply experience’. As argued previously, phenomenal consciousness is independent of any perceptual decision and therefore of any actual report, but its contents are reportable if selectively attended. Phenomenal consciousness is fragile, easily overwritten by new inputs and is therefore short-lived. Evidence of its existence is made available through access to iconic memory which contains a trace of past phenomenal states. Thus iconic memory serves to allow phenomenal consciousness to be accessed and reported when attended. Without attention, phenomenal experience is associated only with the unaccessed iconic memory trace which, following phenomenal experience, continually decays over the following seconds. Phenomenal consciousness can thus be operationally defined as being the potential of a subject to voluntarily report a feature of the world if they

attend to that feature *without further motor action or sensory sampling*. Thus the existence of a phenomenal state can be evidenced by anything that can be reported by shifting attention but without further motor action or sensory sampling – i.e. observers are phenomenally conscious of *x* if they could report *x* without needing to shift their attention I refer to this as *1st order reportability*. If subjects were not asked to attend to *x*, but could have reported it had they attended, I refer to this as *2nd order reportability*. Cases in which further sensory sampling is required in order to enable report, for example where a subject could report on *x* if they reoriented their head position, does index constitute phenomenal consciousness, nor access consciousness, and is referred to as *3rd order reportability*, or ‘instrumental’ reportability (Hurley 2002)

	Defined as...	Contingent on	Index of...
1 st order reportability	Potential for report without reorientation of attention	Working memory	Access consciousness
2 nd order reportability	Potential for report which is contingent on reorientation of attention	Iconic memory	Phenomenal consciousness
3 rd order Reportability	Potential for report which is contingent on further sensory sampling	Availability of information in the environment	Availability of information in the environment to be represented in access consciousness

Table 5 showing a taxonomy of reportability and its relation to different types of consciousness.

The content of phenomenal consciousness can be referred to with the traditional philosophical notion of ‘qualia’ – i.e., the dimension of experience such as ‘redness’ or ‘squareness’ that a stimulus imparts. Cortical areas are specialised for processing

certain stimulus dimensions, and we thus assume that each area contributes these dimensions to phenomenal experience. This is not to assert that one area simply 'secretes' a particular type of consciousness (Morgan 2003), and we recognise the many similarities in cell morphology and laminar structure between different cortical areas. However, each specialised area is embedded in a different sensorimotor dynamical structure (Hurley 2002). This dynamical embedding constrains the developmental and functional plasticity across multiple timescales, and sets up unique inputs, outputs, intra-areal dynamics and sensory weights at all spatial scales within an area. The dynamical embedding will have a different characteristic for each cortical area and thus this will instill a characteristic, structural and functional architecture in each area. By acknowledging that this is drastically different for different areas (e.g. V4 and V5) then we can simply assert that somehow this causes (or simply is) the different phenomenal experience. Thus dimensions of experience are caused only by those cortical areas specialised for their processing. I propose that once this difference is set up then the activity of the different areas causes a particular dimension of experience which cannot be caused by any other region (unless it too goes under the necessary changes caused by the sensorimotor dynamical embedding. This is the claim for cortical deference (Hurley 2002)).

In a normally functioning brain, visual phenomenal experience is determined by the collective activity of all visual cortical areas. The exact mechanism by which this 'collective action' occurs is unknown, though I would propose that it is likely to involve a combination of thalamo-cortical and cortico-cortical recursive processing (Llinas et al 1998). For the phenomenal experience of visual dimensions, the necessary thalamocortical interaction must be between the LGN and visual cortex. As discussed, the LGN is not strictly necessary for phenomenal experience to occur. However, under normal circumstances (i.e. in the absence of TMS pulses or electrical stimulation) the LGN may be necessary to drive cortical areas into the state necessary for phenomenal experience and to subsequently update them according to changes in the outside world. In this sense the LGN can be thought of as playing an enabling role for cortical regions to enter into the correct state for phenomenal experience to occur.

Binding between the different dimensions of phenomenal experience encoded by the cortical regions - is then achieved by cortico-cortical feedback. The resulting

phenomenal experience constitutes a 'rough draft' or model of the outside world which is continually updated by the changes in the world which are relayed by the sensory organs, forever re-arranging to reflect the world. The parallel processing that occurs across the visual cortex means that phenomenal experience for multiple items in the scene is possible without attention or interaction from other systems. I propose that this collective but self-contained coalition of organised neural activity is identical to phenomenal consciousness. This is not to say that attention cannot influence phenomenal consciousness, as there is good evidence that attention can change the perceptual appearance of stimuli (Carasco et al 2004). There is also corresponding evidence in the physiological domain that attention can influence the gain and baseline firing rates of visual cortical neurons (Luck et al 1997).

I have now outlined a theory for the neural basis of phenomenal consciousness, but there are many further distinctions to be made *within* phenomenal consciousness itself. I have argued that one important distinction is the difference between modal and amodal experience, and that any complete theory of phenomenal consciousness must accommodate this difference.

Modal phenomenal consciousness

As described previously, modal phenomenal experience involves perceiving something as directly visible, such that its surface has an immediate sensory quality. In the context of experiment 1, I proposed that the modal experience of a face must incorporate both surface representations (in V1) and the object-level representation of the face (in FFA). Through cortico-cortical feedback the two levels of phenomenal experience are bound together such that the face is modally experienced, as one object comprised of different levels of features. Critically, object perception is modal if the object level representations recursively engage with the lowest level representations within V1. This model generalises beyond objects to other feature domains. The position is that for higher order (more abstract) attributes such as form, colour or motion to be modal the following criteria must be achieved.

1. The higher order feature-specific cortical area must reach a phenomenal state by recursive activity with the LGN (or any other means including artificial stimulation)
2. That subsets of V1 neurons which are spatially compatible (i.e. representing approximately the same region of retinotopic space) with the higher order feature representations also reach a phenomenal state.
3. That sustained cortico-cortical recurrent activity between the higher order areas and the subset of V1 neuron be established, such that a metastable network is transiently set-up.

I propose that the resultant binding of the higher-level phenomenal state to the V1 phenomenal state renders the higher order state modal.

Amodal phenomenal consciousness

As described previously amodal phenomenal experience involves perceiving something as being present but not being able to directly perceive the sensory qualities of its surface. In the context of experiment 1, the amodal perception corresponded to awareness of the presence of the occluded face. This was proposed to be achieved by activity within the FFA reaching a phenomenal state, but without being able to fully engage in cortico-cortical recurrent processing with the V1 cells whose receptive fields were subtended by the occluding screen. The resultant experience is then amodal because the face representation cannot be bound to the surface representations in V1. Again this proposal generalises beyond objects to other feature domains. The position is that for higher order (more abstract) attributes such as form, colour or motion to be amodal the following criteria must be achieved.

The higher order feature-specific cortical area must reach a phenomenal state by recursive activity with the LGN (or any other means including artificial stimulation).

That sustained cortico-cortical feedback between the higher order areas and the subregion of V1 with which it is spatially compatible cannot be properly established (this could be due to occlusion, or due to damage to V1 and/or its feedback connections).

The failure of the higher order phenomenal state to bind to the V1 phenomenal state renders the higher order state amodal. Implicit in this proposal is the notion that the activity of all visual areas except V1 are fundamentally amodal unless able to interact appropriately *with* V1. Two testable predictions from this theory are that V5 activity in the absence of feedback to V1 results in an amodal perception of motion, and V4 activity in the absence of feedback to V1 results in an amodal perception of colour.

Now that I have discussed phenomenal consciousness, and its modal and amodal forms, I return to the third concept described at the start of this section – access consciousness.

9.2 A general theory of access consciousness

In my taxonomy, access consciousness refers to experience which has been accessed, and is fully reportable, such that it is poised for voluntary control of at least one motor modality without further allocation of attention. Access consciousness is indexed by what I refer to as *1st order reportability* (See Table 5 above).

I propose that access is mediated by two distinct network operations.

1. Co-ordination of transcortical functional connectivity. In order for a phenomenal state to be accessed and transformed into a reporting action the appropriate functional connectivity must be established. This involves co-ordinating the fronto-parietal networks which support reportability, in order to establish flexible and short-lived functional networks with the correct areas encoding phenomenal states. I propose that this dynamic coordination requires attention, and is critically dependent on thalamocortical interactions involving the pulvinar nucleus of the thalamus. The pulvinar exhibits high fidelity topographic connectivity with both posterior visual cortices and fronto-parietal regions (Shipp 2003). Fronto-parietal regions can access posterior visual cortices through the pulvinar to selectively promote distinct locations within visual cortical maps for access by the fronto-parietal system. This can be mediated by top-down biasing of visual cortices through the pulvinar, which

acts to remotely co-ordinate the simultaneous changes in baseline firing rates and modulating gain. This coordinates the necessary cortico-cortical functional connectivities between posterior visual cortices and frontoparietal networks supporting decisions (parietal cortex and DLPFC) and actions.

2. Once the pulvinar has co-ordinated the necessary functional connectivity between posterior and frontoparietal cortices, informational access can be achieved, transforming the sensory information encoded by phenomenal states into decision variables in parietal and prefrontal cortices. This informational access is thus mediated entirely by cortico-cortical connections in a manner similar to that proposed by Dehaene and Naccache (2001). They describe informational access as 'mobilisation through ignition'. Once mobilised into the 'fronto-parietal workspace' the availability of a phenomenal state for potential report is mediated by the availability of representations for transformation into motor actions by a number of different effector modalities. This availability of phenomenal states for report by different modalities then satisfies the classical operational definition of consciousness i.e. if the subject is asked they can report voluntarily the content of consciousness with any motor modality (1st order reportability)

10. Conclusions

In summary, this thesis constitutes an investigation into the neural correlates of visibility, invisibility and reportability. Firstly, I have shown that the activity of the fusiform face area is invariant whether a face is modally or amodally perceived. The difference between modal and amodal states is instead reflected in the areas with which the FFA forms functional connectivity. Secondly, I have shown that the processes underlying visual stimulus reportability can be dissociated into 4 stages; **stimulus processing**, which correlates with V1; **attentional selection**, which correlates with fronto-parieto-pulvinar networks; **perceptual decision**, which correlates with a network of temporal parietal and frontal regions; and **motor reports**, which correlate with the classical motor system. The results of these two studies have been used as a platform for a theoretical discussion on the neural correlates of the component stages of reportability, and the relation of each stage to consciousness. Ultimately the key issue of whether objects outside of attention are experienced phenomenally is not decidable on the basis of scientific evidence, and it is hard to think what form such an experiment would take. In terms of core scientific relevance, Experiment 2 contributes a novel approach to understanding reportability that reveals previously unscrutinised components of this previously unitary concept. This has been used to motivate a tentative and highly speculative neural theory of phenomenal and access consciousness. This theory is inspired by my experimental results and rests on several previous theories and experimental investigations in the literature, but I acknowledge that it is not yet fully defended by the available data. I argue that phenomenal and access consciousness are dissociable empirically since they are indexed (that is, provide evidence for) by different kinds of reportability, and that they are supported by dissociable but interdependent thalamocortical systems. The model suggests that phenomenal consciousness is supported by a minimal network consisting of the LGN and specialised visual cortices, and that the correct recurrent dynamic between these areas sets up the phenomenal state for which the area is specialised. If cortico-cortical recurrent activity between spatially compatible representations in V1 and higher areas can be achieved then the resulting phenomenal experience is modal. If such recurrency is prevented then the phenomenal experience of the higher-level feature is amodal. Access consciousness, on the other hand, is mediated by a cortico-pulvo-cortical co-ordination of functional connectivity between

posterior visual cortices and fronto-parietal areas. This allows visual representations to be accessed by a transcortical workspace and mobilises it for transformation into decision and action.

THE END

Appendix 1: Feedback for consciousness

The emphasis that Lamme places on feedback to V1 being essential for consciousness cannot be completely true. Patient 'GY' has damage to V1 causing him to be blind in his left visual field. He cannot see slow motion in his blind field and is no better than chance at guessing in a 2 alternative forced choice task. However GY reports seeing fast motion in his blind field and that his experience correlates with activity in V5 (Zeki and ffytche 1998). Thus if V1 is ablated the feedback from V5 to the corresponding part of V1 cannot occur. Therefore feedback to V1 is not essential for visual consciousness of motion, and therefore feedback to V1 cannot be essential for all types of visual consciousness as is maintained by Lamme (Lamme 2001). There are however caveats to this argument. Firstly there could be feedback from V5 to the contralateral V1 thus allowing GY to see motion. This is reasonably unlikely since this would likely cause GY to misplace the motion sensation to the corresponding retinotopic space encoded by his intact V1. A sensible amendment to Lamme's theory would be to abandon feedback to V1 as being essential *per se* and adopt the stance that feedback to any area is necessary for that area to contribute to consciousness. This standpoint would not be refuted by GY since his V5 can receive its own feedback from other areas.

Appendix 2: Calculating d-prime

D-prime can be estimated from taking just two values into account, proportion of hits (hits/(hits+misses)) and proportion of false alarms (false alarms/(false alarms+misses)).

$$1.0 - p(\text{false alarms}) = Z_n$$

The proportion of false alarms when subtracted from 1, converted into a z-score

$$1.0 - p(\text{hits}) = Z_{sn}$$

The proportion of hits when subtracted from 1, converted to a z-score

D-prime is then

$$d' = Z_n - Z_{sn}$$

Appendix 3: Secretion theory

The localizationist framework could be accused of subscribing to the fallacy of secretion theory a position whereby distinct cells secrete different kinds of experience (Morgan 2003). In this sense it could be said that this position described is tantamount to V1 secreting static achromatic luminance experience. The fallacy is assumed because V1 cells are not particularly different in cellular composition to other cells. (V1 is actually a bad example because of its blobs and striations). So how can they simply just cause a different sensation? I would claim that V1 cells do indeed secrete particular dimensions of experience, but only by virtue of being embedded in a rich hierarchical dynamical system whereby its fine-grained causal structure of inputs and outputs is unique, and more importantly the effect that this has on the morphology and plasticity at the sub-synaptic, synaptic, micro-architecture and laminar levels is also unique. By virtue of that uniqueness I would claim causes or embodies a different phenomenal experience. The prediction would be that if the exact same fine-grained dynamical causal structure were instantiated in other region, (perhaps by sensory substitution) then that other region would secrete that same experience. This position is taking the shape of a speculation by Chalmers that there may be some identity between informational states and consciousness, a form of informational panpsychism (Chalmers 1996). In reference to this context it is the fine-grained dynamical causal structure of a region which instantiates a different informational state within a particular brain area when activated, and thereby causes different experience. How it causes an experience at all or why it causes one experience over any other would still be left to be explained, but the point is that there is sufficient physical and functional difference to account for it causing a different experience.

References

- Adams, D., and Horton, J. (2001). What you see ... *Nature* 412, 482-483.
- Aguirre, G. K., and D'Esposito, M. (1999). Topographical disorientation: a synthesis and taxonomy. *Brain* 122 (Pt 9), 1613-1628.
- Aguirre, G. K., Zarahn, E., and D'Esposito, M. (1998). An area within human ventral cortex sensitive to "building" stimuli: evidence and implications. *Neuron* 21, 373-383.
- 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. *The Journal of comparative neurology, J Comp Neurol* 241, 357-381.
- Baillargeon, R. (1986). Representing the existence and the location of hidden objects: object permanence in 6- and 8-month-old infants. *Cognition* 23, 21-41.
- Baillargeon, R., Spelke, E. S., and Wasserman, S. (1985). Object permanence in five-month
- Baleydier, C., and Morel, A. (1992). Segregated thalamocortical pathways to inferior parietal and inferotemporal cortex in macaque monkey. *Visual neuroscience Vis Neurosci* 8, 391-405.-old infants. *Cognition* 20, 191-208.
- Beck, D. M., and Kastner, S. (2005). Stimulus context modulates competition in human extrastriate cortex. *Nature neuroscience, Nat Neurosci* 8, 1110-1116.
- Becker, M. W., Pashler, H., and Anstis S.M (2000). The role of iconic memory in change detection tasks. *Perception* 29, 273-286.
- Beer, J., Blakemore, C., Previc, F. H., and Liotti, M. (2002). Areas of the human brain activated by ambient visual motion, indicating three kinds of self-movement. *Experimentelle Hirnforschung Expérimentation cérébrale Exp Brain Res* 143, 78-88.

Behrmann, M., Zemel, R. S., and Mozer, M. C. (1998). Object-based attention and occlusion: evidence from normal participants and a computational model. *J Exp Psychol Hum Percept Perform* 24, 1011-1036.

Bender, D. B. (1981). Retinotopic organization of macaque pulvinar. *Journal of neurophysiology, J Neurophysiol* 46, 672-693.

Bender, D. B., and Baizer, J. S. (1984). Anterograde degeneration in the superior colliculus following kainic acid and radiofrequency lesions of the macaque pulvinar. *The Journal of comparative neurology, J Comp Neurol* 228, 284-298.

Bender, D. B., and Youakim, M. (2001). Effect of attentive fixation in macaque thalamus and cortex. *Journal of neurophysiology, J Neurophysiol* 85, 219-234.

Block, N. (1990). Availability: The cognitive basis of experience.

Block, N. (1990). Consciousness and accessibility. *Behavioral and Brain Sciences* 13, 596-598.

Block, N. (1995). Author's Response: On a confusion about a function of consciousness. *Brain and Behavioural Sciences* 18, 227-287.

Block, N. (2001). Paradox and cross purposes in recent work on consciousness. *Cognition* 79, 197-219.

Blythe, I. M., Kennard, C., and Ruddock, K. H. (1987). Residual vision in patients with retrogeniculate lesions of the visual pathways. *Brain : a journal of neurology* Brain 110 (Pt 4), 887-905.

Brindley, G. S., and Lewin, W. S. (1968). The sensations produced by electrical stimulation of the visual cortex. *The Journal of physiology, J Physiol* 196, 479-493.

Bruce, C., Desimone, R., and Gross, C. G. (1981). Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *Journal of neurophysiology, J Neurophysiol* 46, 369-384.

Carrasco, M., Ling, S., and Read, S. (2004). Attention alters appearance. *Nature neuroscience, Nat Neurosci* 7, 308-313.

Chalupa, L. M., Coyle, R. S., and Lindsley, D. B. (1976). Effect of pulvinar lesions on visual pattern discrimination in monkeys. *Journal of neurophysiology, J Neurophysiol* 39, 354-369.

Chalupa, L. M. (1977). A review of cat and monkey studies implicating the pulvinar in visual function. *Behavioral biology, Behav Biol* 20, 149-167.

Chow KL. (1954). Lack of behavioral effects following destruction of some thalamic association nuclei in monkey. *A M A archives of neurology and psychiatry AMA Arch Neurol Psychiatry* 71, 762-771.

Chalmers, D. (1996). *The Conscious Mind: In Search of a Fundamental Theory* (Oxford, Oxford University Press).

Chalmers, D. (1998). On the Search for the Neural Correlate of Consciousness. In *Toward a Science of Consciousness II: The Second Tucson Discussions and Debates*, A. K. S. Hameroff, and A.Scott, ed. (MIT Press).

Chalmers, D. (2004). How Can We Construct a Science of Consciousness? In *The Cognitive Neurosciences III*, M. S. Gazzaniga, ed. (Boston, MIT Press).

Chelazzi, L., Miller, E. K., Duncan, J., and Desimone, R. (1993). A neural basis for visual search in inferior temporal cortex. *Nature* 363, 345-347.

Churchland, M. M., Chou, I.-H., and Lisberger, S. G. (2003). Evidence for object permanence in the smooth-pursuit eye movements of monkeys. *J Neurophysiol* 90, 2205-2218.

Coltheart, M. (1980). The persistences of vision. *Philosophical transactions of the Royal Society of London Series B, Biological sciences*, Philos Trans R Soc Lond B Biol Sci 290, 57-69.

Coltheart, M. (1980). Iconic memory and visible persistence. *Perception and psychophysics*, Percept Psychophys 27, 183-228.

Corbetta, M., Miezin, F. M., Shulman, G. L., and Petersen, S. E. (1991). Selective attention modulates extrastriate visual regions in humans during visual feature discrimination and recognition. *Ciba Foundation symposium*, Ciba Found Symp 163, 165-175.

Critchley, M. (1953). Tactile thought, with special reference to the blind. *Brain : a journal of neurology*, Brain 76, 19-35.

Cumming, B. G., and Parker, A. J. (1997). Responses of primary visual cortical neurons to binocular disparity without depth perception. *Nature* 389, 280-283.

Danziger, S., Ward, R., Owen, V., and Rafal, R. (2001). The effects of unilateral pulvinar damage in humans on reflexive orienting and filtering of irrelevant information. *Behavioural neurology*, Behav Neurol 13, 95-104.

Deco, G., and Rolls, E. T. (2005). Neurodynamics of biased competition and cooperation for attention: a model with spiking neurons. *Journal of neurophysiology* J Neurophysiol 94, 295-313.

Deco, G., Pollatos, O., and Zihl, J. (2002). The time course of selective visual attention: theory and experiments. *Vision research*, Vision Res 42, 2925-2945.

Dehaene, S. (2004). neural mechanisms for access to consciousness.

Dehaene, S., Naccache, L., Cohen, L., Bihan, D. L., Mangin, J.-F., Poline, J.-B., and

Riviere, D. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nat Neurosci* 4, 752-758.

Dennett, D. (1995). The Path Not Taken- Commentary on Block. *Behavioral and Brain Sciences* 183, 252-255.

Desimone, R., Gross CG (1979). Visual areas in the temporal cortex of the macaque. *Brain Research* 178.

Desimone, R. (1996). Neural mechanisms for visual memory and their role in attention. *Proc Natl Acad Sci U S A* 93, 13494-13499.

Desimone, R. (1998). Visual attention mediated by biased competition in extrastriate visual cortex. *Philosophical transactions of the Royal Society of London Series B, Biological sciences, Philos Trans R Soc Lond B Biol Sci* 353, 1245-1255.

Desimone, R., Albright, T. D., Gross, C. G., and Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *The Journal of neuroscience : the official journal of the Society for Neuroscience, J Neurosci* 4, 2051-2062.

Dobelle, W. H., and Mladejovsky, M. G. (1974). Phosphenes produced by electrical stimulation of human occipital cortex, and their application to the development of a prosthesis for the blind. *The Journal of physiology, J Physiol* 243, 553-576.

Driver, J., and Frith, C. (2000). Shifting baselines in attention research. *Nature reviews Neuroscience*
Nat Rev Neurosci 1, 147-148.

Druzgal, T. J., and D'Esposito, M. (2001). Activity in fusiform face area modulated as a function of working memory load. *Brain Res Cogn Brain Res* 10, 355-364.

Druzgal, T. J., and D'Esposito, M. (2003). Dissecting contributions of prefrontal cortex and fusiform face area to face working memory. *J Cogn Neurosci* 15, 771-784.

Dumbrava, D., Faubert, J., and Casanova, C. (2001). Global motion integration in the cat's lateral posterior-pulvinar complex. *The European journal of neuroscience* Eur J Neurosci 13, 2218-2226.

Edelman, G., Tononi, Giulio (2001). *A Universe of Consciousness: How Matter Becomes Imagination*, Basic Books).

Edelman, G. (2001). Consciousness: the remembered present. *Annals of the New York Academy of Sciences*. Ann N Y Acad Sci 929, 111-122.

Epstein, R., Harris, A., Stanley, D., and Kanwisher, N. (1999). The parahippocampal place area: recognition, navigation, or encoding? *Neuron* 23, 115-125.

Fan, J., McCandliss, B. D., Fossella, J., Flombaum, J. I., and Posner, M. I. (2005). The activation of attentional networks. *Neuroimage* 26, 471-479.

ffytche, D. H., and Pins, D. (2003). Are neural correlates of visual consciousness retinotopic? *Neuroreport* 14, 2011-2014.

Field, T. M., Woodson, R., Greenberg, R., and Cohen, D. (1982). Discrimination and imitation of facial expression by neonates. *Science* 218, 179-181.

Frackowiak, R. S. J., Friston, K. J., Frith, C. D., Dolan, R. J., Price, C. J., Zeki, S., Ashburner, J. T., and Penny, W. D. (2004). *Human Brain Function*.

Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., and Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage* 6, 218-229.

Friston, K. J., Frith, C. D., Liddle, P. F., Dolan, R. J., Lammertsma, A. A., and Frackowiak, R. S. (1990). The relationship between global and local changes in PET scans. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism* J Cereb Blood Flow Metab 10, 458-466.

Friston, K. J., Frith, C. D., Liddle, P. F., and Frackowiak, R. S. (1991). Comparing functional (PET) images: the assessment of significant change. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism*, *J Cereb Blood Flow Metab* 11, 690-699.

Friston, K. J., Holmes, A. P., and Worsley, K. J. (1999). How many subjects constitute a study? *Neuroimage* 10, 1-5.

Friston, K. J., Penny, W. D., and Glaser, D. E. (2005). Conjunction revisited. *Neuroimage* 25, 661-667.

Fuster, J. M. (2001). The prefrontal cortex--an update: time is of the essence. *Neuron* 30, 319-333.

Fuster, J. M., and Jervey, J. P. (1982). Neuronal firing in the inferotemporal cortex of the monkey in a visual memory task. *J Neurosci* 2, 361-375.

Gauthier, I., Skudlarski, P., Gore, J. C., and Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature neuroscience* *Nat Neurosci* 3, 191-197.

Gauthier, I., and Tarr, M. J. (1997). Becoming a "Greeble" expert: exploring mechanisms for face recognition. *Vision research*, *Vision Res* 37, 1673-1682.

Gazzaniga, M. S., LeDoux, J. E., and Wilson, D. H. (1977). Language, praxis, and the right hemisphere: clues to some mechanisms of consciousness. *Neurology* 27, 1144-1147.

Gold, J., and Shadlen, M. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends Cogn Sci* 5, 10-16.

Goodale, M. A., Milner, A. D., Jakobson, L. S., and Carey, D. P. (1991). Object awareness. *Nature* 352, 202.

Goren, C. C., Sarty, M., and Wu, P. Y. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics* 56, 544-549.

Graziano, M. S., Hu, X. T., and Gross, C. G. (1997). Visuospatial properties of ventral premotor cortex. *J Neurophysiol* 77, 2268-2292.

Green, D. M., and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*, Wiley).

Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzhak, Y., and Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron* 24, 187-203.

Han, S., Wan, X., and Humphreys, G. W. (2005). Shifts of spatial attention in perceived 3-D space. *Q J Exp Psychol A* 58, 753-764.

Hannula, D., Simons, D., and Cohen, N. (2005). Imaging implicit perception: promise and pitfalls. *Nat Rev Neurosci* 6, 247-255.

Hasselmo, M. E., Rolls, E. T., Baylis, G. C., and Nalwa, V. (1989). Object-centered encoding by face-selective neurons in the cortex in the superior temporal sulcus of the monkey. *Experimental brain research Experimentelle Hirnforschung, Exp Brain Res* 75, 417-429.

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 Sciences of the United States of America Proc Natl Acad Sci U S A* 88, 1621-1625.

Haxby, J. V., Horwitz, B., Ungerleider, L. G., Maisog, J. M., Pietrini, P., and Grady, C. L. (1994). The functional organization of human extrastriate cortex: a PET-rCBF

study of selective attention to faces and locations. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, J Neurosci 14, 6336-6353.

Haxby, J. V., Ungerleider, L. G., Clark, V. P., Schouten, J. L., Hoffman, E. A., and Martin, A. (1999). The effect of face inversion on activity in human neural systems for face and object perception. *Neuron* 22, 189-199.

Haynes, J.-D., Driver, J., and Rees, G. (2005). Visibility reflects dynamic changes of effective connectivity between V1 and fusiform cortex. *Neuron* 46, 811-821.

Haynes, J.-D., and Rees, G. (2005). Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nature neuroscience, Nat Neurosci* 8, 686-691.

He S, Smallman HS, MacLeod DIA (1995) Neural and cortical limits on visual resolution. *Invest Ophthalmol Vis Sci (Suppl)* 36:S438.

He S, Cavanagh P, Intriligator J (1996) Attentional resolution and the locus of visual awareness. *Nature* 383:334–337.

He, S., and MacLeod, D. I. (2001). Orientation-selective adaptation and tilt after-effect from invisible patterns. *Nature* 411, 473-476.

Heekeren, H. R., Marrett, S., Bandettini, P. A., and Ungerleider, L. G. (2004). A general mechanism for perceptual decision-making in the human brain. *Nature* 431, 859-862.

Helmholtz, H. L. F. von (1867). *Handbuch der Physiologischen Optik*. Leipzig.

Horton JC, Hoyt WF (1991) Quadrantic visual field defects. A hallmark of lesions in the extra-striate (V2/V3) cortex. *Brain* 114:1703–1718.

HUBEL, D. H., and WIESEL, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *The Journal of physiology, J Physiol* 148, 574-591.

Hurley, S. (2002). *Consciousness in Action* (Boston, Harvard University Press).

Inouye, T. (2000). Visual Disturbances following Gunshot Wounds of the Cortical Visual Area (translated by M. Glickestein and M. Fahle) (*Brain* 123, Suppl. 1-101). In (Oxford Univ. Press).

Ishai, A., Haxby, J. V., and Ungerleider, L. G. (2002). Visual imagery of famous faces: effects of memory and attention revealed by fMRI. *Neuroimage* 17, 1729-1741.

Itti, L., Rees, G., Tsotsos, JK, ed. (2003). *Neurobiology of Attention* (Academic Press Inc.,U.S.).

Johnson, S. P. (2004). Development of perceptual completion in infancy. *Psychol Sci* 15, 769-775.

Joseph, J. S., and Nakayama, K. (1999). Amodal representation depends on the object seen before partial occlusion. *Vision Res* 39, 283-292.

Kanwisher, N. (2000). Domain specificity in face perception. *Nature neuroscience* Nat Neurosci 3, 759-763.

Kanwisher, N. (2001). Neural events and perceptual awareness. *Cognition* 79, 89-113.

Kanwisher, N., Chun, M. M., McDermott, J., and Ledden, P. J. (1996). Functional imaging of human visual recognition. *Brain research Cognitive brain research* Brain Res Cogn Brain Res 5, 55-67.

Kanwisher, N., McDermott, J., and Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *The Journal of neuroscience : the official journal of the Society for Neuroscience* J Neurosci 17, 4302-4311.

Kastner, S., and Ungerleider, L. G. (2001). The neural basis of biased competition in human visual cortex. *Neuropsychologia* 39, 1263-1276.

Kastner, S., O'Connor, D. H., Fukui, M. M., Fehd, H. M., Herwig, U., and Pinsk, M. A. (2004). Functional imaging of the human lateral geniculate nucleus and pulvinar. *J Neurophysiol* 91, 438-448.

Kellman, P. J., Yin, C., and Shipley, T. F. (1998). A common mechanism for illusory and occluded object completion. *J Exp Psychol Hum Percept Perform* 24, 859-869.

Kleinschmidt, A., Bülthel, C., Zeki, S., and Frackowiak, R. S. (1998). Human brain activity during spontaneously reversing perception of ambiguous figures. *Proceedings Biological sciences / The Royal Society Proc Biol Sci* 265, 2427-2433.

Kohler S, C. J., Milner B (2002). Differential contributions of the parahippocampal place area and the anterior hippocampus to human memory for scenes. *Hippocampus* 12, 718-723.

Kurata, K., and Tanji, J. (1986). Premotor cortex neurons in macaques: activity before distal and proximal forelimb movements. *J Neurosci* 6, 403-411.

Lamme, V. (2003). Why visual attention and awareness are different. *Trends Cogn Sci* 7, 12-18.

Lamme, V. A. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *J Neurosci* 15, 1605-1615.

Lamme, V. A. (2001). Blindsight: the role of feedforward and feedback corticocortical connections. *Acta Psychol (Amst)* 107, 209-228.

Lamme, V. A., and Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in neurosciences* *Trends Neurosci* 23, 571-579.

Landis, T., Cummings, J. L., Benson, D. F., and Palmer, E. P. (1986). Loss of topographic familiarity. An environmental agnosia. *Arch Neurol* 43, 132-136.

Landis, T., Regard, M., Bliestle, A., and Kleihues, P. (1988). Prosopagnosia and agnosia for noncanonical views. An autopsied case. *Brain : a journal of neurology* Brain 111 (Pt 6), 1287-1297.

Landman, R., Spekreijse, H., and Lamme, V. A. F. (2003). Large capacity storage of integrated objects before change blindness. *Vision research, Vision Res* 43, 149-164.

Landman, R., Spekreijse, H., and Lamme, V. A. F. (2004). The role of figure-ground segregation in change blindness. *Psychonomic bulletin and review, Psychon Bull Rev* 11, 254-261.

Leopold, D. A., and Logothetis, N. K. (1996). Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry, *Nature* 379, 549-553.

Levy, R., and Goldman-Rakic, P. S. (2000). Segregation of working memory functions within the dorsolateral prefrontal cortex. *Experimental brain research Experimentelle Hirnforschung, Exp Brain Res* 133, 23-32.

Llinás, R., Ribary, U., Contreras, D., and Pedroarena, C. (1998). The neuronal basis for consciousness. *Philosophical transactions of the Royal Society of London Series B, Biological sciences, Philos Trans R Soc Lond B Biol Sci* 353, 1841-1849.

Logothetis, N. K., Leopold, D. A., and Sheinberg, D. L. (1996). What is rivalling during binocular rivalry? *Nature* 380, 621-624.

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., and Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412, 150-157.

Logothetis, N. K., and Schall, J. D. (1989). Neuronal correlates of subjective visual perception. *Science* 245, 761-763.

Lu, Z.-L., Neuse, J., Madigan, S., and Doshier, B. A. (2005). Fast decay of iconic memory in observers with mild cognitive impairments. *Proc Natl Acad Sci U S A* 102, 1797-1802.

Luck, S. J., Chelazzi, L., Hillyard, S. A., and Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *Journal of neurophysiology*
J Neurophysiol 77, 24-42.

Lumer, E. D., Friston, K. J., and Rees, G. (1998). Neural correlates of perceptual rivalry in the human brain. *Science*
Science 280, 1930-1934.

Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., Ledden, P. J., Brady, T. J., Rosen, B. R., and Tootell, R. B. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences of the United States of America*
Proc Natl Acad Sci U S A 92, 8135-8139.

Mareschal, D. (2000). Object knowledge in infancy: current controversies and approaches. *Trends Cogn Sci* 4, 408-416.

McAdams, C. J., and Reid, R. C. (2005). Attention modulates the responses of simple cells in monkey primary visual cortex. *J Neurosci* 25, 11023-11033.

McCarthy, G., Puce, A., Belger, A., and Allison, T. (1991). Electrophysiological studies of human face perception. II: Response properties of face-specific potentials generated in occipitotemporal cortex. *Cerebral Cortex* 9, 431-444.

Meng, M., Remus, D. A., and Tong, F. (2005). Filling-in of visual phantoms in the human brain. *Nat Neurosci* 8, 1248-1254.

Merabet, L., Desautels, A., Minville, K., and Casanova, C. (1998). Motion integration in a thalamic visual nucleus. *Nature* 396, 265-268.

Metzinger, T. (2000). *Neural Correlates of Consciousness: Empirical and Conceptual Questions*, Academic Press Inc., U.S.).

Michael, G. A., and Buron, V. r. (2005). The human pulvinar and stimulus-driven attentional control. *Behavioral neuroscience, Behav Neurosci* 119, 1353-1367.

Morel, A., Magnin, M., and Jeanmonod, D. (1997). Multiarchitectonic and stereotactic atlas of the human thalamus. *The Journal of comparative neurology J Comp Neurol* 387, 588-630.

Michotte, A. (1950). On phenomenal permanence: facts and theories. *Acta Psychologica*, 7, 293-322

Milner, D., Goodale, MA (1995). *The Visual Brain in Action.*, Vol 17, Oxford: Oxford University Press.).

Miyashita, Y., Date, A., and Okuno, H. (1993). Configurational encoding of complex visual forms by single neurons of monkey temporal cortex. *Neuropsychologia* 31, 1119-1131.

Moutoussis, K., and Zeki, S. (2002). The relationship between cortical activation and perception investigated with invisible stimuli. *Proc Natl Acad Sci U S A* 99, 9527-9532.

Morgan, M.J. (2003) *The Space Between Our Ears'* (winner of the Wellcome Turst Book Prize). London: Weidenfled and Nicholson

Murata, A., Fadiga, L., Fogassi, L., Gallese, V., Raos, V., and Rizzolatti, G. (1997). Object representation in the ventral premotor cortex (area F5) of the monkey. *J Neurophysiol* 78, 2226-2230.

Nakamura, K., and Kubota, K. (1995). Mnemonic firing of neurons in the monkey temporal pole during a visual recognition memory task. *J Neurophysiol* 74, 162-178.

Newby, E. A., and Rock, I. (1998). Inattention blindness as a function of proximity to the focus of attention. *Perception* 27, 1025-1040.

Nichols, T., Brett, M., Andersson, J., Wager, T., and Poline, J.-B. (2005). Valid conjunction inference with the minimum statistic. *Neuroimage* 25, 653-660.

Niemann, K., Mennicken, V. R., Jeanmonod, D., and Morel, A. (2000). The Morel stereotactic atlas of the human thalamus: atlas-to-MR registration of internally consistent canonical model. *NeuroImage* 12, 601-616.

O Scalaidhe, S. P., Wilson, F. A., and Goldman-Rakic, P. S. (1997). Areal segregation of face-processing neurons in prefrontal cortex. *Science* 278, 1135-1138.

O'Connor, D., Fukui, M., Pinsk, M., and Kastner, S. (2002). Attention modulates responses in the human lateral geniculate nucleus. *Nat Neurosci* 5, 1203-1209.

O'Craven, K. M., and Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *J Cogn Neurosci* 12, 1013-1023.

O'Hear, A., ed. (1998). *Contemporary Issues in the Philosophy of Mind* (Cambridge University Press).

O'Regan, J. K., and Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *The Behavioral and brain sciences, Behav Brain Sci* 24, 939-973.

Palmer, S. E. (1999). *Vision science: Photons to phenomenology*. Cambridge, MA: MIT Press

Palmer, J., Huk, A. C., and Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *Journal of vision, J Vis* 5, 376-404.

Pessoa, L., and Ungerleider, L. G. (2004). Neural correlates of change detection and change blindness in a working memory task. *Cerebral cortex (New York, NY : 1991)* Cereb Cortex 14, 511-520.

Peters, R. J., Iyer, A., Itti, L., and Koch, C. (2005). Components of bottom-up gaze allocation in natural images. *Vision Res* 45, 2397-2416.

Petersen, S. E., Robinson, D. L., and Keys, W. (1985). Pulvinar nuclei of the behaving rhesus monkey: visual responses and their modulation. *Journal of neurophysiology. J Neurophysiol* 54, 867-886.

Petersen, S. E., Robinson, D. L., and Morris, J. D. (1987). Contributions of the pulvinar to visual spatial attention. *Neuropsychologia*, 25, 97-105.

Pollen, D. A. (1999). On the neural correlates of visual perception. *Cerebral cortex* 9, 4-19.

Rauschenberger, R., and Yantis, S. (2001). Masking unveils pre-amodal completion representation in visual search. *Nature* 410, 369-372.

Raz, A., and Buhle, J. (2006). Typologies of attentional networks. *Nat Rev Neurosci* 7, 367-379.

Rensink, R. A. (2000). Seeing, sensing, and scrutinizing. *Vision research* Vision Res 40, 1469-1487.

Ress, D., Backus, B. T., and Heeger, D. J. (2000). Activity in primary visual cortex predicts performance in a visual detection task. *Nat Neurosci* 3, 940-945.

Ress, D., and Heeger, D. J. (2003). Neuronal correlates of perception in early visual cortex. *Nat Neurosci* 6, 414-420.

Robinson, D. L., and Petersen, S. E. (1985). Responses of pulvinar neurons to real and self-induced stimulus movement. *Brain research, Brain Res* 338, 392-394.

Robinson, D. L., Petersen, S. E., and Keys, W. (1986). Saccade-related and visual activities in the pulvinar nuclei of the behaving rhesus monkey. *Experimental brain research Experimentelle Hirnforschung Expérimentation cérébrale, Exp Brain Res* 62, 625-634.

Robinson, D. L., McClurkin, J. W., Kertzman, C., and Petersen, S. E. (1991). Visual responses of pulvinar and collicular neurons during eye movements of awake, trained macaques. *Journal of neurophysiology, J Neurophysiol* 66, 485-496.

Robinson, D. L., and Petersen, S. E. (1992). The pulvinar and visual salience. *Trends in neurosciences, Trends Neurosci* 15, 127-132.

Robinson, D. L. (1993). Functional contributions of the primate pulvinar. *Progress in brain research. Prog Brain Res* 95, 371-380.

Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G., and Matelli, M. (1988). Functional organization of inferior area 6 in the macaque monkey. II. Area F5 and the control of distal movements. *Exp Brain Res* 71, 491-507.

Rizzolatti, G., Scandolara, C., Matelli, M., and Gentilucci, M. (1981). Afferent properties of periarculate neurons in macaque monkeys. II. Visual responses. *Behav Brain Res* 2, 147-163.

Rolls, E. T. (1984). Neurons in the cortex of the temporal lobe and in the amygdala of the monkey with responses selective for faces. *Hum Neurobiol* 3, 209-222.

Schall, J. D. (2001). Neural basis of deciding, choosing and acting. *Nat Rev Neurosci* 2, 33-42.

Scholl, B. J., and Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: clues to visual objecthood. *Cognit Psychol* 38, 259-290.

Schneider, K. A., Richter, M. C., and Kastner, S. (2004). Retinotopic organization and functional subdivisions of the human lateral geniculate nucleus: a high-resolution functional magnetic resonance imaging study. *J Neurosci* 24, 8975-8985.

Sekuler A, Palmer S, Flynn C (1994) Local and global processes in visual completion. *Psychol Sci* 5:260–267

Sergent, C., Baillet, S., and Dehaene, S. (2005). Timing of the brain events underlying access to consciousness during the attentional blink. *Nat Neurosci* 8, 1391-1400.

Sergent, C., and Dehaene, S. (2004). Is consciousness a gradual phenomenon? Evidence for an all-or-none bifurcation during the attentional blink. *Psychological science : a journal of the American Psychological Society / APS Psychol Sci* 15, 720-728.

Sergent, J., Signoret, J L (1991). Varieties of functional deficits in prosopagnosia. *Cerebral Cortex* 2, 375-388.

Shallice, T. (2003). Functional imaging and neuropsychology findings: how can they be linked? *NeuroImage* 20, 146-154.

Sheinberg, D. L., and Logothetis, N. K. (1997). The role of temporal cortical areas in perceptual organization. *Proc Natl Acad Sci U S A* 94, 3408-3413.

Sherman, S. M., and Guillery, R. W. (1998). On the actions that one nerve cell can have on another: distinguishing "drivers" from "modulators". *Proc Natl Acad Sci U S A* 95, 7121-7126.

Sherman SM, Guillery RW (2004) The visual relays in the thalamus. In: *The Visual Neurosciences* (Chalupa LM, Werner JS, eds), pp 565-591. MIT Press.

Shipp, S. (2003). The functional logic of cortico-pulvinar connections. *Philos Trans R Soc Lond B Biol Sci* 358, 1605-1624.

Shipp, S. (2004). The brain circuitry of attention. *Trends Cogn Sci* 8, 223-230.

Ungerleider, L. G., and Christensen, C. A. (1977). Pulvinar lesions in monkeys produce abnormal eye movements during visual discrimination training. *Brain Res* 136, 189-196.

Silver, M. A., Ress, D., and Heeger, D. J. (2005). Topographic maps of visual spatial attention in human parietal cortex. *J Neurophysiol* 94, 1358-1371.

Simons, D. J., Chabris, C. F., Schnur, T., and Levin, D. T. (2002). Evidence for preserved representations in change blindness. *Conscious Cogn* 11, 78-97.

Simons, D. J., Franconeri, S. L., and Reimer, R. L. (2000). Change blindness in the absence of a visual disruption. *Perception* 29, 1143-1154.

Simons, D. J., and Rensink, R. A. (2005). Change blindness: past, present, and future. *Trends in cognitive sciences*
Trends Cogn Sci 9, 16-20.

Stoerig, P., and Barth, E. (2001). Low-level phenomenal vision despite unilateral destruction of primary visual cortex. *Conscious Cogn* 10, 574-587.

Tanaka, J. W., and Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly journal of experimental psychology A, Human experimental psychology*
Q J Exp Psychol A 46, 225-245.

Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annual review of neuroscience* 19, 109-139.

Tanner, J. W. P., and Swets, J. A. (1954). A decision-making theory of visual detection. *Psychol Rev* 61, 401-409.

Thompson, J., Clarke, M., Stewart, T., and Puce, A. (2005). Configural Processing of Biological Motion in Human Superior Temporal Sulcus. *J Neurosci* 25, 9059-9066.

Thurstone, L. L. (1927). Psychophysical Analysis. *The American journal of psychology* 100, 587-609.

Tong, F. (2003). PRIMARY VISUAL CORTEX AND VISUAL AWARENESS. *Nat Rev Neurosci* 4, 219-229.

Tong, F., Nakayama, K., Vaughan, J. T., and Kanwisher, N. (1998). Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron* 21, 753-759.

Tse, P. U., Martinez-Conde, S., Schlegel, A. A., and Macknik, S. L. (2005). Visibility, visual awareness, and visual masking of simple unattended targets are confined to areas in the occipital cortex beyond human V1/V2. *Proc Natl Acad Sci U S A* 102, 17178-17183.

Umiltà, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., and Rizzolatti, G. (2001). I know what you are doing. a neurophysiological study. *Neuron* 31, 155-165.

Ungerleider, L. G., and Christensen, C. A. (1979). Pulvinar lesions in monkeys produce abnormal scanning of a complex visual array. *Neuropsychologia* 17, 493-501.

Villeneuve, M. Y., Kupers, R., Gjedde, A., Ptito, M., and Casanova, C. (2005). Pattern-motion selectivity in the human pulvinar. *NeuroImage* 28, 474-480.

Weiskrantz, L. (1997). *Consciousness Lost and Found*. (New York, Oxford, Oxford University Press,).

Weiskrantz, L., Barbur, J. L., and Sahraie, A. (1995). Parameters affecting conscious versus unconscious visual discrimination with damage to the visual cortex (V1). *Proc Natl Acad Sci U S A* 92, 6122-6126.

Yi, D.-J., Kim, M.-S., and Chun, M. M. (2003). Inhibition of return to occluded objects. *Percept Psychophys* 65, 1222-1230.

Zeki, S. (1993). *Vision of the Brain*.

Zeki, S., and Bartels, A. (1999). Toward a theory of visual consciousness. *Conscious Cogn* 8, 225-259.

Zeki, S., and Ffytche, D. H. (1998). The Riddoch syndrome: insights into the neurobiology of conscious vision. *Brain* 121 (Pt 1), 25-45.

Zeki, S., and Shipp, S. (1988). The functional logic of cortical connections. *Nature* 335, 311-317.

Zeki, S. M. (1978). Functional specialisation in the visual cortex of the rhesus monkey. *Nature* 274, 423-428.

Zihl, J., and von Cramon, D. (1979). The contribution of the 'second' visual system to directed visual attention in man. *Brain* : 102, 835-856.

Zipser, K., Lamme, V. A., and Schiller, P. H. (1996). Contextual modulation in primary visual cortex. *J Neurosci* 16, 7376-7389.