Detection without awareness: an investigation of implicit change detection using combined EEG and fMRI

Catriona L. Scrivener

PhD in Neuroscience

A thesis presented for the degree of

Doctor of Philosophy

School of Psychology and Clinical Language Sciences

University of Reading

July 2019

Abstract

The failure to detect differences between visual scenes is known as change blindness. When changes to an image are disrupted in some way, for example by a distractor screen or eye movement, we are often blind to any differences. It was once assumed that change detection is dichotomous; we either see a change, or we don't. However, the presence of a change can influence our behaviour, even in the absence of full conscious report. It may be possible for us to *sense* that a change has occurred, even if we cannot specify exactly where or what it was. Using electroencephalogram (EEG), functional magnetic resonance imaging (fMRI), and behavioural measures, we found multi-modal evidence to suggest that *sensing* a change is distinguishable from being *blind* to it. In EEG, the late positivity potential, N2pc, and N1 amplitudes were larger for *sense* trials compared to *blind*. Additionally, a range of visual (BA18), parietal (BA40), and midbrain (anterior cingulate) areas showed increased fMRI BOLD activation when a change was *sensed*. These visual and parietal areas are commonly implicated as the storage sites of visual working memory, and we therefore argue that *sensing* may not be explained by a lack of representation of the visual display. In addition, we compared the EEG recorded inside and outside of the MRI scanner to investigate the influence of the MRI environment. Increased amplitudes were identified in the visual N1 in combined EEG-fMRI data, and almost all peak latencies were reduced. Based on our experience with EEG-fMRI data, we provide a guide for researchers considering combined recording and suggest when simultaneous EEG-fMRI may or may not be necessary.

For my nan, Margaret Adam

Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Chapter 3 contains the manuscript for a paper that has been accepted at Experimental Brain Research. The experiment design, analysis, and manuscript write-up were 95% my own work, with comments provided by the co-authors at weekly lab meetings. Jade Marsh helped with data collection for all participants. All other co-authors provided comments on at least one manuscript draft. These were Asad Malik, Jade Marsh, Michael Lindner, and Etienne Roesch.

All other chapters were 95% my own work, with comments provided at weekly lab meetings. Michael Lindner, Asad Malik, and Zola Dean contributed with advice for the fMRI and EEG-informed fMRI analysis. Additional comments on the write-up of this thesis were provided by Arran Reader and David Scrivener.

Catriona L. Scrivener

Acknowledgements

Thank you to my MSc supervisor Joseph Brooks, whose support and belief gave me the confidence to pursue a career in academia, and without whom I probably would not have attempted a PhD.

Thank you to my parents for their constant support, and to my dad, David Scrivener, who has proof-read everything that I have written from BSc to PhD.

Thank you to my lab members for all of their help and stimulating conversation. In particular Asad Malik, who helped me learn to code and always prompted me to ask the important questions, even when they were difficult to answer.

Thank you to Michael Lindner, who helped with all aspects of my PhD journey, from programming to experiment design to fMRI analysis. His unlimited support and confidence were invaluable.

Thank you to Zola Dean for her advice on fMRI analysis and comments on the EEG-fMRI manuscript.

Thank you to my PhD supervisor, Etienne Roesch, for giving me the freedom and opportunity to explore my ideas and approach everything with a critical eye.

Thank you to Arran Reader for providing valuable comments on the majority of this thesis, and for his support in all of my academic endeavours.

Finally, thank you to all of the students who assisted me with data collection, most notably Jade Marsh, Ioana Gaftoneanu, Isil Bilgin, and Sophie Szymkowiak.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 A brief history of change blindness

When you are watching a film, how often do you notice mistakes in the continuity? Do you notice if the main character's scarf changes position across shots, or if the background wall changes colour? The answer - probably not! When film makers started to experiment with cross-cutting, the practice of linking alternating shots, a very important practical question arose. How can a range of shots be combined in such a way that the viewer will perceive them as a single piece of action, rather than the discrete and possibly unrelated images? The ability to do this lies in the capacity of the audience to distinguish changes between interrupted images of the same scene, and in the phenomenon of change blindness.

A few decades before psychologists became interested in change blindness (ie., the finding that observers often fail to notice changes between images), it was already a well-known concept to those working in film (Simons & Levin, 1997). For example, Russian film makers Kuleshov and Pudovkin (1974) realised that they could combine shots from several locations, and, providing that no cut created unnatural motion from one scene to the next, viewers would not notice the difference (Prince & Hensley, 1992; Levin et al., 2000). Similarly, in the 1977 film 'The Obscure Object of Desire', an actor was replaced in a scene within the space of a few seconds to illustrate the transitive nature of love. Despite this seemingly obvious substitution, many audience

members did not notice the change. How can this possibly be the case? Continuity supervisors began to discuss the aspects of a scene that required the highest level of control; Rowlands (2000) suggests that the largest moving object should be prioritised, as "a viewer's attention will be drawn to it" (p.93). The same logic was applied to actors who were speaking and to objects in bright colours. Although some film makers were writing about continuity long before empirical research on change blindness, many models of attention now support these assumptions.

Several years later, researchers in psychology began to investigate the same concepts empirically {(Grimes, 196). In fact, two pioneers of change blindness research, Simons and Levin, quote reports from the film industry as inspirations for their interest in how and why we fail to see changes across image presentations (Simons & Levin, 2003). However, early studies focused primarily on the finding that changes presented during a saccadic eye movement were missed (Bridgeman et al., 1975; Wallach & Lewis, 1966), and often change blindness was an indirect finding rather than the main question of interest. It was only later that findings in the fields of film, eye-movement, visual short-term memory, and attention literature were synthesised into a coherent concept of change blindness (Irwin, 1991; Rensink et al., 1997; Simons & Levin, 1997). This shift was due in part to the work of Ronald Rensink who popularised the use of a 'flicker paradigm', in which a blank screen served the same function as a saccade without the need for expensive eye-tracking equipment. The short presentation of a blank screen or distractor image prevented any transient movement across image presentations, and therefore facilitated change blindness.

Other real-world examples of inattention blindness from Daniel Simons and colleagues sparked increasing public interest in related visual phenomenon; for example, the well-known video where observers failed to notice a confederate dressed up as a gorilla walking across a scene as they attended to actors throwing a ball (Simons & Chabris, 1999). In an experimental manipulation

of the effect found in films, Levin and Simons also determined that two-thirds of observers did not detect the change of an actor in a short film and were surprised about their lack of awareness (Levin & Simons, 1997). In fact, participants often overestimate their ability to detect changes (Levin et al., 2000). Although none of their observers identified a change to an actor's scarf, 90% of subjects believed that they would have noticed such an obvious difference in a film. This overestimation of ability has been termed "change blindness blindness", reflecting the extent to which people are unaware of their own lack of awareness (Simons & Levin, 2003).

"Change blindness blindness" has practical implications, in particular for situations in which people rely on the assumption that their attention will be directed towards what they need to see. For example, if we expect that our attention will be drawn towards a cyclist entering the path of our car, we may be less proactive in our search for such hazards. Even worse, we may divide our attention between driving and other tasks (Strayer & Johnston, 2001), increasing the likelihood of an accident. The phenomenon of change blindness therefore challenges many of the assumptions that we make about our own capabilities, and our blindness to it reflects a general misunderstanding of the mechanisms of attention (Simons & Levin, 2003). The dissociation between our phenomenological experience of complete awareness, and our actual capabilities, makes change blindness an interesting an relevant phenomenon for investigation.

1.2 The change blindness paradigm

In psychological research, two of the most commonly used methods for the manipulation of change blindness are the one-shot paradigm and the flicker paradigm. In the one-shot paradigm, the observer is shown the original image (A) and the changed image (A') once before being asked to report what they have seen. These two images are interrupted by a brief blank display,

in order to prevent detection using transient motion within the image. The advantage of this paradigm is that trials can be easily divided into those where the participant did or did not see the change.

In the flicker paradigm (Rensink et al., 1997), the observer is repeatedly presented with an (A) and (A'), interrupted in each iteration by a short distractor image. This sequence repeats, usually in the format of A, A', A, A', until the observer responds to say that they have seen the change, or a time limit is reached. Because of this, observers are aware of their own change blindness, as they continue to search the display until they see the change. One advantage of the flicker paradigm is that the search time can be used to investigate what occurs in the period leading up to change detection.

When designing a change blindness paradigm, it is important to consider the expectations of the participant. If all trials contain a change, then participants may learn to respond to changes even if they did not see them. Consequently, many paradigms will also include catch-trials, or no-change trials, where both images are identical. This prevents the participants from learning one correct answer, and facilitates signal detection analysis (Stanislaw & Todorov, 1999). One option is to include 50% catch trials. However, as researchers are often most interested in the trials where a change occurred, the percentage of catch trials is often reduced to 40 or 30%.

It should be noted that both the one-shot and flicker paradigm are examples of intentional change detection, where the participant is intentionally trying to detect changes in a laboratory environment. This contrasts to real-world or incidental detection experiments, where observers are only made aware of the task aims after they have watched a video or interacted with experiment confederates (Varakin et al., 2007; Simons & Levin, 1997).

Trial definitions

Based on participant responses during a change blindness paradigm trials are traditionally divided into the following conditions, corresponding to those typ-

ically used in signal detection analysis:

- ***See***: the participant correctly reported a change during a change trial (hit)
- ***Blind***: the participant incorrectly did not report seeing the change during a change trial (miss)
- ***Correct rejection***: the participant correctly did not report seeing a change during a no-change trial
- ***False alarm***: the participant incorrectly reported a change during a no-change trial

1.3 What causes change blindness?

The extent of the academic and anecdotal evidence available confirms the existence of the change blindness phenomenon; sometimes, under some circumstances, we miss changes that occur in the external environment. However, what exactly causes this lack of awareness, or what parameters lead to our increased/decreased awareness of changes, is still unclear (Simons, 2000).

Highly relevant to our understanding of change blindness is the debate over the extent to which we form detailed internal representations of the outside world. Many models of visual perception are based on the premise that we somehow reconstruct the world internally (Breitmeyer et al., 1982; Simons & Levin, 2003; Angelone et al., 2003; Sperling, 1960). For example, object-file theory (Kahneman, 1984) posits that we recreate detailed working memory representations called object-files. These contain details about real-world items that can be updated over time, and therefore used to identify differences. This view is tempting, as it correlates with our phenomenological experience of the world; our perception is that we are aware of objects around us, and therefore that this information must be stored somewhere in the brain. However, the fact that change blindness occurs at all seems to provide evidence against the

theory of complete internal representations (O'Regan & Noë, 2001). At the very least, if they do exist, they must be sparse or incomplete (Rensink et al., 1997).

Five possible explanations for change blindness have been suggested (Simons, 2000), which rely to varying extents on the theory of internal representations:

1) Initial representations are overwritten

One explanation for the occurrence of change blindness is that internal representations do exist, but that information regarding the changed object overwrites the initial representation (Simons, 2000). Without this visual record of the first image, no comparison can be made, and change blindness results. This is supported by findings that participants are more accurate when describing details of the second image than those of the original (Mitroff et al., 2004).

2) Only initial representations are stored

An opposite explanation is that only the initial scene is stored. If the primary goal of perception is to understand the context of our environment, then it may be plausible to suggest that we only encode information about an initial scene. Once the meaning has been abstracted, there is no need for us to re-evaluate (Friedman, 1979; Digirolamo & Hintzman, 1997), and therefore the changed scene is not encoded. For example, observers reported details about an initial view in a motion picture, rather than the changed view (Levin & Simons, 1997).

3) Features are combined

In this explanation, the observer is unable to store two complete representations, and instead combines some feature of the first view with others from the second. This combination creates a reasonable representation, and therefore it is not obvious to the observer that the details have merged. Examples

of feature combination within short term working memory tasks, or binding errors, have been noted in relation to cognitive decline (Cowan et al., 2006) which increase with age (but also see Pertzov et al., 2015).

4) Representations are not compared

There is also evidence to suggest that both the initial and the changed scene may be encoded or represented to some extent, but that their failed comparison is what leads to change blindness. For example, Mitroff et al. (2004) found that observers were able to identify pre- and post-change objects above chance level in a two-alternative forced choice paradigm (2AFC), even if they did not notice that a change had occurred.

5) Representations do not exist

Other theories of change blindness dismiss the concept of internal representations completely. In a perhaps extreme view, Gibson supported an "outside memory" hypothesis, in which internal representations are unnecessary as information continues to exist in the external world (Gibson et al., 1969). The world does not have to be represented internally, because all we need to access it is to look around (Simons & Levin, 2003). Further, if a subtle change is introduced into our environment that does not affect our interaction with the environment, then we would not notice the difference; given that we have no reference with which to compare (Stroud, 1967; Shallice, 1964).

In a refined version of this theory, Noe and colleagues suggest that the idea of 'representations' in vision is misunderstood; "rather than being a process whereby the brain produces detailed representations of what it experiences, [vision] is taken to be an activity of exploring the environment drawing on sensorimotor skills' (Noë, 2005; O'Regan & Noë, 2001; Noe et al., 2000). Unless you are physically interacting with the environment, you are not experiencing it, and therefore it is logical that you would not see changes within it. Although few researchers argue for the complete absence of internal representation of

information, many support the idea that these exist in a 'sparse' or 'incomplete' form (O'Regan, 1992; Irwin, 1991; Simons & Ambinder, 2005).

6) More than one explanation?

While it is convenient to have distinct hypotheses about the causes of change blindness, finding evidence that supports one theory while refuting all others is difficult. For example, how can we determine whether change detection failed due to a lack of complete representation, or due to the absence of a successful comparison? Perhaps, in some cases, both are true. In an experiment intending to dissociate these two explanations, Varakin et al. (2007) found that both explanations could be applied to distinct subsets of participants. Observers who missed changes and had low confidence in their ability also had poor memory for the pre- and post-change items, suggesting that they failed to represent the information. By contrast, observers who missed changes but had high confidence demonstrated good memory for these objects, indicating a comparison failure as the cause of their change blindness. Therefore, even within one experiment, several explanations were plausible.

1.4 Is there a *'sense'* condition?

To date, research has established a set of core principles regarding the nature of change blindness (Simons & Ambinder, 2005). First, attention is necessary but not sufficient for change detection; changes outside of the focus of attention are often missed, but change blindness can also occur for attended items (Levin & Simons, 1997; O'Regan et al., 2000). Second, changes to items or objects central to a visual scene are more likely to be detected (Rensink et al., 1997). Regardless of these core principles, open questions still remain. As discussed above, one open question relates to the cause of change blindness, and the extent to which we represent the outside world internally. Following on from this, if we only partially represent the objects around us, how might this

influence our ability to detect when they change? Can we then be partially aware of these changes in our environment? This partial awareness has been termed *'implicit'* detection or *'sensing'* in the change blindness literature, and the aim of this thesis is to explore both its existence and nature; does the *sense* condition exist, and if so, what is it?

As outlined in section 1.2, most versions of the change blindness paradigm ask participants to detect the presence of a change across two image presentations, meaning that trials can only be categorised as one of four types: *see*, *blind*, *false alarm*, or *correct rejection*, depending on whether the participant reports seeing a change. However, in an early experiment Rensink (2004) suggested the presence of a *sense* condition, in which observers could detect a change without fully identifying it. This can occur when observers fail to identify the object that has changed, its location, or some other detail. He argued that this condition is both phenomenologically and perceptually distinct from the traditionally reported *see* condition in which participants are fully aware of what change occurred (note that the terms *sense* and *sensing* can be used interchangeably).

The *sense* condition has high face validity, as participants in change blindness paradigms often report the sensation of suspecting a change, without knowing exactly what it was or where it occurred. Subsequently, several other researchers have explored the possibility of an awareness condition that lies somewhere between the traditional *see* and *blind* dichotomy (Fernandez-Duque et al., 2003; Laloyaux et al., 2006; Thornton & Fernandez-Duque, 2001; Galpin et al., 2008; Busch et al., 2009; Ball & Busch, 2015; Kimura et al., 2008; Hollingworth et al., 2001). For example, Fernandez-Duque et al. (2000) found that the location of a change could be identified above chance level even when participants did not report to see the change itself. This suggests that even when participants were not aware of the change, some details about it were represented or stored. Further, in Mitroff et al. (2004) participants were able to identify pre- and post-change object stimuli above chance level even when they

did not notice that a change had occurred (see Chapter 2 for a detailed review of the *sensing* literature). The presence of a *sense* condition has therefore been proposed as evidence that change blindness may arise from a failure to compare two displays or images (explanation 4), rather than a failure to encode the visual information (explanation 5) (Simons & Ambinder, 2005; Hollingworth et al., 2001).

However, the evidence supporting the *sense* condition is varied and inconclusive, and its existence has been heavily debated within the change blindness literature. The results from Fernandez-duque and colleagues (2000) were disputed by Mitroff et al. (2002) who argued that the *sense* condition, or implicit awareness of a change, can be explained by explicit mechanisms such as guessing or a process of elimination. They also detailed several experimental confounds in the original experiment which may have influenced the results (but see Fernandez-duque 2003 and Laloyaux 2006 for a discussion). The original paper by Rensink was also challenged by Simons (2005), who suggested that participants who *sense* a change simply apply a more liberal response criteria while completing the task. A detailed description of the relevant *sensing* literature can be found in Chapter 2.

Sense trial definitions

The lack of consensus regarding the nature of the *sense* condition is further demonstrated by the range of definitions that have previously been used to describe it. In addition to the traditional trial definitions outlined in section 1.2, researchers investigating the *sense* condition may refer to the following trial types (Mitroff, 2002):

- **Sensing** (or mindsight): where participants respond directly that they 'suspect' something has changed, but cannot identify what has changed.
- **Identification without detection** (or implicit awareness): where participants fail to detect a change during a change trial, but can identify

the pre- or post-change object above chance level

- **Detection without identification and/or localisation**: where participants correctly report a change during a change trial but cannot correctly identify the object that changed, and/or cannot localise where the change occurred

- **Registration without detection**: where participants fail to detect a change during a change trial, but their behaviour (usually reaction times or confidence rating) differs from trials where no change occurred

1.5 Separate processes or continuum of awareness?

Rensink (2004) originally suggested that *seeing* and *sensing* are the result of separate mechanisms that facilitate different types of awareness. This was based on the finding that reaction times were similar for trials with and without the participant reporting to *sense* a change before *seeing* it. He argued that if they were based on a continuous process, trials including *sensing* would be slower. However, Simons et al. (2005b) challenged these conclusions, suggesting that the increased time lag between the *sense* and *see* responses reported by the 'can-sense' participants simply reflected a liberal response strategy, where they were likely to indicate *sensing* even when uncertain.

An alternative hypothesis is that *sensing* and *seeing* lie on a continuum ranging from blindness to full awareness, and therefore that *sensing* is simply a weaker form of *seeing*. At some point, the threshold between *sensing* and *seeing* is met, and full knowledge of the change is achieved. At the lowest level, this threshold could be determined by neural activity. At the highest level, it could be determined by participant confidence and report strategy. In support of this hypothesis in a recent pre-print, Railo et al. (2020) used signal detection analysis to indicate that unconscious perception 'lies on the same continuum' as conscious vision by comparing behavioural responses to different trial types. While they did not use a change detection paradigm,

a growing number of recent studies are investigating this specific question regarding unconscious visual awareness, which is undoubtedly related to the phenomenon of *sensing*.

1.6 Methods of investigation

Measuring an implicit level of awareness using explicit measures, such as participants' responses, has been questioned. If we are capable of processing information that is not complete enough for explicit report, then it is possible that we will underestimate the ability of the visual system to detect changes if we only rely on self-report measures (Lamme, 2004; Fernandez-Duque & Thornton, 2000).

In response, several researchers turned to neuroimaging methods, such as electroencephalography (EEG), to identify neural correlates of the *sense* condition (Busch, 2010; Fernandez-Duque et al., 2003; Kimura et al., 2008; Lyyra et al., 2012). Although appearing to provide evidence for the *sense* condition, the definitions of *sensing*, paradigm choices, and ERP analysis parameters vary greatly within the existing literature (see Chapter 2 for a review).

While *sensing* has been explored using EEG, to our knowledge no functional magnetic resonance imaging (fMRI) data has been collected for this condition; the existing literature has only investigated the traditional *see* and *blind* conditions. For example, in a comparison between trials in which observers correctly detected changes and change blind trials, awareness was associated with activation in the bilateral superior parietal lobule, right middle frontal gyrus, and fusiform gyrus (Beck et al., 2001). An alternate comparison between correctly identified changes and no change trials found activation in a network of frontal and parietal regions, as well as the pulnivar, cerebellum, and inferior frontal gyrus (Pessoa, 2004). A similar pattern was identified for false alarm trials, where participants reported a change when no change occurred, suggesting that activity was related to the participants' perception of the change rather

than properties of the visual stimulus. Few regions were specifically activated when participants were blind to the change.

Overall, ventral areas of the brain are thought to provide the substance for visual awareness, whereas frontal and parietal activation facilitate its conscious experience (Kanwisher, 2001; Dehaene et al., 2006; Pins, 2003). The results from EEG experiments (Busch et al., 2010; Fernandez-Duque & Thornton, 2003; Kimura et al., 2008; Lyyra et al., 2012), suggest that *sensing* and *seeing* may rely on two separate mechanisms, however, it is not clear whether these mechanisms rely on distinct or overlapping networks of brain activation (Rensink, 2004; Busch et al., 2009; Howe & Webb, 2014). Investigating the *sense* condition using fMRI, with its high spatial resolution, is therefore a valuable contribution to the field.

However, the high spatial resolution of fMRI is not matched by its temporal resolution, which is in the order of seconds. As action potentials take only milliseconds to occur in the neuron, this relatively long period of a time fails to differentiate individual activity, and a large amount of temporal information is lost with fMRI BOLD signals. The opposite limitations are associated with EEG recordings, which have high temporal resolution but low spatial resolution, as recordings at each electrode represent the summation of activity across large areas of the brain. The concurrent acquisition of EEG and fMRI arose with the aim of improving the spatial and temporal limitations of each measure respectively, and therefore holds promise for the investigation of neural and hemodynamic activity associated with the *sense* condition in change blindness.

While EEG and fMRI data can also be recorded separately, an advantage of concurrent EEG-fMRI lies in the possibility of single-trial analysis, which is useful for processes that vary over time. It is also useful for experiments where the design cannot directly manipulate the trials falling into each condition; for example, because the condition label for each trial depends on participant responses. These processes include perception, attention, and awareness,

where performance/activation will vary greatly across time and between participants. Paradigms investigating these phenomena validate the use of concurrent EEG-fMRI recording as single-trial analysis would not be possible for separately recorded data sets where participant performance, accuracy, and confidence, are likely to vary between separate experimental sessions. This is true for an investigation of the *sense* condition, as we cannot manipulate when participants will *sense* or *see* a change, and it is likely that the number and timings of each trial type would vary across sessions. In Chapter 6 we provide a more detailed summary of when combined EEG-fMRI is necessary, rather than individual recording sessions.

1.7 Thesis outline

1.7.1 Aims

A consensus on the neurological signature of the *sense* condition, if it exists, is yet to be reached. The overall aim of this thesis was therefore to explore the existence and nature of the *sense* condition in the change blindness paradigm, using a combination of EEG, fMRI, and behavioural measures. More specifically, we had the following theoretical and methodological aims:

1. To build upon existing EEG results comparing the *sense* condition with other levels of awareness, such as when participants are completely *blind* to the change

2. To ascertain whether the brain activity related to the *sense* condition is different from that related to other levels of awareness, using fMRI

3. To identify brain areas with activity that co-varies with fluctuations in the EEG signal, using ERP-informed fMRI analysis

4. To investigate the influence of the MRI environment on EEG and behavioural results

1.7.2 Structure

As stated above, the main focus of this thesis was to explore the existence of a possible *sense* condition in the change blindness paradigm, using a combination of EEG, fMRI, and behavioural measures. It contains a combination of manuscripts intended for publication, as well as an opening introduction chapter and closing discussion chapter. Consequently, some information - for example the experimental paradigm and analysis methods - will be repeated across chapters.

Chapter 1 provides an introduction to the topic of change blindness, as well as an explanation of the main aims of this thesis. As Chapter 2 contains a detailed account of the *sense* condition, Chapter 1 is intended to provide a broader background to the overall topic of change blindness.

Chapter 2 contains the manuscript for a review paper entitled 'Detection without awareness in change blindness: a review of 'sensing' and implicit awareness'. This contains a detailed overview of the literature exploring if a *sense* condition exists, and the ways in which it has previously been defined.

Chapter 3 is the manuscript for an EEG experiment entitled 'An EEG study of detection without awareness in change blindness'. This was submitted to Experimental Brain Research on 7th January 2019, and a pre-print was also uploaded to BioRxiv. Corrections following peer-review were submitted on 8th May 2019 and the paper was accepted for publication on 12th June. A small pilot study conducted before the full experiment is reported in appendix A.1, and supplementary results are reported in A.2.

Chapter 4 is the manuscript for a combined EEG-fMRI experiment, entitled 'Simultaneous EEG, fMRI, and behavioural measures of detection without localisation in change blindness'. The pre-registration for this experiment can be found in appendix A.3, and supplementary results can be found in A.4.

Chapter 5 is the manuscript for a paper comparing the results from the EEG only experiment and the EEG-fMRI experiment, examining the effects of environment on both EEG and behavioural measures.

Chapter 6 contains a discussion of when combined EEG-fMRI is necessary, rather than running separate EEG and fMRI. This manuscript is titled 'When is simultaneous recording necessary? A guide for researchers considering combined EEG-fMRI'.

Chapter 7 offers an overall discussion of our findings, explains how these fit into the wider literature, and highlights the main contributions to the field.

The focus of this thesis is on the presence, or absence, of the so called 'sensing' condition in change blindness. The following review paper summarises the main definitions of the 'sensing' condition in previous literature, and provides an overview of the existing findings. Although there are several existing literature reviews outlining general change blindness findings, a review focusing on the 'sensing' condition does not yet exist.

Chapter 2

Detection without awareness in change blindness: a review of 'sensing' and implicit awareness

2.1 Introduction

The failure to detect changes between visual scenes, despite all visual information being available, is often referred to as change blindness (Rensink et al., 1997). To manipulate this in an experimental setting, the change blindness paradigm typically consists of two images displayed in quick succession that are interrupted by a blank screen or a distractor image. In some instances, the second image is identical to the first, and in others, some aspect will have changed. Participants are then asked to report if the trial contained a change or not (Simons & Levin, 2003).

In traditional versions of the change blindness paradigm, trials are divided based on participant responses into those where they were either aware, or unaware, of a change. This suggests that awareness can be neatly divided into two conditions, in which observers either have full access to information regarding the changed item, or none at all (Simons, 2000). However, many studies have challenged this viewpoint, suggesting the presence of an additional level of awareness called *sensing* or implicit awareness (Rensink, 2004; Fernandez-Duque & Thornton, 2000; Busch et al., 2010). Although some variation exists in the definition of the *sensing* condition, it represents a state of awareness

in which complete knowledge of a change is not achieved. Despite this, information about the object of change can influence behaviour in some way. This challenges the original view that vision is always accompanied by a conscious 'snapshot' of our visual world, as we are not always able to consciously report differences between visual scenes (Levin & Simons, 1997; Noe et al., 2000).

The aim of this article is to review the current understanding of *sensing* and implicit change detection, including behavioural, electroencephalogram (EEG) and eye-tracking data. An overarching definition of *sensing* is a state of awareness in which the observer is not completely aware of a change, but the presence of this change still influences their behaviour in some way. However, different researchers have operationalised this definition in varying ways. The structure of this review is therefore derived from the four main operationalised definitions of the *sense* conditions, which are each constrained by the experimental paradigms that were designed to investigate them. Given the large variation in these definitions and paradigms, it is important to consider their divergence and what this can reveal about the nature of 'sensing' in change detection.

We outline the four main operationalised definitions of implicit awareness as follows: a) **sensing**, where participants respond directly that they 'suspect' something has changed, but cannot identify what has changed, b) **identification without detection** (also commonly referred to as implicit awareness), where participants fail to report the change during a change trial, but can identify the pre- or post-change object above chance level, c) **detection without identification and/or localisation**, where the participant correctly reports a change during a change trial but cannot correctly identify the object that changed, and/or cannot localise where the change occurred, and d) **registration without detection**, where the participant fails to report the change during a change trial, but their behaviour (usually reaction times or confidence rating) differs from trials where no change occurred.

For each definition of *sensing*, we will consider the explanation of change

blindness for which the results provide the most support. Further, we will consider whether the results indicate that *sensing* is simply a weaker form of *seeing*, or that these two types of awareness may be supported by separate mechanisms.

2.2 Sensing

Definition: the participant reported that they could *sense* a change, i.e., they thought that something had changed but were not sure what. This condition can also be referred to as 'mindsight'.

Rensink (2004)

Rensink (2004) challenged the traditional view that vision must always be accompanied by a conscious visual experience, or complete internal representation of what we see. He suggested the presence of a *sense* condition during the change blindness paradigm, in which observers can detect a change without fully identifying it. Using the flicker paradigm that he popularised (Rensink et al., 1997) in which the original and change stimuli are presented in a repeating sequence, participants were asked to make two different responses during each trial. First, they responded when they 'thought' something had changed. Second, they responded when they were sure of the change, such that they could identify what change had occurred.

Participant performance was calculated by taking the difference in reaction time between their first and second response, with a differences of less than 1s indicating no sensing, and over 1s indicating a 'significant duration of sensing'. Based on these differences, participants were placed in an 'only-see' group (<5% of trials contained a significant duration of sensing), a 'can-sense' group (>5% of trials with sensing, and also a false alarm rate of <50%), or a 'guess group' for those with higher false alarm rates. From the total of 40 participants, 19 were classed as 'only-see', 12 as 'can-sense', and 9 as 'guess'.

The average response time for the first response (*sensing*) in the false alarm trials was increased (by more than 9s) than when there was a real change. Rensink used this difference as evidence to suggest that the trials in which the participant 'senses' a change were different to the false alarm trials where they are just guessing. One possible explanation for the *sensing* condition is that is is a weaker form of seeing, using the same mechanism but with reduced neurological activation. However, reaction times for 'seeing' were roughly the same, regardless of whether the participants reported also *sensing* the change.

Rensink also described the *sensing* condition as 'mindsight', as the condition involved a 'conscious (or mental) experience without an accompanying visual experience'. He suggested that the mechanisms underlying 'mindsight' and 'seeing' operate concurrently, and that the increased reaction times for *sensing* indicate that this mechanism may be slow. Both mechanisms work concurrently, and the conscious experience of the participant is determined by the mechanism that is first aware of the change.

Simons, Nevarez, and Boot (2005)

Simons et al. (2005b) challenged the conclusions of Rensink (2004). They suggested that the increased time lag between the *sense* and *see* responses reported by the 'can-sense' participants simply reflected a liberal response strategy, where they were likely to indicate *sensing* even when uncertain. They therefore refute the hypothesis that *sensing* and *seeing* are facilitated by separate mechanisms, suggesting that participants simply waited until they were certain before reporting that they saw the change. For participants with high false alarm rates, the *sense* response reflected a liberal criteria, and therefore occurs earlier in time than their final detection response. For those with more conservative criteria, the *sense* response occurs closer in time to their final detection, as they are less likely to report *sensing* the change when they are not sure. Overall, they argue that the variance in proportion of *sense* or *see* trials across participants, as well as the divergence in reaction times, could then be

explained by differences in response strategies, rather than separate underlying mechanisms. However, this argument only applies to flicker paradigms, where participant are shown the images more than once, and therefore have time to wait and confirm their suspicion (Ball & Busch, 2015).

In a follow up study, Simons et al. (2005) increased the number of no change trials and used larger range of images. Participants were also asked to make two responses; when they 'sensed' a change, and when they 'saw' it. Once responding that they saw the change, they had to localise it on the screen using a mouse. All incorrectly localised trials were removed from analysis (this is important as some studies would classify these trials as *sense* trials, using a different definition).

The distribution of performance was similar to Rensink (2004), with 15 'only-see' participants (<5% *sense* trials), 16 'can-sense' (>5% *sense* trials), and 9 'guess' (false alarm rate >50%). Simons et al. (2005) note that, in both studies, *sensing* rates did not exceed false alarm rates, suggesting that they maay be more similar than Rensink claimed. Also demonstrated is the influence of changing the 'arbitrary' time constraints used by Rensink to determine when a participant was *sensing*. When taking a time difference of 1.5s, rather than 1s, the number of 'can-sense' participants was reduced from 16 to 3.

Simons et al. (2005) suggested that 'can-sense' participants simply adopted a more liberal criterion for the *sense* response, pressing earlier than those who waited until they were more certain. This explains the differences in reaction times between 'can-sense' participants and 'only-see'. They also highlight the fact that the false alarm rate was highly correlated with the rate of *sensing*, further emphasising the difference in response bias across individual participants. Those who had more *sense* responses also had more incorrect false alarm responses, so may have been more happy to respond even when uncertain.

Both studies made the assumption that, if a change is 'sensed', it can also be 'seen', given enough viewing time. The trials where the participant reported to see the change, but then incorrectly localised them were not included in analy-

sis. Arguably, this could also be considered *sensing*, since the participants may have 'sensed' a change, but never become fully aware of the change location. In fact, this definition of *sensing* was adopted by a number of other researchers (Busch, 2009; Busch et al., 2009; Ball & Busch, 2015; Scrivener et al., 2019). However, it is difficult to dissociated this from trials where participants simply press the wrong response, or were completely unaware.

Galpin, Underwood, and Chapman (2008)

Galpin et al. (2008) also measured *sensing* by explicitly asking participants to report when they could sense a change, as well as when they could *see* it. It is not clear how these conditions were defined to the participants, but they define the sense condition in their introduction as 'a conscious feeling that a change was occurring with no accompanying visual experience'.

A comparative visual search paradigm (CSV) and eye tracking were used. In this paradigm, two images were presented on the same screen for an unlimited amount of time. The authors suggest that the traditional 'flicker' paradigm used by Rensink (2004) artificially places time constraints on the processes of encoding and comparison, which could be impaired if the stimulus is removed too early.

During a trial, participants had three response options; they could sense a change, see a change, or quit the trial. No option was given to explicitly identify a trial with no change, other than to quit the trial. One trial could therefore have multiple responses at different times. If participants pressed a response key for see or quit, then the trial would end. If participants pressed for sense, however, they would then be asked to rate their confidence on a scale of 1 (low confidence) to 7 (high confidence). They were instructed to press 7 if they had detected any physical difference between the two images. After this rating, the trial would then continue until the participant selected either see or quit. The sense trials that were used in the analysis contained a mixture of trials where the participants pressed sense and then see, as well as sense and

then quit. Only 10.5% of all trials across the group were sense trials, whereas 15.2% were false alarms, once again making it difficult to distinguish between sense trials and false alarms.

The main behavioural finding was that participants reported higher confidence in their responses during the sense trials where a change was present, compared to false alarm trials where no change was present. Confidence was also higher on trials where participants went on to see the change, as opposed to trials where they went from sense to quit. In the eye-tracking data, participants were more likely to be looking at the correct location of the change than any other region, at the time of their sense response. This suggests that they were not completely unaware of the change location during trials where they *sensed* the change.

Sense summary

In the original experiment by Rensink (2004), participants were asked to explicitly indicate when they 'suspected' a change. This required explicit knowledge of an implicit awareness, and relied on the participants' understanding of what *sensing* feels like. While this paradigm provides the closest measure of the original phenomenological sensation of the participants, that they 'suspected' a change, it is likely that some trials will be falsely categorised using this method; as pointed out by Simons et al. (2005), these definitions are heavily influenced by individual response strategies.

Rensink (2004) suggested that *seeing* and *sensing* may be facilitated by separate mechanisms working concurrently, based on the pattern of reaction times observed in his experiment. However, Simons et al. (2005) argued that this could be explained by participant response strategies, and dispute the idea of *sensing* as an individual process. Similarly, it is difficult to link *sensing* in this definition to the causes of change blindness, as in some cases, *sense* trials went on to become *see* trials. *Sensing* was not always associated with blindness, and therefore explanations of change blindness may not be

useful to explain how or why *sensing* occurs. It is difficult to assess what information was available to participants when they reported to *sense* the change in this paradigm, and therefore difficult to explain why participants did not always go on to *see* changed that they originally *sensed*. Again, this may be explained simply by participant confidence, and their desire to respond only when completely sure they were correct.

2.3 Identification without detection

Definition: the participant incorrectly reported no change during a change trial, but could identify the pre- or post-change object above chance level. This is also commonly referred to as implicit awareness, given the lack of explicit knowledge of the change.

Fernandez-Duque and Thornton (2000)

In 2000, Fernandez-Duque and Thornton criticised the use of explicit report measures to investigate change detection, suggesting that they may be insensitive to some levels of information processing in our visual system due to their reliance on conscious awareness. They referenced evidence from blindsight and hemi-spatial neglect patients (Weiskrantz, 1986; Driver et al., 1992), as well as visual masking studies where blindsight was induced (Kolb & Braun, 1995), where changes in a visual scene that were not be explicitly reported could still influence participant behaviour. If we are capable of processing information that is outside of the focus of attention and conscious report, then it is possible that we will underestimate the ability of the visual system to detect changes if we only rely on self-report. They therefore aimed to examine the extent to which non-reported changes could influence participant behaviour.

The authors used a two-alternative forced choice version of a change detection paradigm, where participants were asked to select which of two rectangles had been rotated (changed) during each trial. Each display contained 16 rect-

angles arranged in a 4x4 grid, and the distractor rectangle was always in the opposite corner to the one that had changed. Participants were first asked to report if they had been aware or unaware of the change. They were then presented with two of the 16 rectangles, and asked to indicate which of the two had changed, even if they had been unaware of the change. Participants selected the correct rectangle at a level above chance, even when they reported to be unaware of the change. This suggests that they had some knowledge of the change, even though they did not report it.

As mentioned above, Simons et al. (2005) suggested that *sense* trials may reflect a liberal response strategy, rather than a different level of awareness. In response to this, Fernandez-Duque et al. (2000) gave their participants two explicit response strategies for two different blocks: in a conservative block, they were told to select 'aware' only if they had seen the change, whereas in a liberal block they were told to select 'aware' if they 'noticed or felt' a change had occurred. Participants identified a change on 29% of trials in the conservative block, compared to 45% in the liberal block. They were therefore less likely to identify a change in the block where they were instructed to be sure that they had seen something. Further, even when participants were unaware of the change, the correct item was identified in 57% of trials, compared to 55% in the liberal block. (It is worth noting, however, that only 16 out of 20 people exceeded chance performance in the liberal block, and 19 out of 20 in the conservative block. This means that not all participants were able to report the correct change location in unaware trials.)

One concern that the authors noted regarding their paradigm is the possibility that participants will respond incorrectly during a trial, therefore labelling a trial where they did see the change as one where they did not. This would subsequently increase the number of trials where participants are unaware of the change, but can name the correct object of change. Although they use the above chance localisation from the conservative condition as evidence against this occurrence, it is just as likely that participants respond incorrectly in both

blocks. They suggested the use of neuroimaging data, as well as tasks that deviate from the two-alternative forced choice method, as future directions.

Thornton and Fernandez-Duque (2001)

A further paper by Thornton and Fernandez-Duque in 2001 aimed to extend their initial findings regarding implicit localisation of unseen changes. Specifically, they were looking for a priming effect of stimuli orientation, using a probe that was either congruent or incongruent with the orientation of the changed stimuli. They used a ring of 8 rectangles, equidistant from a fixation, and asked the participants to complete two tasks per trial; firstly, to indicate the orientation of a probe stimuli (vertically or horizontal), and secondly to indicate if any of the stimuli had changed orientation. In half of the change trials, the probe and the changed items were identical, with the same location and orientation (valid trials). In the other half, the probe was diametrically opposite the changed item (invalid) and either the same orientation (congruent) or different (incongruent). Their hypothesis was that, if changes are not seen or represented at all, then the orientation of the probe should not have any priming effect on the change response, meaning that their accuracy and reaction times should not vary across congruent and incongruent trials.

Observers were asked to adopt a liberal criterion for their responses, meaning that they should report seeing a change even if they just suspected that something may have changed. Reaction times were faster when the change was detected, and when the location of the cue was valid. This validity effect was smaller for the unaware trials, but still significant. The congruency effect was not present in the unaware trials, meaning that the orientation of the probe had no influence on detection reaction times when participants were not aware of the change. However, more errors were made for both aware and unaware trials when the probe orientation was different to the change item. Therefore, even when participants did not report seeing the change, their responses were still influenced by the congruency of the probe and target orientation.

The authors acknowledge that the error rates found in their first experiment were low, as performance was close to ceiling, and therefore that statistical power could be artificially increased. They also note that the use of a no-go response for the detection of a change that coincided with feedback for the orientation task, which may have led to errors in response.

Across both experiments, the authors provide evidence that even in the absence of awareness, participant responses were still influenced by validity and congruency effects. For example, the fact that validity effects occurred for both aware and unaware trials suggests that attention was directed to the location of the changed item, even when the observer did not notice the change. The authors note, however, that the validity effect may be influenced by contamination from aware trials, if the participant makes an incorrect response. The fact that the reaction times of aware and unaware trials differ is used as evidence that contamination did not occur to a large extent.

Mitroff, Simons, and Franconeri (2002)

In a follow up paper in 2002, Mitroff, Simons and Franconeri evaluate the claims made by Fernandex-duque and others by replicating their findings. They argue that the presence of of a *sensing* or implicit awareness condition can be explained by explicit awareness, and therefore that the *sensing* condition does not exist.

As noted in the original paper by Fernandez-Duque and Thornton (2000), participants may have been able to identify the correct object that changed through a process of elimination, rather than the storage of implicit information. If the participants knew that one of the two objects did not change across displays, then it is clear that the other object contained the change. They could therefore perform well without having any additional knowledge about the change stimuli. Mitroff et al. (2002, experiment 2) aimed to identify if explicit elimination could contribute to the main finding of the Fernandez-Duque paper (that participants could guess the correct object of change above

chance level, despite indicating that they did not see a change).

One group of participants were placed into a guess condition, where they were asked to guess where the change could have occurred if they did not see the change. A second group were placed into an eliminate condition, where they were instead asked to report all of the items that they knew did not change. All participants were asked to adopt a 'liberal criterion', meaning that they should report a change if they were 'mostly certain'. Based on a prediction of the levels of chance guessing, when taking into account set size and the number of items that can be eliminated, results indicated that participants could not perform any better than this chance level. It may be concluded from this that participants were not using implicit awareness, but that elimination strategies enabled their localisation.

However, as Mitroff points out, if participants were always and reliably able to rely on elimination strategies, meaning that they could always rule out one or two objects per trial, then their performance in the Fernandez-Duque study should have been higher. The fact that it was around 50% suggests that observers are not always able to improve their performance by relying on an elimination strategy, and therefore, the 'chance level' accuracy for each condition estimated by Mitroff can not assume that this occurs on every trial.

Fernandez-Duque et al. (2000) also found that participants were slower in identifying the orientation of a changed bar if its orientation was inconsistent with a cued bar, even if they did not report noticing the change. This incongruency effect should not be possible if they did not see the change occur, and this result therefore provided some evidence for implicit awareness. However, Mitroff et al. (2002) point out that the cued item was always diametrically opposite from the changed item, and that a learned association could facilitate this performance. They therefore ran another study (experiment 4B), in which the spatial association between a cued object and a changed object was removed. The influence of incongruency was removed from both aware and unaware trials.

Overall, Mitroff et al. (2000) provide strong arguments, and supporting evidence, to suggest that implicit awareness can be explained by explicit strategies, such as elimination of no change objects or confounds in task design. This therefore refutes the hypothesis that *seeing* and *sensing* are the result of separate mechanisms.

Fernandez-Duque and Thornton (2003)

Fernandez-Duque and Thornton (2003) further evaluated the alternative explanations for *sensing* that were suggested by Mitroff et al. (2000). In response to the elimination hypothesis, they asked participants to identify the no change item in a two-alternative forced-choice paradigm. If observers were using their knowledge of which items did not contain the change to identify the change item, then their performance should be similarly above chance level. They suggest that less than chance level identification for the no-change items would dispute the elimination strategy hypothesis. If the change item was in fact noticed at some level, even if not explicitly, then a bias may be produced towards this location, and cause a reduction in accuracy for the no-change items.

From a ring of 8 rectangles, the observers were asked to identify the item opposite the one that had changed. They were then asked to report if they had been 'aware' or 'unaware' of a change, and instructed to adopt a liberal response criteria to minimise the number of aware trials that fell into the unaware condition. When they reported to be aware of the change, they identified the rectangle opposite to the change on 81% of change trials. When they reported to be unaware, they identified these items in only 46% of trials; a level that was not better than chance performance.

The authors use this finding to dispute the possibility of elimination providing grounds for the above chance performance in unaware trials. However, being able to identify the item opposite to the one that had changed is very similar to being able to identify the object of change. If the participants stored some location information at some level, then their performance may be the

same whether they have to identify the item of change or the one opposite. The authors make the assumption that the implicit knowledge of the change location causes a bias towards that location, and therefore reduces the accuracy of those opposite. This hypothesis could be improved by asking participants to not only report those opposite, but those at other specific locations. If reporting those furthest away to the item of change is least accurate, then they would have more evidence to suggest some attentional bias in space as an explanation for their findings.

In a second experiment, Fernandez-Duque and Thornton (2003) examined the relationship between a change item and probe item. Mitroff pointed out that there was a relationship between these two items in their original study, which may have given participants strategies to boost their performance. In an adaptation of their original paradigm, Fernandez-Duque et al. (2003) removed this relationship, controlling for the relationships between position and orientation of a probe item in relation to the change stimuli.

In general, observers were slower in responding when they were aware of the change than when they were unaware. Observers made more errors when the orientation was incongruent than when the orientation was congruent, regardless of whether they were aware of unaware of the change. In reaction times, larger congruency affects were found for the aware condition, but only for trials where the location of the probe was close in distance to the change item. For accuracy, this proximity led to large congruency effects only in the unaware trials. The authors concluded that the spatial relationship in their original experiment could not be the cause of the congruency effect, and challenge the Mitroff study for the elimination of all congruency effect, even unexpectedly in the aware condition. If observers are encouraged to make speeded reactions, and the performance is not at floor or ceiling, they suggest that the effects can be found in both aware and unaware conditions.

With paradigms such as this, there is always a question of whether participants' incorrect responses could contaminate one condition or reduce the

distinction between conditions, either reflecting a lack of confidence or simple mistaken response. The authors here claim that the differences in reaction times across aware and unaware conditions, as well as the relatively small standard deviations, can be used as some evidence that contamination did not occur. However, this contamination is difficult to test. Overall, Fernandez-Duque and Thornton (2003) provide some convincing evidence against Mitroff's alternate explanation, but it should be noted that they used different paradigms, potentially driving the divergent results.

Laloyaux, Destrebecqz, and Cleeremans (2006)

As mentioned previously, Mitroff et al. (2002) challenged many of the conclusions made by Fernandez-Duque et al. (2000) with new experimental findings. Fernandez-Duque et al. (2003) ran subsequent studies, with yet another set of findings, leaving the true mechanisms unclear. However, the divergence in results between the three studies may have been compounded by methodological differences between them. In an attempt to bridge the gap between these studies, Laloyaux et al. (2006) further replicated the original experiments, controlling for all possible biases that had been identified up to that point.

Layloyaux succeeded in replicating Fernandez-Duque's (2000) findings, identifying congruency effects in reaction times and error rates for trials where the participant did not report seeing the change. This suggests that, even though the participants did not perceive the change, information about the changed item still influenced their behaviour.

Identification without detection summary

Fernandez-Duque et al. (2000) defined implicit awareness as the ability to detect the object of change above chance level, even when the change is missed. We refer to this as identification without detection, or implicit awareness. Results suggest that information regarding the changed object is available for conscious report, even if the comparison between two displays fails. This

supports the conceptualisation of change blindness as a failure to compare the pre- and post-change images, rather than a lack of visual representation of the stimuli. Despite the identification of several experimental confounds (Mitroff et al., 2002) and some back and forth between researchers (Thornton & Fernandez-Duque, 2001; Fernandez-Duque et al., 2003), Laloyaux et al. (2006) managed to successfully replicate the initial findings from Fernanez-Duque et al. (2000), suggesting that this paradigm may be a useful measure of implicit awareness. Similar findings were also reported in a study by Mitroff et al. (2004) using a paradigm with household objects instead of diagonal bars.

Of course, there is always the possibility that observers will be able to select the object of change through a process of elimination, especially with a small stimulus set (Mitroff et al., 2004). However, if participants were always able to rely on this method then you would expect their performance to be higher than recorded. It is possible that this elimination strategy does not aid performance on every trial, and participants may not be explicitly aware that this strategy is possible.

Further, if participants are able to spot 'novel' stimuli or colours in the post-change display, then they may rely on its pop-out nature to 'detect' the change (Ball & Busch, 2015). Recognising that red is a 'new' colour in the display requires a different process to successful comparison between the pre- and post-change image. There is also the possibility that trials will be incorrectly labelled as *sense* trials due to a response error (Thornton & Fernandez-Duque, 2001). Participants who were aware of the change, but respond 'no change' by mistake, will be able to identify the object of change because they did see it. Such experimental confounds could heavily alter the validity of results attributed to implicit awareness, and it is essential to consider then when designing a change blindness paradigm.

In this paradigm, participants were able to *sense* the change without completely *seeing* it. On the one hand, this could suggest that these types of awareness are supported by separate mechanisms, as one could occur without

the other. However, these results could also be used to suggest that *sensing* is simply a weaker form of *seeing*; the threshold for *seeing* may not have been reached in *sense* trials, and therefore participants were unaware of the change. It is also difficult to assess the cause of change blindness given these results. Knowledge of the location of the changed item would suggest that participants were aware of the difference between the pre- and post-change item. However, they still failed to detect that a change had occurred, suggesting that the comparison between displays may have been unsuccessful.

2.4 Detection without identification and/or localisation

Definition: the participant correctly reported a change during a change trial, but incorrectly identified the object that changed, and/or incorrectly reported the location at which the change occurred.

Busch, Fründ, and Herrmann (2009)

Busch at al. (2009) emphasised the use of electroencephalography (EEG) to enable improved distinctions between levels awareness. They define sensing as 'the correct conscious detection of change (reporting that something has changed) without correct identification of what has changed'. They make the distinction between this, and 'implicit processing' in which observers are completely unconscious of any changes.

In a change blindness paradigm using a 4x4 array of objects, participants were asked to respond if they saw a change or not, and then to identify the object that was changed. Trials in which participants were successful at both were classified as full awareness trials, or 'identified' (*see*), and those where participants could not identify the object as 'detected' (*sense*). Specifically, identification of either the pre or post-change object was considered a correct identification. Participants were instructed to respond with a liberal criterion, and report that they saw a change even if they were unsure which object

had changed. However, they were also instructed not to guess, but to respond 'unsure' if they felt uncertain (specifically, when they could not decide). These trials were excluded from further analysis (just 15% of trials). Participants correctly identified the change on 68% of change trials, and selected the correct identity on 37% of change trials.

Several event-related potentials were examined in the EEG data. The visual awareness negativity (VAN) was defined as the mean amplitude from 130-320 ms after the changing stimuli. VAN amplitudes were greater in both the identification (*see*) and detection (*sensing*) conditions compared to change blindness, although the difference was greater in the identification condition. This effect was greatest in a central ROI, and was not present for correct rejections or false alarms.

The P3 peak was defined as the mean within 400-600 ms, and amplitudes followed the same trend as the VAN, maximally at posterior electrodes. A difference in P3 was also found between false alarms and correct rejections at frontal electrodes. The change-related positivity was defined as an asymmetry at 90-150 ms in contralateral recording electrodes. Statistically significant peaks were found only for identified trials, in a posterior ROI. Similarly, the N2pc (at 220-370 ms) was found only when the change was identified, but not when participants were blind to the change, or in the detection condition (*sensing*). The fact that the N2pc and change-related positivity were found exclusively in the identification condition suggests that the *sense* condition is not simply trials where the participants respond with the wrong button. If this were the case, these change related asymmetries would be similar.

The same analysis was run again with an addition between-subject factor of false alarm rate. Simons et al. (2005) suggested a relationship between *sensing* and false alarm rates, suggesting that the *sensing* condition could be explained by a more liberal response criteria in some participants. The aim of this analysis was therefore to identify if there were any differences in ERPs between participants with high and low false alarm rates. Two sub-groups

were created by dividing the group based on the number of false alarms made by each participant (note that the study sample size was 16, leaving 8 in each sub-group). The VAN amplitude for detection (*sensing*) at central electrodes was actually greater in the low false alarm rate group. None of the other analyses found a significant effect of group. However, the behavioural results indicated that those with false alarms also had a higher rate of *sensing*.

The authors conclude that seeing a change is not a stronger version of sensing the change, as ERPs can be found for seeing that are not present for sensing. This conforms to the hypothesis from Rensink (2004) that seeing and sensing represent two separate and competing mechanisms. However, conclusions can not be drawn from null results, and the lack of significant N2pc for *sensing* should not be used as evidence that the two awareness conditions in this experiment were statistically different.

Overgaard, Jensen, and Sandberg (2010)

In a comment on Busch et al. (2009), Overgaard, Jensen, and Sandberg (2010) outlined 'methodological pitfalls' in their approach to identify a*sensing* condition.

Subjects were instructed not to guess in the paradigm used by Busch et al. (2009), meaning that they only reported a change when they were conscious of it. Overgaard points out that this is not consistent with the aims of the study, as they aimed to identify 'implict' change detection, without full knowledge of the change. Although participants were asked to adopt a liberal criteria, and therefore may respond to a change without being able to identify it, participants may have found it difficult to distinguish between these trials, and those where they should respond 'unsure'. The ability for participants to respond 'unsure' also impacts the use of signal detection on the data, as these trials were excluded rather than being included as 'miss' trials.

Busch, Dürschmid, and Herrmann (2010)

In another EEG study of the *sensing* condition, Busch et al. (2010) distinguished between localisation and identification of a change in object. Although similar to their 2009 study, the paradigm differed in that participants had to identify where the change occurred on the screen (change left/change right/no change), and then identify which object had changed. In their previous study (2009) participants were only asked to respond change or no change, rather than suggest the location of the change. Trials could therefore be divided into 'blind' trials, where the participants incorrectly responded no change, 'correct localisation' trials where the participants got the location of the change correct, 'incorrect localisation' where the location was incorrectly reported, and 'change identified', where both the localisation and object identification were correct.

The visual awareness negativity (VAN) was defined as the mean within 250-350 ms from the second display onset. The change related-positivity was defined as the asymmetry within 120-190 ms, the N2pc as the asymmetry within 250-270ms, and the late posterior contralateral positivity (LPCP) from 480-600 ms. Increased VAN and LPCP amplitudes were found when changes were both localised and identified, at posterior electrodes. No effects were found for trials that were only localised and not correctly identified. This suggests that the VAN and LPCP may be indexing knowledge of the changing object's form, rather than its location. This is supported by their previous finding that VAN amplitudes were increased when subjects correctly identified the object that changed (Busch et al., 2009). Further, the fact that participants could sometimes localise the change without identifying it suggests that these two processes may not be dependent.

In comparison, the N2pc and change-related positivity were found for correct localisation and identification trials, as well as those where the change was only localised. This was not found for incorrectly localised or change

blind trials. Unlike the VAN, these peaks were significantly increased for both identification and localisation, suggesting that they are not specific to one type of knowledge.

Busch and Ball (2015)

In a behavioural study, Ball and Busch (2015) suggested that the 'sense' condition, detection without identification, occurs when a change in visual stimuli lies outside of the focus of attention. These changes would go unnoticed unless accompanied by the appearance, or disappearance, of a unique feature. To test this, they manipulated changes in coloured stimuli, using either novel or repeating colours from an initial display. This change could occur in one of two following displays. In a detection task, participants had to report on which display the change occurred. In a localisation task, they had to click on the location of the change.

Detection was better for colour changes that were unique, meaning that the colour of change was not used in any other objects in the display. Increasing set size resulted in poorer change detection, but detection of unique features was less influenced by the number of items on the display. It could be argued that using a unique colour created a 'pop-out' stimuli, meaning that it could be identified purely based on the presence of a 'new' colour, rather than on a recognition that the previous colour was different. Localisation was also better for unique colour changes, and decreased with set size. However, there was no difference in localisation for unique and non-unique features. Is is therefore that participants were simply remembering the new colour name, rather than the object of change. This would explain why unique colours could be identified more easily, but localisation was not improved.

Sensing trials (detection without localisation) were similar across set size for unique features, but decreased with set size for non-unique changes. The number of distractor objects did not influence the rate of 'sensing' when the change was unique, suggesting that less attentional resources were required

for unique changes. However, set-size reduced the rate of sensing when the change was non-unique. Further analysis found that, when incorrectly localising the change, participants reported the location at random, rather than choosing an incorrect location that was close to the site of the change. This is evidence against the hypothesis that participants were just imprecise in their localisation.

The aim of a second experiment was to determine if participants who could not localise a change would also be unable to identify it. The set size was fixed to 7 objects, and additional question asked participants 'which colour was part of the change?'. The correct colour could be the one before of after the change. In concordance with experiment 1, they found that sensing rates were higher for unique colour changes than non-unique, only exceeding chance performance for unique features. This finding is explained by the fact that participants were more likely to select non-unique colours in both the unique and non-unique conditions. This reduced performance in the unique conditions, but increased performance in the non-unique.

Overall, when participants were not able to localise the change (sensing), they were also unable to identify the colour that changed. They were also more likely to guess a non-unique colour that had changed, suggesting that they weren't able to use the 'pop-out' nature of unique colour to aid them in colour selection. The researchers also note the low occurrence of detection without localisation, referencing other studies with similar findings (Galpin et al., 2008; Rensink, 2004; Simons & Ambinder, 2005; Howe & Webb, 2014; Haberman & Whitney, 2011). This, highlights individual differences in task performance, and may have been because the task was too easy for some of the participants with only 8 stimuli.

In summary, researchers found that a failure to localise a change in space was always accompanied by a failure to identify the colour of the change, again suggesting failures are the pre-attentive level. Further, when participants incorrectly localised the change, they guessed another location at random. When

dividing participants based on the number of *sense* trials they demonstrated, the response bias of both 'detectors' and 'non-detectors' was conservative, suggesting that another explanation is needed for the distinction between these two groups. Ball and Busch also highlight the distinction between *sensing*, where participants know that they are details about the change that they cannot report, and implicit change detection, where they are unaware of a change but the change still influences their performance (eg: Fernandex-Duque et al. 2003).

Howe and Webb (2014)

Howe and Webb (2014) sought to establish if the *sensing* condition can be explained by response bias, or participant guessing. They used a one-shot paradigm, where the pre and post-change images are shown only once, to remove the possibility of participants simply waiting to verify their answer before responding, as could be possible in the flicker paradigm. They also included catch trials. Participants were shown colour portrait photographs, and asked if they saw a change (yes/no). If they responses yes, then they were asked to click on one of 9 possible items that changed within the picture.

Using an equation based on conditional probability, they estimated the number of trials that participants would guess the correct answer, rather than get the detection correct due to seeing the change. They found that participants had more trials where they could detect but not identify the change than you would expect if they used a guessing strategy alone.

While Howe and Webb found evidence that change detection can occur without accompanying localisation or identification, others such as Fernandez-Duque (Fernandez-Duque & Thornton, 2000) suggest that localisation can occur in the absence of change detection. However, the authors suggest that their results are not necessarily in conflict with other groups. Based on this divergence in results, Howe and Webb suggest that there may be a double dissociation between detection and localisation; two distinct aspects of change

detection that may be facilitated by different underlying processes.

Detection without identification and/or localisation summary

This specific definition of *sense* trials identifies them based on the participants' ability to provide further information about the object that changed. We refer to this as detection without identification and/or localisation. Although participants are able to detect that a change has occurred, they do not know what has changed, or where.

Is is important to note the distinction between definitions; in some experiments the participant is asked to identify the object that changed, independent of its location. In others, the participants have to locate the change in space, and some experiments use a combination of both. Busch (2013) found that participants were able to identify the changing object at a level above chance, even on trials where the localisation of the change was incorrect. This suggests a dissociation between identification of an object and localisation. Mitroff and Simons (2002) further extend this dissociation with the finding that changes could not be localised before they were explicitly detected. Researchers should therefore consider the definition based on their research question.

Again, there is the possibility that trials will be incorrectly labelled as *sense* trials due to a response error. Participants may be aware of the change location, but press the wrong location button in response. However, Ball & Busch (2015) found that when incorrectly reporting the location of change, the location reported was random with no relationship to the correct response. This suggest that participants were not simply imprecise in their localisation, at least for their particular paradigm.

Busch et al. (2009) used the presence of ERPs for their *see* condition that were absent for *sensing* as evidence that these two mechanisms are separate. If the N2pc is only present when we are fully aware of the change, then perhaps *sensing* relies on a different set of activations within the brain. However, the null result for *sensing* could also be due to a weaker signal and therefore

reduced power to detect it. It is difficult to draw conclusions from a null result in one condition, and you may also expect a non significant ERP if the threshold for *seeing* had simply not been met.

2.5 Registration without detection

Definition: the participant incorrectly reported no change during a change trial, but their behaviour (usually reaction time or confidence rating) differed from that in trials where no change occurred

Mitroff, Simons, and Franconeri (2002)

Williams and Simons (2000) identified implicit 'registration' by comparing reaction times for trials where no change was present, versus those where a change occurred but the participant failed to identify it (they were blind to the change). While these two trial types both involve a response of 'no change' from the participant, any difference in the reaction times of these responses could indicate an implicit influence of the change on their behaviour. The researchers reported that 68% of their observers were slower to respond 'no change' when a change was present, which they use as evidence to suggest implicit change detection may have occurred. However, Mitrofff et al. (2002) argued that this could be explained by the participants making erroneous 'no change' responses, which would be accompanied by lower confidence, and therefore slower reaction times. They also highlight the bias towards reporting 'no change' for trials where they were not certain of the change, which would also increase reaction times.

To test this hypothesis, Mitroff et al. (2002, experiment 1) recorded levels of confidence (from 1 to 5) in a simple change detection paradigm. Similarly to Williams and Simons (2000), they found that 80% of observers responded 'no change' more quickly when there was no change than when there was a change, and that the differences were significantly different. However, when

regressing the confidence scores and the presence or absence of a change onto the reaction times, they found that confidence levels explained more variance in reaction times (29.24%) than the presence-absence of a change (15.07%). This therefore suggests that participant confidence had a greater influence on reaction times, and that these differences may be useful to infer implicit detection. However, it may be difficult to differentiate between awareness and participant confidence, as they are likely to be correlated; participants will be less confident when they cannot report all details regarding a change.

Fernandez-Duque, Rossi, Thornton, and Neville (2003)

A follow-up paper from Fernandez-Duque (2003) used EEG to distinguish between awareness conditions. It should be noted that they did not explicitly test for the *sensing* condition within this paradigm. However, the abstract states an aim of the paper is to find 'neural substrates of implicit representation of change'. Another aim of the paper was to distinguish between attention and awareness in change detection, which is often hard to disambiguate.

A flicker paradigm was used, similar to Rensink (2000), where the change and no change stimuli are consistently switched until the participant noticed the change. One advantage of this paradigm is that, for each image, you will have trials where the participant does not see the change, as well as trials where they do. The presentations of the stimulus occurring before the participants reported the change were classified as 'unaware' trials, and those after the identification as 'aware'. Notably, the scenes contained either a location change or colour change.

Throughout this paper, the authors refer to *blind* stimuli presentations (where a participant had not yet detected a change) as 'unaware' or 'implicit' trials, making their definition of a *sensing* condition broader than other papers. By using all of the *blind* trials in their important 'unaware' condition, that indexes implicit awareness, they make the assumption that implicit information about the change was available for all trials. Is this necessarily the

case? They can only make the distinction between *blind* and *see* trials, like the traditional paradigms, rather than a separate 'implicit' level of awareness condition.

In comparison to traditional paradigms, however, a more complex paradigm was used. This involved an initial search for a change, the removal of that change, a potential second change occurrence, and then a repeat of an initial change. This complex design allowed the trials to be divided into several conditions, including attention search for the initial change, and focused attention for a re-occurrence of a change.

For aware versus unaware changes, a frontal negativity was found at 100-300 ms, larger for aware trials. At posterior sites, greater positivity for aware trials was found at 120-310 ms. As this pattern was very similar, they hypothesised that the differences between aware and unaware were based on the differences in search strategy during the trial. To test this, they compared focused attention minus attention search, to aware change minus unaware change. As expected, they found no significant differences in the 100-300 ms time window.

At medial sites, a significant differences was found between aware and unaware trials at 350-600 ms, predominantly at electrodes Fz, Cz, Pz, and IPz. As this effect was not found in the previous analysis, comparing focused and non-focused attention, it is unlikely that this effect was driven by differences in attention. A further comparison, between unaware trials and no change trials, revealed a bilateral positivity at anterior sites between 240-300 ms, and at midline sites (particularly Fz and Cz). Although the state of awareness was the same across conditions (not aware of a change), these differences in activation were present, suggesting the presence of some activity that was elicited by the change.

Due to their variation on the standard change detection paradigm, the researchers were able to distinguish between differences in activation related to attention, compared to differences related to awareness. Overall, they conclude that frontal activation may be related to the attentional control mechanisms

(Kastner, 1998), whereas the posterior activation may be related to awareness (Beck et al., 2001). However, as mentioned, they are only comparing trials where the participant has noticed the change, to those where they are still blind to it. They do not explicitly measure a *sensing* condition, as seen in other studies (Fernandez-Duque & Thornton, 2000; Rensink, 2004; Galpin et al., 2008).

Kimura, Katayama, and Ohira (2003)

The title of this paper declares it to be 'a replication of Fernandez-Duque et al. (2003)'. However, the paradigm and analysis methods used are very different to the paper mentioned above. One similarity is that they also use *blind* trials to indicate implicit awareness, in comparison to *see* trials where the change is explicitly detected.

In the paper by Fernandez-Duque et al. (2003), a flicker paradigm of real life images was used, during which changes were added, removed, and returned, to focus on the relationship between attention and awareness. Crucially, in the flicker paradigm, the original and change stimuli are swapped repeatedly until the change is detected (or an arbitrary threshold is reached). The paradigm used by Kimura et al.(2008) is very different. Here, participants were presented with a ring of 6 grey circles of different luminance, and therefore different colour. Participants were presented with consecutive displays, interrupted with a blank screen, and asked to report if they noticed a circle change in colour.

In order to increase cognitive load, and subsequently task difficulty, the participants were given an initial task, located at the fixation. In the centre of the ring was a smaller circle, that would change in size on some trials. On each trial, participants had to respond as quickly as possible when detecting a change in size at the fixation point. They were then given the opportunity to report colour changes in the surrounding circle. A combination of colour/size, change/no change, resulted in four possible trials types.

Performance on the task at fixation was 85%, regardless of whether the size

change was accompanied by a colour change or not. For colour detection in the outer circle, accuracy was 3.9% when a size change also occurred, and 16.8% when a size change was absent. The task at the fixation therefore succeeded in making the colour discrimination task more difficult. This was a crucial manipulation for subsequent analysis, which focused on *blind* trials.

For EEG analysis, a difference wave was calculated between no change trials, and change trials where the participant failed to notice the change (*blind*). Significant peaks were found at central electrodes (Fz and Cz) at 160-180 ms after the onset of the changed stimuli. This difference occurs just after the initial P1/N1 complex, potentially in a P2 window, and is much earlier than the peak found by Fernandez-Duque et al. (2003). They claim, however, that their results are 'highly consistent' with the Fernandez-Duque study, and suggest that the shorter latency may reflect the use of simpler visual stimuli.

Registration without detection summary

In this definition, *blind* trials are compared with trials containing no change. The finding that reaction times are increased when participants are *blind* to the change is used as evidence that the presence of the change influenced behaviour, and therefore some aspect of the change was registered. Based on this definition and the use of the flicker paradigm, the assumption is made that all *blind* trials could eventually become *see* trials where the observer detects the change, given enough iterations of the stimuli. Therefore, the assumption follows that observers can *sense* changes on all trials. Is this necessarily the case? Models of conscious awareness acknowledge the fact that observers will not have any knowledge of stimuli on all trials, due to a lack of attention (Lamme, 2003, 2004; Dehaene et al., 2006). To avoid this, researchers often exclude trials where participants fail to correctly identify or localise the changing object, using this as a marker for unidentified changes. This therefore conflicts with the definition of implicit awareness used by other researchers.

It is also possible that the *blind* condition contains a combination of trials.

In some cases the participant is certain that nothing occurred and there is no possibility of explicit report. This may be explained by a lack of attention towards the stimuli, preventing any knowledge of the change (Lamme, 2003, 2004). In other cases, participants may have some knowledge of the change but are uncertain about their detection. Taking a mean value over all trial combined may therefore dilute the results, and as reported by Mitroff et al. (2002) participant certainty may provide a better explanation for differences in reaction time.

The results from paradigms investigating registration without detection suggest that some information about the change is stored in some trials, even when participants fail to report a change. Once again, this could provide evidence that change blindness occurs as a failure to compare images, rather than represent them, as if nothing was represented then no effect on behaviour would be expected. It could also occur due to a binding error, such that aspects of the pre- and post-change scene become mixed in their representation and therefore reduce the confidence of the participant. In the context of the flicker paradigm, if all *sense* trials can become *see* trials given enough time, then this would suggest that *sensing* is just a weaker form of seeing, rather than a separate mechanism, as the evidence for the change increases to the point of conscious subject report. The fact that some changes are only ever *sensed* could provide evidence against this theory. However, these trials may be more difficult, and therefore need more time than allowed during the experiment for the participant to *see* the change completely.

2.6 Others

This section contains additional papers where the definition of implicit awareness did not match any of our definitions.

Layolaux, Devue, Doyen, David, and Cleeremans (2008)

An important distinction between this paper and others, is that the participants were not told that changes may occur between displays. Therefore, all of the participants included in the results believed that the images they were presented were static, without any changes. This is arguably very different to the traditional version of the change detection task, where participants are explicitly told to detect differences between stimuli presentations, and are actively looking for them (Laloyaux et al., 2008).

Their stimuli consisted of face stimuli that continuously morphed from neutral to emotional, or emotional to neutral, over 12 seconds. The level of emotional content was also controlled, at 25%, 50%, or 75%. Observers were also presented the static versions at 0% and 100% emotion (0% is the neutral face). They were asked to memorise the faces, for the purpose of a subsequent recognition test. After each presentation, participants had to select the face that they just viewed (from a selection of 5), and rate their confidence from 1 to 4. Any participants that reported noticing anything unusual or 'strange' we not included, as they could not be considered 'unaware' of the changes. This resulted in 14 out of the 49 participants being excluded.

The face morph with either 0/100% emotion was chosen most frequently for the static stimuli, and significantly more often than in the changing condition (this was the only face shown in the static condition). The emotion of the face had an influence on accuracy in the static condition, with higher accuracy for the emotional face.

In the changing condition, the direction of change had no effect. Participants were more likely to select a face morph that was close to the final one, suggesting a recency effect, but less likely to chose the initial face, demonstrating the lack of primacy effect. Confidence was significantly higher for the static condition, suggesting that even though the participants were not explicitly aware of the changing stimuli, they were perhaps noticing that something

was different. As they did not report this after the experiment, the authors suggest that the participants were doubting their own judgment during the changing trials, rather than questioning the stimuli.

The finding that participants reported the 0/100% morph more frequently in the static condition, although the final image in the changing condition was also the 0/100% morph, does suggest that something different may be occurring between these two trial types. However, as the authors note, this could be reflecting the greater variability in the changing condition, which makes the task more difficult. The difference in confidence also provides evidence that the two trial types are experienced differently by the participant. This experiment has to advantage that only participants who were unaware were included, meaning that the differences in response are not due to different perceptual experiences; all participants believed that the trials were completely static. It would be interesting to see a comparison with those participants who were aware that something was changing, but this is not reported here.

Chetverikov, Kuvaldina, Macknnes, Hohannesson, and Kristjansson (2018)

Chetverikov et al. (2018) aimed to investigate the role of covert attention on the rate of change detection. Given that changes occurring within the focus of covert attention are more likely to be detected, as well as those that are close to the point of fixation, evidence suggests that changes are more likely to be detected when covert attention is allocated towards them. However, the presence of covert attention may not be sufficient to facilitate change detection, as fixations over the target do not always lead to awareness of a change "attentive blank stares", Caplovitz et al., 2008). Using a gaze-contingent display, the authors attempted to 'tether' covert and overt attention with the hypothesis that this would reduce so called attentive blank stares.

In a gaze-contingent display, the visual scene is restricted to include only the location of gaze in the visual scene. This eliminates the possibility of peripheral information processing, and therefore reduces the extent of possible covert

attention. In a slight modification to the traditional paradigm, Chetverikov et al. increased the size of the visual display around each gaze position, in order to allow for implicit processing on neighbouring regions of the display. However, if participants fixed their gaze for too long, the size of the visual display would decrease, to encourage active exploration of the display. A mouse-contingent paradigm was used as a control, in which covert attention is less restricted, as the visual display is determined by mouse position rather than the participant's gaze. They hypothesised that gaze-contingent change detection would be lower than in the mouse-contingent paradigm, as covert attention could not be used to aid identification of a change. Conversely, "attentive blank stares" would also be reduced in a gaze-contingent display, as the dissociation between covert and overt attention was restricted.

In order to facilitate direct comparisons in eye-movements between catch trials (with no change) and change trials, one object per catch trial was randomly assigned as the 'target'. Change blindness occurred at a similar rate for mouse and gaze-contingent paradigms. However, participants had a greater number of fixations on the target in the mouse-contingent condition, and change detection was faster. This suggests that uncoupled covert and overt attention facilitates better performance in change detection. When controlling for total trial duration, the time spent fixating on the target was higher in change trials than no change trials, even when the change was not detected. This indicates implicit processing of unreported changes. No evidence was found for a reduction in "attentive blank states" in a gaze-contingent paradigm, and the authors suggest this as an avenue for future research.

Overall, the fact that fixations on the target were higher in change trials than no change trials, despite no explicit change detection, suggests that some information about the change was processed. Although the de-coupling of covert and overt attention did facilitate faster change detection, accuracy did not increase as expected. It is therefore difficult to establish the relationship between implicit change detection and covert attention to changes.

Reynolds and Withers (2015)

Reynold and Withers (2015) used eye-tracking as an objective measure of participant's knowledge of changes in a flicker paradigm. This was based on the assumption that a participant's eye gaze is correlated with the location of their directed attention (Finlay & Gilchrist, 2003). The stimuli were based on those used in Smilek et al. (2000) and Mitroff et al. 2002, with the digits 2, 4, and 8 arranged in an invisible 6 x 6 matrix. Digits changed in pairs; 2 to 4 (with five feature changes) or 2 to 8 (with two feature change). Set sizes of 4, 10, and 16 were used. Participants were asked to indicate when they were aware of the change. Location accuracy was assessed by asking them to fixate on the area where they saw the change occur. Note that trials in which the location was incorrect were excluded from analysis, so the researchers were not explicitly testing *sensing*.

As set size increased, so did reaction times and number of fixations. The search slopes were shallower for the five-feature changes compared to the two-feature changes, as found previously (Smilek et al., 2000; Mitroff et al., 2002). The researchers conclude that unattended implicit changes can guide attention towards the location of a change, resulting in explicit awareness of the change. *sensing* of unattended changes is therefore a useful process for guiding explicit change detection, as attention cannot be allocated everywhere at once.

Lyyra, Wikgren, Ruusuvirta, and Astikainen (2012)

Lyyra at al. (2012) used the visual mismatch negativity (vMMN) ERP, occurring at posterior electrodes between 150-300 ms, as an indicator of implicit change detection. They hypothesised that if explicit change detection relies on implicit detection, then detection cannot occur in the absence of a vMMN response. Using a flicker paradigm, change and no change trials were compared using the time period before the participant identified a change, for which they were encouraged to press a button at first thought of detection. They there-

fore used a liberal criteria for awareness of the change. A greater vMMN was found for change trials than no change trials, suggesting that the information regarding the changed item was processed to a certain extent before explicit detection.

Another interesting finding is that the vMMN effect was not present in trials with a longer ISI of 500 ms, compared to 100 ms, suggesting a possible link to working memory decay. This therefore goes against the hypothesis that vMMN is necessary for explicit detection, as changes were identified even with an ITI of 500 ms.

It is important to note that there was no distinction made between change trials where participants could or could not later detect the change. However, only 3% of trials were reported as misses or false alarms, making these a small proportion of all trials. Also, *sensing* was not directly manipulated here; all types of change trial were included, whether participant would have been able to correctly identify and/or localise changes or not.

2.7 Discussion

Participants in change blindness experiments often report that they 'suspect' a change has occurred, but cannot provide any details about the change (Rensink, 2004; Simons & Ambinder, 2005). This experience appears to be phenomenologically different from complete change blindness and from full awareness. But how does this relate to neural activity? Are *sense* trials characterised by a different pattern of brain activation, preventing full awareness of the change? Or, is the sensation of limited awareness solely linked to response errors and a lack of participant confidence? The large body of evidence considered here suggests that *sensing*/implicit awareness can be separated from full awareness and change blindness, both behaviourally and in neuroimaging data. Even when all aspects of a change cannot be reported, the behaviour of the observer appears to be influenced by the presence of the change, indicated

by increased reaction times, fixation durations, or ERP amplitudes.

However, one difficulty faced in the investigation of the *sense* condition is how to effectively measure it. By definition, *sensing* reflects a state of awareness where full report is not possible, and therefore experimental paradigms are designed to infer the trials in which *sensing* occurs. In sum, the lack of single definition for this condition has led to a range of methods being employed in the literature, and thus a lack of consensus for the nature of this specific level of awareness.

How do these definitions overlap?

Within the four main definitions within the literature, two are directly opposing; one with detection but not identification, and the other with identification but not detection. Are these definitions two ways of measuring the same underlying construct, where complete explicit knowledge is not possible, or are these indicators of two separate mechanisms of change blindness? Is it the case that both trial types are possible at different times, or will they always overlap?

It is rare for experiments to explicitly test more than one definition of *sensing*, and often exclude trials conforming to a different definitions. For example, in the original experiments where participants reported when they 'suspected' a change as a measure of *sensing*, trials in which the participant incorrectly localised the change were removed from analysis (Simons & Ambinder, 2005). However, Howe and Webb (2014) suggest that there is a double dissociation between detection and localisation; they are two distinct aspects of change detection that are facilitated by separated underlying processes. Similarly, Watanabe et al. (2003) found that change identification was impaired when distracting 'mudsplashes' were presented at the point of change, whereas localisation was impaired when 'mudsplashes' were presented preceding the change. They therefore suggest that identification and localisation are supported by separate mechanisms. Future research could attempt to measure a combina-

tion of definitions within the same paradigm, in an attempt to distinguish between these trial types.

Can sensing be explained by explicit mechanisms?

Another important question is whether *sense* trials are behaviourally different to *blind* or *false alarm* trials, as others have suggested (Fernandez-Duque et al., 2003; Galpin et al., 2008), or whether they can be explained by explicit mechanisms (Mitroff et al., 2002). One explanation for the presence of a *sense* condition in change blindness is that it reflects a liberal response criteria, such that participants report seeing a change even though they were not certain that it occurred (Simons & Ambinder, 2005). In other words, they make a *'false alarm'* during change trials. If this is the case, then these trials may be similar in number to *false alarm* trials, where participants incorrectly report a change for identical displays where they could not have seen a change. The ability to make this comparison is reliant on the presence of 'catch trials', or trials in which no change occurs. Catch trials are essential for distinctions between *sense* and *false alarm* trials to be made, and for measures of d'prime to be calculated, although some previous experiments omit them (Fernandez-Duque & Thornton, 2000).

A second explanation for the *sense* condition is that it contains trials for which the participant mistakenly reported a change, even though they were not aware of it. In this case, reaction times for *sense* trials should be similar to those for *blind* trials, particularly those where participants were uncertain of their responses. Participant certainty may also be a valuable method in assigning condition labels for analysis (Galpin et al., 2008). Theoretically, *blind* trials with an *uncertain* response could indicate *sensing*, as participants did not report seeing a change but were not confident in their response.

Unfortunately, the trial distributions are often uneven, with many studies reporting low numbers of *sense* and *false alarm* trials (Ball & Busch, 2015; Galpin et al., 2008; Rensink, 2004; Howe & Webb, 2014). This reduces the

statistical power available for comparisons to other trial types. One explanation for low trial numbers may be the unwillingness of participants to report their detection of the change in the face of uncertainty. A solution is to give explicit instructions to the participants before the task, encouraging them to adopt a liberal response criterion or explaining when they should make certain responses (Galpin et al., 2008; Fernandez-Duque & Thornton, 2000; Mitroff et al., 2002; Thornton & Fernandez-Duque, 2001). When observers were encouraged to adopt a liberal response criteria and identify a change even if they only suspected it, the rate of *sensing* increased (Fernandez-Duque & Thornton, 2000). Without direct instructions, it is likely that participants will respond conservatively, for fear of being wrong.

The low number of *sense* and *false alarm* trials across experiments may also be an indication that task difficulty is incorrectly specified. If a participant finds the task easy, then they will be able to both detect and identify the change during most trials. Thornton & Fernandez-Duque (2001) reported low error rates due to near ceiling performance. Conversely, if the task is too difficult, then participants will not be able to detect many changes or detect them without identifying them. For example, Fernandez-Duque & Thornton (2000) increased the number of trials featuring implicit awareness by reducing the set size from 16 to 8. Although large individual differences in performance are well documented within the wider working memory literature (Luck & Vogel, 1997; Vogel et al., 2005), to our knowledge, no study investigating *sensing* has attempted to adjust the task difficulty to suit participants' individual capability. This may help to increase the number of trials where participants lack complete awareness of the change, and therefore facilitate comparisons with higher statistical power.

Links with theories of change blindness

As discussed in Chapter 1, there are a number of explanations for the phenomenon of change blindness. One theory is that blindness occurs due to a

failure to encode either the pre- or post-change display. If no information is stored, then differences cannot be detected. However, previous researchers have argued that the presence of the *sense* condition provides evidence against this hypothesis, as observers can identify object from both displays above chance level (Simons et al., 2005a; Hollingworth et al., 2001). Even when observers are not explicitly aware of a change, their ability to correctly identify items from both displays suggests that they have stored some information about them. In this case, it is more likely that a failure to compare the pre- and post-change display resulted in the lack of full awareness.

However, it is also possible there are multiple explanations for change blindness and *sensing*. In an experiment intending to dissociate between different explanations, Varakin et al. (2007) found that multiple explanations could be applied to distinct subsets of participants. Observers who missed changes and had low confidence in their ability also had poor memory for the pre- and post-change items, suggesting that they failed to represent the information. By contrast, observers who missed changes but had high confidence demonstrated good memory for these objects, indicating a comparison failure as the cause of their change blindness. Therefore, even within one experiment, several explanations were plausible.

Links with other literature

Similarities can be drawn between implicit awareness and 'phenomenal awareness', as defined by Lamme (2003). While a large amount of visual input reaches the point where conscious awareness could be achieved, Lamme (2003) suggests that this vulnerable visual experience is short-lived without accompanying attention. Conscious stimuli that are not attended to, and therefore cannot be explicitly reported, only achieve 'phenomenal awareness'. This is defined as a non-cognitive form of seeing, independent of attention, that can contain information about many items in a visual scene (Lamme, 2003, 2004). Similarities can therefore be drawn between phenomenal awareness and the

sense condition, where participants are unable to report all aspects of the change. In contrast, stimuli that benefit from the protective mechanism of attention enter 'access awareness', and can be explicitly reported. Within this framework, unconscious stimuli can never be reported, even if attended to, and would result in complete change blindness.

In the wider working memory literature, there are concepts known as activity-silent working memory (Stokes, 2015), as well as implicit working memory (Baumann et al., 2008). While the principle is the same, in that they seek to explain the fate of unnoticed stimuli, the paradigms used are often different. A key feature of change blindness paradigms is that all information is visible and available for the participant to see. Despite this, changes are sometimes missed. However, implicit working memory paradigms construct stimuli that are themselves below conscious awareness. For example, they may be displayed on the screen for a duration too fast for explicit report, or at a luminance/contrast too low to detect. Despite this, the presence of such stimuli can influence behaviour, by priming the participant or producing congruency effects. While this is still an interesting avenue of research, it is perhaps different to the implicit awareness that we discuss in the context of change blindness paradigms.

Conclusions

Given the large body of evidence discussed above, we believe that the *sense* condition is likely to occur during change detection paradigms. On some occasions, participants will suspect that a change has occurred, and the presence of that change will influence their behaviour. However, they will not always be able to explicitly report every detail about this change.

The main difficulty faced by researchers interested in this distinct level of awareness is how to create an operationalised definition and design an experimental paradigm that will capture *sensing* trials without additional confounds. As a consequence, the range of evidence available is based on varying defini-

tions of *sensing*, which up to this point have not be combined into a unified theory of this state of awareness.

Chapter 3 contains the manuscript for an EEG experiment titled 'An EEG study of detection without localisation in change blindness'. Many of the design choices were based on the conclusions drawn from the literature review in the previous chapter. The majority of the analysis is reported within the main text. However, additional figures and reports can be found in the appendix (A.2).

The manuscript was submitted to Experimental Brain Research on 7th January 2019, and a pre-print was also uploaded to BioRxiv, doi: 10.1101/513697. Corrections following peer-review were submitted on 8th May 2019 and the paper was accepted for publication on 12th June.

Chapter 3

An EEG study of detection without localisation in change blindness

3.1 Introduction

Change Blindness is a phenomenon in which changes to a visual scene are often missed (Rensink, 2004; Simons & Levin, 1997). To manipulate this in an experimental setting, the change blindness paradigm typically consists of two images displayed in quick succession that are interrupted by a blank screen or a distractor image. In some instances, the second image is identical to the first, and in others, some aspect will have changed. Participants are then asked to report if the trial contained a change or not. The complexity of these images varies across paradigms, ranging from coloured rectangles (Koivisto & Revonsuo, 2003) and coloured dots (Schankin & Wascher, 2007), to facial expressions (Eimer & Mazza, 2005), detailed visual scenes (Fernandez-Duque et al., 2003) and household objects (Busch et al., 2010). In all cases, although complete visual information is available, participants often fail to notice or identify changes.

Most versions of the change blindness paradigm ask participants to detect the presence of a change across two image presentations, meaning that trials can only be categorised as one of four types: hit (or *see* trials), miss (or *blind* trials), false alarm (FA), or correct rejection (CR), depending on whether the participant reports seeing a change. Several researchers have

challenged the traditional view that vision must always be accompanied by a complete conscious visual experience, or the activation of complete internal representation of what we see (Rensink, 2004; Fernandez-Duque & Thornton, 2000), and subsequently suggested the possibility of further trial divisions in the change blindness paradigm. In an early experiment reported by Rensink (2004), participants were asked to indicate when they 'thought' that something had changed in a flicker paradigm, and again when they were certain that they could see the change. In a flicker paradigm, the original image and changed image are presented sequentially until the participant is able to detect the change (Rensink et al., 1997). Trials in which these responses had a time difference greater than 1 second were labeled as trials with a 'significant duration of sensing', where the participant suspected a difference but was not confident in their perception of the change. Rensink (2004) termed the ability to detect a change without fully identifying it as *sensing*, suggesting that this condition is both phenomenologically and perceptually distinct to the traditionally reported *see* condition.

Several other researchers have explored the possibility of an awareness condition that lies somewhere between the traditional *see* and *blind* dichotomy (Fernandez-Duque et al., 2003; Laloyaux et al., 2006; Thornton & Fernandez-Duque, 2001; Galpin et al., 2008; Busch et al., 2009; Ball & Busch, 2015; Kimura et al., 2008; Hollingworth et al., 2001). For example, Fernandez-Duque et al. (2000)found that the location of a change could be identified above chance level even when participants did not report seeing the change itself (but see Mitroff et al. 2002 and Laloyaux et al. 2006 for a discussion of these results). Further, in Mitroff et al. (2004) participants were able to identify pre- and post-change object stimuli above chance level when they detected a change, as well as when they did not. The presence of a *sense* condition has therefore been suggested as evidence that change blindness may arise from a failure to compare two displays or images, rather than a failure to encode the visual information (Simons & Ambinder, 2005; Hollingworth et al., 2001).

Further, *sense* trials may occur when features of a changing object only reach a pre-attentive stage, and are not fully integrated at later stages of visual processing (Galpin et al., 2008; Busch et al., 2009).

Results from change blindness experiments using EEG appear to support this assertion. In previous EEG research, the trials types of *see* and *blind* are often distinguishable in an early visual attention component around 200-300 ms after the change onset at contralateral electrode sites, known as the N2pc (Luck & Hillyard, 1994; Schankin & Wascher, 2007). The presence of an N2pc reflects the allocation of attention towards an attended object (Luck & Ford, 1998), and the amplitude is increased for 'aware' stimuli (Schankin & Wascher, 2007). However, the N2pc also been found for 'unaware' stimuli in a masking paradigm, and therefore does not necessarily represent conscious awareness of a change (Woodman & Luck, 2003). It is therefore suggested that the N2pc, in the context of change blindness, reflects processing that is necessary, but not sufficient, to facilitate conscious change detection (Schankin & Wascher, 2007; Busch et al., 2009).

There is also evidence that the amplitude of early visual components, such as P1 and N1, may be dependent on the awareness level of the participant during a change detection task, given that larger peaks are identified for stimuli occurring in an attended location (Pourtois et al., 2006; Railo et al., 2011; Luck & Ford, 1998). However, not all change blindness EEG studies succeed in replicating this effect (Koivisto & Revonsuo, 2010).

In a similar time window to the P1/N1 complex (around 200 ms), the visual awareness negativity (VAN), typically occurring at posterior electrode sites, is thought to indicate detection of a stimulus and be dependent on spatial attention (Koivisto et al., 2008, 2009; Wilenius & Revonsuo, 2007). It has been suggested that the VAN is associated with phenomenal visual awareness (hence the name 'visual awareness negativity'), and is present even when successful identification of a changed object is not achieved (Lamme, 2004; Busch et al., 2009).

VAN is often followed by later positive ERP at posterior electrode sites called the late positivity (LP) (Koivisto et al., 2009). This overlaps with the P3 component, also peaking around 400 ms, and can also be referred to as such in the literature (Busch et al., 2009). In comparison to the VAN, the LP is associated with conscious aspects of task processing (Railo et al., 2011), and has been shown to correlate with participants' confidence in their responses (Eimer & Mazza, 2005).

Several EEG papers have also identified differences between *see*, *sense* and *blind* conditions. In a comparison between trials in which the participants were able to detect a change and identify the object of the change (*see*), and those where they could detect a change but not name it (*sense*), Busch et al. (2010) found an increase in amplitude of the VAN. The same effect was found in a later LP ERP at posterior electrodes. However, the N2pc peak was found only when participants could both detect and identify the change, and was not present when participants were change blind, or could not identify the object. The authors concluded that *seeing* a change is not simply a stronger version of *sensing* a change, as the N2pc can be found for *see* trials but not *sense* trials. This supports the hypothesis of Rensink (2004) that *seeing* and *sensing* may be facilitated by separate mechanisms. Other studies have also found differences in ERP amplitudes when comparing *see* and *sense* (Fernandez-Duque et al., 2003; Kimura et al., 2008; Busch, 2013; Ball & Busch, 2015), but the definition of *sense* trials varies across studies (Mitroff et al., 2002), leading to divergent results.

The main aim of the present study was to compare behavioural and ERP effects for trials in which participants could report the presence of a change but not localise it (*sense*), versus those in which participants could report and localise the change correctly (*localise*). Specifically, we divided the visual display into quadrants, and asked participants to select the quadrant in which the change occurred. Our *sense* condition therefore requires registration of the change, but not necessarily knowledge of its location (Mitroff et al., 2002).

Further, participants were asked to rate how confident they were in their responses at every trial, in order to distinguish between trial types (Galpin et al., 2008). We used a simple paradigm with an array of coloured squares (see figure 3.1).

As increased amplitudes in the N2pc and LP have previously been found in the *see* condition compared to the *blind* condition, we hypothesised that we would replicate these findings (Railo et al., 2011). Although modulation of P1 amplitudes have been reported in some change detection paradigms (Busch et al., 2009; Pourtois et al., 2006), others report no such effect (Eimer, 2000; Turatto, 2002; Niedeggen et al., 2001), so our hypothesis was not directed. When comparing *localise* versus *sense* trials, we hypothesised that we would find increased amplitudes in the VAN, LP, and N2pc for *localise* trials (Busch et al., 2010; Fernandez-Duque et al., 2003).

A further aim of the study was to identify if *sense* trials are behaviourally different to *blind* or *false alarm* trials, as others have suggested (Fernandez-Duque et al., 2003; Galpin et al., 2008), or whether they can be explained by explicit mechanisms (Mitroff et al., 2002). If the *sense* condition (where participants can detect but not localise a change in coloured square) can be explained by participant pressing the incorrect response when they did not see a change, then reaction times for *sense* trials should be similar to *blind* trials. Or, if *sense* can be explained by a liberal response criteria, such that participants report seeing a change despite not being sure, then uncertain *sense* trials should have similar reaction times to *false alarms*. By using EEG measures of neural activity, as well as additionally asking participants to rate their confidence at each trial (Galpin et al., 2008), we aimed to distinguish between these distinct types of awareness.

3.2 Materials and Methods

3.2.1 Participants

Twenty subjects (mean ± SD, age = 20 ± 5, 6 left handed, 2 male) with no history of psychiatric or neurological disorders participated in this EEG study. All had corrected-to-normal vision and were not colour blind (based on self report). The experiment was approved by the University of Reading ethics committee (UREC: 17/03), and was conducted in accordance with the Declaration of Helsinki (as of 2008). All participants gave informed consent to take part, including consent to share their anonymised data. Three participants were removed from the original sample size of 23 for having less than 200 usable trials after pre-processing (out of a maximum of 250 trials). Trials were classified as unusable if they contained muscle or eye-movement artifacts that could not be removed during pre-processing.

3.2.2 Stimuli and Presentation

Participants were presented with a change blindness task using Psychtoolbox (Kleiner et al., 2007), on a 1920 x 1080 LCD monitor with a 60 Hz refresh rate. Participants were seated comfortably on an armchair, at approximately 60cm away from the screen, alone, in a quiet room (Faraday cage) with constant dim light. They were asked to fixate on a central fixation cross and identify changes between consecutive displays of coloured squares. These were interrupted by a short fixation display to facilitate the change blindness phenomenon(see figure 3.1 for details on display durations). On change trials, one of the squares changed colour from the first to the second display. On no-change trials, the displays were identical. This was followed by two or three questions, depending on the participant's response to the first question. Each participant completed 5 blocks of 50 trials, leaving a total of 250 trials. Within these trials, two-thirds contained a change in coloured square (165 trials), and the rest contained no

change (85 trials).

Question 1 asked 'Did you see a change?' to which participants could respond 'yes' or 'no' using a keyboard. Question 2 asked participants to localise the change, based on a 2x2 grid from top left to bottom right. Question 3 asked how certain participants were of their responses, ranging from '1: Very Uncertain' to '4: Very Certain'. If participants responded 'no' change to question 1, they were moved straight to question 3. This decision was made as our hypotheses did not relate to 'implicit' change detection, as reported in Fernandez-Duque & Thornton (2000), and removing this question allowed for a greater number of trials within the same period of time. Participants were asked to respond within a limit of two seconds for each question, and trials with any response missing were not included in further analysis (3.6 $\pm$ 2.9). Participants made their response on a keyboard, using their index and middle finger from each hand.

Difficulty was modulated in real time by adding and removing two squares from the display, based on the assumption that more distractors increases task difficulty (Vogel et al., 2005). This was to prevent floor and ceiling performance during the task as a result of individual differences (Luck & Vogel, 2013), and optimise for performance rather than to establish specific individual thresholds. Performance over the previous two trials was used to update the current trial; two consecutive correct answers added two squares, two incorrect deducted two squares, and one correct and one incorrect resulted in no change. The decision to increase or decrease the number of squares was made using responses to the localisation question (Q 2), as we were specifically interested in controlling the number of *sense* and *localise* trials. The display was divided into a 6 x 6 grid of possible change locations, meaning that a maximum of 36 squares could be presented during each trial. The location of the change on each trial was random, but the change occurred an equal number of times on the left and right hemifield of the screen. The number of squares always changed by two, to balance the number on the left right hemifields of the screen, and all

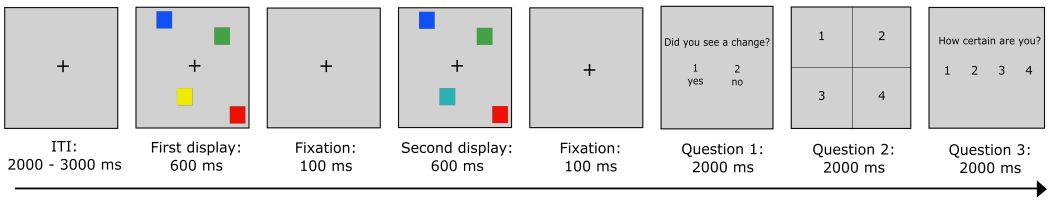

Figure 3.1: Illustration of the experimental paradigm. The number of squares presented varied from 2 to a maximum of 36. Question 1 asked 'Did you see a change?' to which participants could respond 'Yes' or 'No'. Question 2 asked participants to localise the change, based on a grid from top left to bottom right. Question 3 asked how certain participants were of their responses, ranging from '1: Very Uncertain' to '4: Very Certain'. If participants responded 'no change' to question 1, they were moved straight on to question 3.

participants began the experiment with two squares presented. Each block began with the number of squares presented on the last trial of the previous block. As the colour of the squares was not related to our main hypotheses, we used seven default MATLAB colours; blue, cyan, yellow, green, white, red, and magenta (MathWorks, Inc., version 2016b).

3.2.3 Behavioural Analysis

The trials in which a change occurred were divided into three conditions: *blind* (no change detection), *localise* (change detection and localisation), and *sense* (change detection without localisation). Trials in which no change occurred were divided into *correct rejection* (no change reported) and *false alarm* (change incorrectly reported). The number of false alarm trials was low, with a mean of 12.45 trials (range $= 2 - 33, SD = .65$), and therefore EEG analysis comparing *false alarm* to *sense* trials was not possible. The percentage of *false alarm* trials was calculated in relation to the the total number of no-change trials, whereas the percentage of *sense* trials was calculated in relation to the total number of change trials.

Detection accuracy for each participant was calculated based on the percentage of change trials in which they correctly detected a change. Localisation accuracy was calculated as the percentage of correctly detected changes where the localisation was also correct. We also recorded each participant's mean

and maximum difficulty scores, with the maximum referring the the highest number of squares that were displayed to them during the experiment.

D'prime was calculated as a measure of participant response bias. This was calculated using the equation $d = z(\text{hit rate}) - z(\text{false alarm rate})$ (Stanislaw & Todorov, 1999), and is defined as the difference between the means of signal and noise distributions, normalised by the variance. Response bias, or criterion, was also calculated, where $c = -0.5 * (z(\text{hit rate}) + z(\text{false alarm rate}))$ (Stanislaw & Todorov, 1999) . $c = 0$ indicates no response bias to either 'yes' or 'no' responses. $c > 0$ indicates a bias towards 'no' responses, with fewer hits and fewer false alarms. $c < 0$ indicates bias towards 'yes', with more hits but also more false alarms. We expected that participants would display a range of response strategies.

One problem faced in identifying a *sense* condition is that it is difficult to distinguish these trials from *false alarm* trials, or those where participants press the wrong response key (Simons & Ambinder, 2005; Mitroff et al., 2002). Rensink et al. (2004) found that reaction times for *sense* trials were shorter for change trials than no-change trials, meaning that participants were slower when they were simply making a false alarm. Galpin et al. (2008) also found greater certainty associated with *sensing* during change trials, compared to *false alarms*. We therefore compared reaction times across awareness conditions, as well as between levels of confidence. As trial numbers were low, 'very uncertain' and 'uncertain' responses were combined, and 'certain' and 'very certain' were combined. Each awareness condition therefore had two levels of certainty; for example, *localise certain* and *localise uncertain*.

3.2.4 EEG Data Acquisition

EEG data was recorded with a BrainVision EasyCap (Brain Products), with 64 passive electrodes including an IO channel, arranged according to the 10-10 layout. The reference electrode was placed at FCz and the ground at AFz. Impedance was kept below 10kΩ for all the EEG channels, and 5kΩ for the

IO channel. EEG signals were recorded using BrainVision Recorder (Brain Products, version 1.20) at a sampling rate of 5000 Hz.

3.2.5 EEG Pre-processing

Raw EEG data was pre-processed using BrainVision Analyzer (Brain Products, version 2.1). The data was first downsampled to 500 Hz to reduce computation time, then filtered with a high-pass filter of 0.01 Hz to remove low frequency drift (Butterworth, 2nd order). A low-pass filter of 50 Hz and a notch filter of 50 Hz were chosen to remove line noise. Independent component analysis (ICA) was used to remove eye movement artifacts (FastICA). Two components were removed for each participant; one corresponding to eye-blinks and the other to lateralised eye-movements.

Further analysis was completed using EEGLab (Delorme & Makeig, 2004). Trials were marked as outliers if any ERP value was greater than 3 standard deviations from the mean value of that ERP across all trials (using the MATLAB function 'isoutlier'). Note that we only searched for outliers in the electrodes used for analysis (P07, P08, Cz, Pz, and CPz). Trials marked as containing outliers were excluded from further analysis (3.25 trials per participant $\pm$ 2.46), as well as those where a response to any question was not made within the response time (3.60 trials per participant $\pm$ 2.94).

Segments were then taken from -200 to 7000 ms to include the whole trial, and baseline corrected using a mean of the data within -200ms to 0ms, where 0ms was the start of the first display of coloured squares (see figure 3.1). We chose the baseline period to be before the first display onset, rather than the second, as we were interested in visual ERPs that occurred in response to the both displays. It has also been suggested that ERPs in response to the first presentation of a stimuli are related to the subsequent perception of change (Pourtois et al., 2006).

3.2.6 EEG Analysis

To identify the peaks of the visually evoked potentials (P1 and N1), a grand average ERP was calculated across all conditions and participants, as advised in Luck & Gaspelin (2017), from electrodes P07 and P08. From here, the peaks of interest were determined by identifying the local maxima/minima of the expected peaks, using the peak detection function in BrainVision Analyzer. The mean value within a window around the peak was used instead of the peak value, as the mean is more robust against noise (Luck, 2014). A window of 40ms around the mean was chosen as the appropriate window for visual ERPs P1 and N1. In relation to the first display onset, the first P1 was identified at 122ms, and the first N1 at 212ms. In relation to the second display onset, the second P1 was identified at 114ms, and the second N1 at 222ms.

Based on previous literature (Busch et al., 2010; Tseng et al., 2012; Fernandez-Duque et al., 2003), the N2pc was defined as the mean within 200-400 ms after the second display at occipital electrodes PO7 and PO8. Over central parietal electrodes Cz, CPz and Pz, the VAN was defined within a window of 130-330 ms after the second display, and the LP within a window of 400-600ms. We used window sizes of 200 ms, defined a-priori, in an attempt to be conservative given the large variation within the literature.

To assess how differences between early visual components across detection conditions were reflected at each stimulus presentation, P1 and N1 amplitudes were compared in two separate 2x3 repeated measures ANOVAs, with display (first/second) and awareness (*blind*/*localise*/*sense*) as the independent variables. Differences across hemispheres in the N2pc were analysed with another 2x3 repeated measures ANOVA, with the independent variables of hemisphere (contralateral/ipsilateral) and awareness (*blind*/*localise*/*sense*). Amplitudes of the VAN and the LP were compared in two separate repeated measures ANOVAs with awareness (*blind*/*localise*/*sense*) as the independent variable. Where Mauchly's Test of Sphericity indicated that the assumption had been

violated, Greenhouse-Geisser correction was used. All post-hoc comparisons were two-tailed, and corrected for multiple comparisons using false discovery rate where $q = .05$ (Benjamini & Hochberg, 1995). Effect sizes are reported as partial eta squared for ANOVA, and repeated measures Hedge's g for t-tests (Lakens, 2013).

To determine if the visual ERPs (P1 and N1) varied as a function of the task difficulty (the number of squares presented per trial) we correlated the single-trial P1 and N1 amplitudes with the number of squares presented at each trial. To determine if the LP amplitude varied with participant confidence, as previously suggested (Eimer & Mazza, 2005), single-trial LP values were correlated with participant confidence ratings. For single-trial analysis, time courses were constructed for each participant from the single-trial values of each ERP, at each channel (7 ERPs, 64 channels, 20 participants). Note that midline electrodes were not included in N2pc analysis, as the N2pc values were calculated as the difference between ipsilateral and contralateral amplitudes, which by definition is not meaningful for electrodes on the midline. Each single-trial value was calculated as the mean amplitude within the pre-defined ERP window at each trial. These values were baseline corrected by subtracting the mean of the trial from which they were selected. P-values were corrected for multiple comparisons using false discovery rate where $q = .05$ (Benjamini & Hochberg, 1995).

3.3 Behavioural Results

3.3.1 Accuracy and Difficulty

Accuracy for question 1, in which participants had to identify a change, had a mean of 49% (range $= 32 - 73\%$, SD $= 13$). Accuracy for question 2, in which participants had to localise the change, had a mean of 70% ($55 - 87\%, 8$). The mean difficulty level given to each participant was 14 squares ($10-18, 3$), with the mean maximum difficulty experienced by each participant

at 26 squares $(20-36, 4)$. D'prime scores had a mean of .61 $(.74-1.64, .27)$. In a one-sample t-test, D'prime was significantly different from zero, suggesting that participants were able to distinguish between change and no-change trials $t(19) = 19.293, p < .001$. Two participants had a negative criterion, meaning that they had a response bias towards false alarms. All other participants had positive criterion, indicating a conservative response strategy $(.60 \pm .42)$.

Mean difficulty did not correlate with detection accuracy $(r = -.022, p = .928)$, location accuracy $(r = .136, p = .566)$, or d'prime $(r = -.229, p = .332)$, suggesting that the difficulty of the task did not influence task performance. Maximum difficulty also did not correlate with detection accuracy $(r = .067, p = .779)$, location accuracy $(r = -.077, p = .748)$, or d'prime $(r = -.148, p = 535)$.

3.3.2 Comparison of *sense* and *false alarm* trials

The percentage of *false alarm* trials $(14.64\% \pm 11.35)$ was lower than the percentage of *sense* trials $(30.31\% \pm 8.02)$ $t(19) = -7.107, p < .001, g_{rm} = 1.48$, suggesting that *sense* trials occurred more often than participants made false alarms. However, the percentage of false alarms was positively correlated with the percentage of *sense* trials $(r = .527, p = .017)$. Therefore, participants with a more liberal response strategy who made more false alarms, also had more *sense* trials.

Reaction times for *sense* and *false alarm* trials were compared, to determine if *sense* trials were different to trials where the participant incorrectly reported a change during a no change trial. Reaction times for all *sense* trials $(0.744 \pm 0.149$ s), regardless of certainty, were not significantly different to *false alarm* trials $(0.778 \pm 0.179$ s), $t(19) = -1.229, p = .234, g_{rm} = 0.193$. However, *sense certain* trials $(0.619 \pm 0.133$ s) were significantly faster than *false alarm* trials, $t(19) = -4.741, p < .001, g_{rm} = 0.939$. Therefore, when participants were certain that a change occurred, they responded more quickly than when they were simply making a false alarm.

Reaction times for *sense certain* trials (0.619 ± 0.133 s) were also significantly faster than *false alarm uncertain* trials (0.817 ± 0.211 s), $t(19) = -4.510, p < .001, g_{rm} = 1.081$. However, this may be explained by the general finding that, across all conditions, certain trials ($.628s \pm .142$) were faster than uncertain trials (0.849 ± 0.129 s), ($t(19) = -7.831, p < .001, g_{rm} = 1.563$)

3.3.3 Comparison of *sense* and *blind* trials

Reaction times for *sense* trials (0.744 ± 0.149 s) were not significantly different to *blind* trials (0.731 ± 0.176 s), $t(19) = -.285, p = 779, g_{rm} = .082$. However, reaction times for *sense uncertain* trials (0.801 ± 0.189 s) were significantly slower than *blind* trials, ($t(19) = 4.424, p < .001, g_{rm} = .373$). Therefore, on trials where the participant did not see the change (*blind*), they responded more quickly than when they suspected a change but could not provide additional information about it (*sense*).

Comparatively, reaction times for *sense certain* trials (0.619 ± 0.133 s) were significantly faster than *blind uncertain* trials (0.860 ± 0.231 s), ($t(19) = 4.424, p < .001, g_{rm} = 1.224$), which again may be explained by the fact that uncertain trials were slower over all conditions.

3.3.4 Comparison of *blind* trials and no-change trials

Out of the 20 participants included in the analysis, 15 were slower to respond when they were *blind* to the change, compared to no-change trials (75%). This difference in reaction times was not significant when comparing all no-change trials (0.704 ± 0.167 s) to *blind* trials (0.731 ± 0.176 s), ($t(19) = -2.084, p = .051, g_{rm} = .143$). However, *blind uncertain* trials (0.860 ± 0.231 s) were significantly slower than no-change trials (0.704 ± 0.167 s), ($t(19) = 3.637, p = .002, g_{rm} = .718$). Therefore, despite being *blind* to the change, the presence of a change in the display increased reaction times, particularly for trials where the participant was uncertain.

3.4 EEG Results

3.4.1 Single-trial Correlations

The purpose of this analysis was to check whether single-trial ERPs varied as a function of difficulty, i.e. the number of squares presented on the screen during each trial. After correcting for multiple comparisons using FDR correction ($q = .05$), no significant correlations were found.

The second analysis was to test whether single-trial ERPs varied with the confidence ratings of the participants. Several researchers have suggested that ERPs, particularly those in later time windows such as the LP, may be more influenced by participant confidence in their response than by the level of conscious awareness (Koivisto et al., 2005; Eimer, 2005). None of the tests were significant, with all $p > .34$. This result suggests that confidence ratings were not directly correlated with single-trial ERP amplitudes.

3.4.2 P1 and N1

Overall, no significant differences were found between the three awareness conditions for either the P1 or N1 (figure 3.2). For P1 amplitudes, the main effect of awareness was not significant, $F(1.473, 19) = 1.117, p = .338$, $\eta^2 = .056$. The main effect of display was also not significant, $F(1, 19) = .355, p = .558, \eta^2 = .018$, nor was the interaction between awareness and display, $F(1.80, 34.35) = .307, p = .305$, $\eta^2 = .060$.

For the N1, the main effect of awareness was not significant, $F(1.36, 19) = 3.534, p = .060$, $\eta^2 = .157$. The main effect of display was also not significant, $F(1, 19) = .209, p = .653, \eta^2 = .011$, nor was the interaction between awareness and display, $F(1.87, 35.61) = .377, p = .675$, $\eta^2 = .019$.

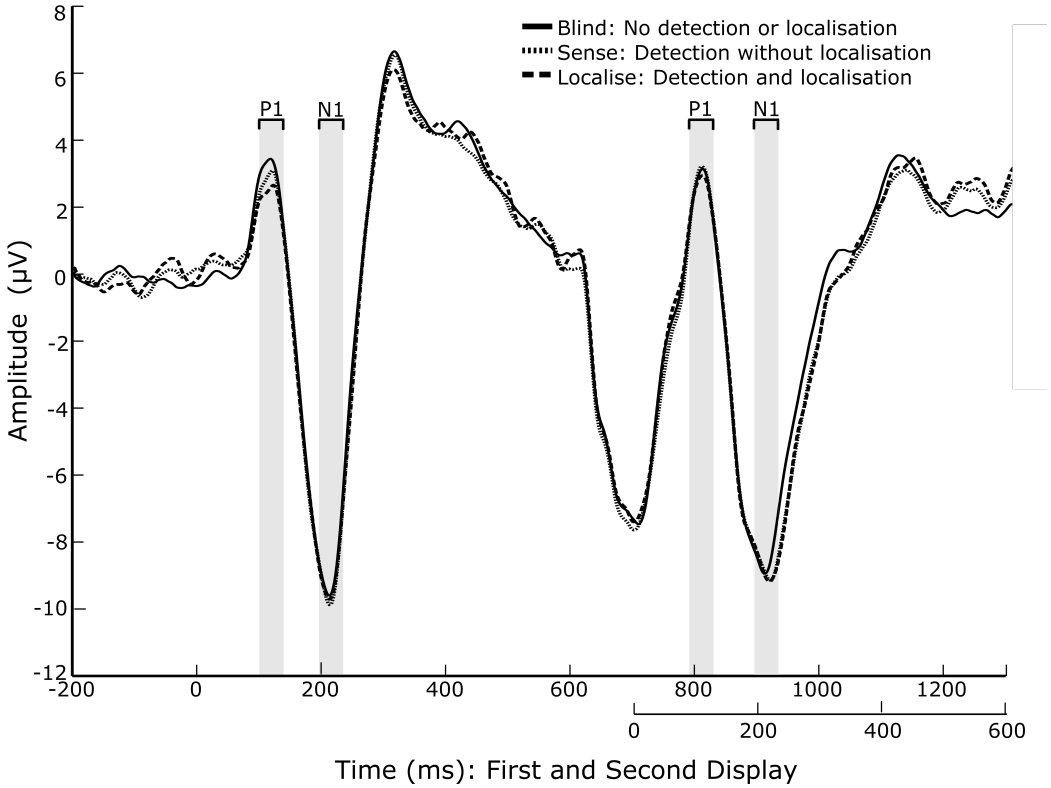

Figure 3.2: ERP plot showing the mean of electrodes PO7 and PO8, for each awareness condition. Condition means for the values within the shaded time windows were used for ERP analysis.

3.4.3 N2pc

In line with our hypothesis, there was a significant main effect of awareness on N2pc amplitudes, $F(2, 18) = 4.043, p = .026$, $\eta^2 = .175$ (figure 3.3). There was also a significant main effect of hemisphere, $F(1, 19) = 4.594, p = .045$, $\eta^2 = .195$, with a greater negativity in the contralateral hemisphere (-2.89 ± 3.97 μV) than the ipsilateral (-2.33 ± 4.26 μV). The interaction was not significant, $F(2, 18) = 1.048, p = .361$, $\eta^2 = .052$.

Post-hoc pairwise comparisons across awareness levels with a FDR corrected threshold of $p = 0.03$ showed that *blind* (-2.055 ± 1.23 μV) had a significantly smaller N2pc amplitude than localise *localise*, (-2.941 ± 1.80 μV), $t(19) = 2.340, p = .030, g_{rm} = .197$, and *sense* ($-2.847$ ± 1.19 μV), $t(19) = 2.525, p = .021, g_{rm} = .181$. However, *sense* and *localise* were not significantly different, $t(19) = -.283, p = .780, g_{rm} = .022$.

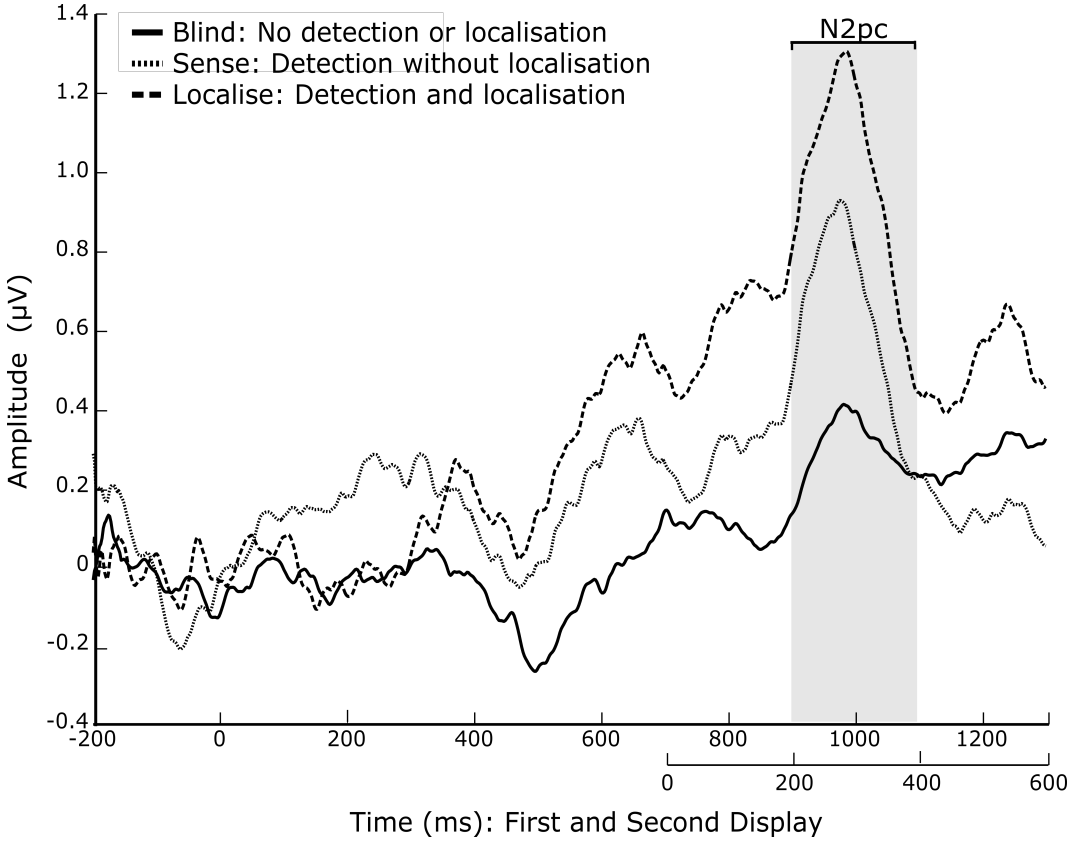

Figure 3.3: ERP plot showing the mean of electrodes PO7 and PO8, for each awareness condition. Asymmetry was calculated by subtracting contralateral from ipsilateral waveforms, and therefore a greater amplitude indicates a greater N2pc. Condition means for the values within the shaded time window (200-400 ms after the second display) were used for N2pc analysis.

3.4.4 Visual Awareness Negativity (VAN)

Confirming our hypothesis, there was a significant main effect of awareness on the VAN (figure 3.4), $F(1.374, 18) = 3.931, p = .046$, $\eta^2 = .171$. However, in post-hoc pairwise comparisons across awareness levels with a FDR corrected threshold of p = 0.04, *blind* ($-1.474 \pm 2.52\ \mu V$) was not significantly different to *localise* ($-2.167 \pm 3.09\ \mu V$), $t(19) = 2.158, p = .044, g_{rm} = .217$, or *sense* ($-1.961 \pm 1.92\ \mu V$), $t(19) = 1.950, p = .066, g_{rm} = .161$. *Localise* and *sense* were also not significantly different, $t(19) = 1.235, p = .232, g_{rm} = .062$.

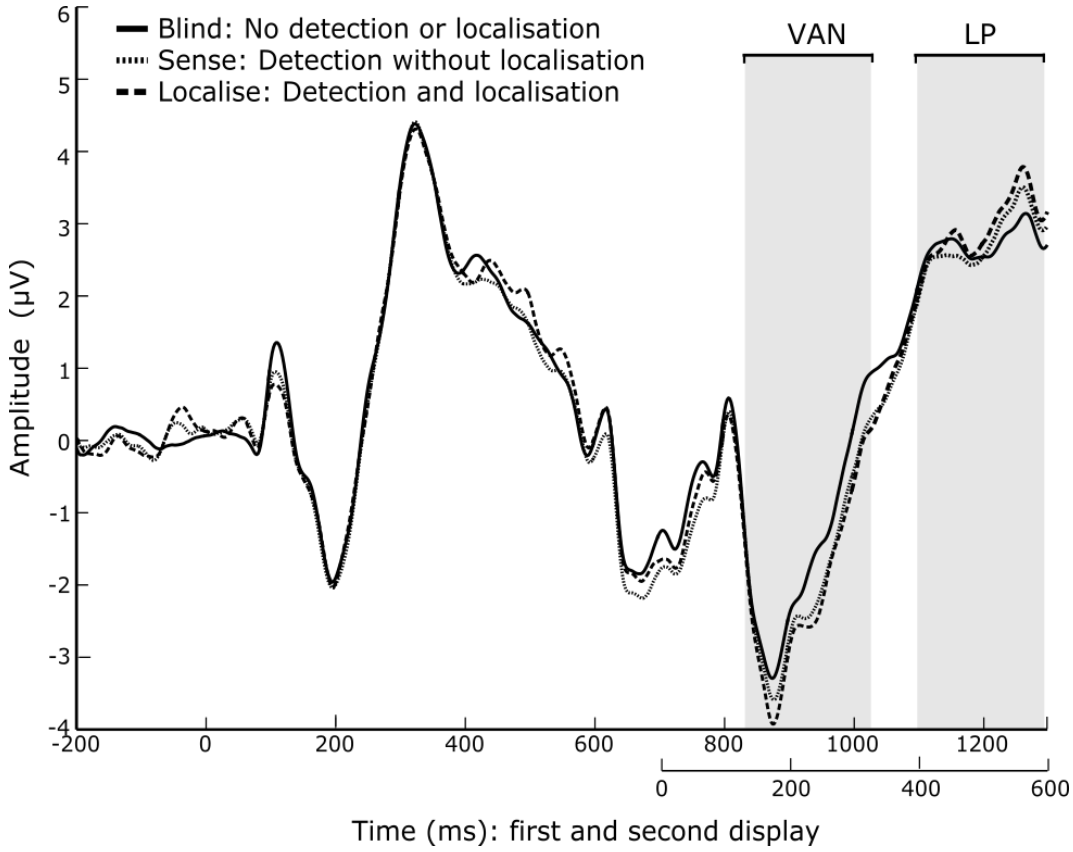

Figure 3.4: ERP plot showing a mean of electrodes Cz, CPz, and Pz, for each awareness condition. Condition means for the values within the shaded time window were used for ERP analysis. The first shaded area was used for the visual awareness negativity (130-330 ms after the second stimulus), and the second shaded area was used for the late positivity (400-600 ms) .

3.4.5 Late Positivity (LP)

In support of our hypothesis, there was a significant main effect of awareness on LP amplitudes (figure 3.4), $F(1.355, 8) = 7.000, p = .008$, $\eta^2 = .269$. In post-hoc pairwise comparisons across awareness levels with a FDR corrected threshold of p = .048, *blind* (2.931 $\pm$2.02 μV) was significantly smaller in amplitude to both *localise* (3.905 $\pm$2.53 μV), $t(19) = -3.094, p = .006, g_{rm} = .383$, and *sense* (3.591 $\pm$2.40 μV), $t(19) = -2.193, p = .041, g_{rm} = .275$. *Localise* was also significantly greater in amplitude than *sense*, $t(19) = 2.110, p = .048, g_{rm} = .118$.

3.5 Discussion

The main aim of this change blindness experiment was to distinguish between trials in which participants could both detect and localise a change in coloured square (*localise*), versus those in which they could only detect it (*sense*), or not detect it at all (*blind*). We found significant differences between *blind* trials and both *sense* and *localise* trials in the N2pc ERP. We also found that *sense* and *localise* were significantly different in the late LP window. Behaviourally, reaction time results allowed us to distinguish *sense* trials from *false alarm* and *blind* trials, when taking participant certainty into account. Overall, our results suggest that the *sense* condition may be distinguishable from the traditional *see* condition, and that utilising participant confidence is a valuable method to distinguish between types of awareness in change blindness.

3.5.1 EEG

Our results indicated a difference between *sense* and *localise* trials within the LP range, which were significantly different to each other, as well as to *blind*. An increased late positivity for change detected trials versus change *blind* trials is the most commonly reported finding within the EEG literature, and all of the papers considered in the review by Koivisto et al. (2010) report this finding. This may be due to the relatively large size of this ERP, peaking anywhere between 300 and 700ms after a change stimulus and across large time windows.

While the earlier negativity, VAN, is typically thought to be associated with phenomenal consciousness, the later positivity is linked to access consciousness and greater subject report ability. The repeated finding that the LP can be significantly reduced by specific stimuli, such as non-targets and repeated stimuli, suggests that it is not a direct correlate of visual awareness (Koivisto & Revonsuo, 2010). Instead, it is generally thought to reflect higher level or fully conscious aspects of task processing (Railo et al., 2011; Koivisto & Revonsuo, 2003). It has also be shown that the LP correlates with confidence

in participant responses (Eimer & Mazza, 2005). However, when correlating single trial LP amplitudes with confidence ratings, we did not find a significant effect.

The majority of change blindness papers listed by Koivisto (2010) reported enhanced negativity in the N1-N2 range (with the exception of Fernandez-Duque et al. 2003; Niedeggen et al. 2001). Busch et al. (2010) found that an N2pc was evoked only when the change was fully identified, and not in the *sense* or *blind* conditions. Based on this, they draw the conclusion that for *sense* trials, the change did not induce a shift in attention towards the location of the change, and therefore the features of the change were not available for further recognition. This is based on the assumption that the N2pc represents the allocation of attention towards the object of interest, which is supported by a number of previous studies (Luck & Ford, 1998).

Contrary to this, we found that both awareness conditions were significantly different to *blind* trials, indicating a shift in the allocation of attention for all identified changes, regardless of subsequent success/failure to localise. It may also be that *sense* trials elicited a shift in attention to the correct hemifield of change (and therefore subsequently an N2pc), but that it was not specific enough to determine whether the change occurred in the upper or lower field within that hemifield. Woodman and Luck (2003) also identified an N2pc for 'unaware' stimuli which were masked by object substitution masking, suggesting that the N2pc does not necessarily represent conscious awareness of changes (Woodman & Luck, 2003). It is suggested, however, that the amplitude is increased for 'aware' stimuli (Schankin & Wascher, 2007), which our findings support.

Other studies have reported a larger N2pc for more attention-demanding tasks (Luck & Hillyard, 1994). It was therefore a concern before analysis that *sense* trials would occur more often when the task was more difficult, and therefore that the N2pc would be larger for this condition as a result of uneven trial distribution. We found the opposite, however, with a smaller

N2pc in the *sense* condition compared to the *localise* condition. We also found no significant correlation with the number of *sense* trials and the difficulty of task given to the participant, suggesting that the trial distribution was even enough to avoid this confound.

Although there was a main effect of awareness within the VAN at central parietal sites, the corrected post-hoc tests were not significant, and only *localise* was significantly different to *blind* using an uncorrected threshold ($p = .044$). In comparison, (Busch et al., 2010) were able to identify a VAN for their *sense* condition, compared to *blind*. The VAN is thought to be dependent on spatial attention, and requires both the location and identity of an object to be stored such that it is available for conscious report (Koivisto et al., 2008). As participants were not able to identify the location of change in our *sense* condition, this may explain the lack of significant VAN ERP. In another study (Koivisto et al., 2008), VAN was found to be reduced when participants were asked to keep their eyes fixated at the centre of the screen. This was the case in this experiment, which may also have contributed to the lack of significant finding withing the VAN window.

Unlike previous findings from Pourtois et al., the amplitude of the P1 during the first stimuli display was not influenced by the level of awareness (Pourtois et al., 2006). In fact, no significant modulations of awareness were identified within either of the visual ERPs, P1 and N1, across either display, which fails to support previous findings that P1 amplitude during a visual display varies with attention (Wilenius & Revonsuo, 2007) and identification of changes (Mathewson et al., 2009). One possible reason for this could be that the number of squares varied across trials, unlike other experiments where the number was fixed (Pourtois et al., 2006), and therefore possibly driven by inter-individual differences in performance. However, when correlating single trial P1 and N1 amplitudes with difficulty across time, no significant correlations were found, after correcting for multiple comparisons. This suggests that the amount of squares presented during each trial had no direct influence on the amplitude

of the P1 and N1, and therefore that it did not create an obvious confound in the data.

In a review of the ERP correlates of visual awareness, Koivisto & Revonsuo (2010) list a number of change blindness EEG studies that also failed to detect modulation of an early P1 peak (Eimer, 2000; Koivisto & Revonsuo, 2003; Fernandez-Duque et al., 2003; Schankin & Wascher, 2007; Turatto, 2002; Niedeggen et al., 2001), compared to two studies which did (Busch et al., 2010; Pourtois et al., 2006). One criticism of the change blindness paradigm is that success relies on the participant paying attention to the first visual display, in order for the change to be integrated into the short term memory and the change detected (Simons & Levin, 1997). Attention levels, and perhaps ERPs, in response to the first display, may therefore have a large influence on the success of the following trial. We did not find any electrophysiological evidence for this occurring, as the amplitude of the P1 and N1 during the first visual display did not correlate with subsequent ERPs, or with performance. It may be, however, that this effect presented itself in a section of the EEG that was not analysed, or that the effect was not strong enough to detect across participants, some of whom may have been more vigilant than others.

The relationship between attention and awareness in change blindness is complex, and we did not attempt to explicitly dissociate the two in our paradigm. In fact, Koivisto & Revonsuo (2010) argue that the change blindness paradigm is not optimal for investigating the relationship between attention and awareness, as change detection is reliant on memory and therefore also on attention (given that attention facilitates working memory). It is very possible that attention directed towards a particular stimuli or region of the display increased the probability of detection, and enabled participants to localise the change successfully. As previously found, attention may be necessary but not sufficient for change detection; changes outside of the focus of attention are often missed, but change blindness can also occur for attended items (Levin & Simons, 1997; O'Regan et al., 2000; Chetverikov et al., 2018).

In an attempt to define the independent roles of attention and awareness, Lamme (2004) hypothesised that attention does not determine which stimuli reach a conscious state, but facilitates explicit report of these stimuli. While a large amount of visual input reaches the point where conscious awareness could be achieved, this vulnerable visual experience is short-lived without accompanying attention. Conscious stimuli that are not attended to, and therefore cannot be explicitly reported, only achieve 'phenomenal awareness'. This is defined as a non-cognitive form of seeing, independent of attention, that can contain information about many items in a visual scene (Lamme, 2003, 2004). Similarities can therefore be drawn between phenomenal awareness and the *sense* condition in our experiment, where participants could not successfully report the location of a change. In contrast, stimuli that benefit from the protective mechanism of attention enter 'access awareness', and can be explicitly reported. It should also be noted that, within this framework, unconscious stimuli can never be reported, even if attended to.

3.5.2 Behavioural

One explanation for the presence of a *sense* condition in change blindness is that it reflects a liberal response criteria, such that participants report seeing a change even though they were not certain that it occurred (Simons & Ambinder, 2005). In other words, they make a *'false alarm'* during change trials. If this is the case, then these trials may be similar in number to *false alarm* trials, where participants incorrectly report a change for identical displays where they could not have seen a change. We found that participants had a significantly higher percentage of *sense* trials than *false alarm* trials, suggesting that *sense* trials occurred more often. This finding cannot be explained by the fact that more trials contained a change, as the percentages were calculated in relation to the total number of change/no-change trials, respectively.

However, we also found a significant correlation between the percentage of *sense* and *false alarm* trials, suggesting that participants with a more liberal

response strategy were more likely to report the presence of a change when they were not completely sure where the change occurred. To further compare *sense* and *false alarm* trials, we also examined reaction times. Although all *sense* trials combined were not significantly different to *false alarms*, *sense certain* trials were significantly faster. Therefore, *sense* trials where the participant was certain that they saw something change may be distinguishable from simple *false alarms*.

Another explanation for the *sense* condition is that it contains trials for which the participant mistakenly reported a change, even though they were not aware of it. In this case, reaction times for *sense* trials should be similar to those for *blind* trials, particularly those where participants were uncertain of their responses. We found that *sense uncertain* trials were significantly slower than *blind* trials, suggesting that participants took longer to respond to trials where they suspected that something had changed, but were uncertain.

Previous studies have also reported that participants responded 'no change' more quickly for no-change trials, compared to change trials (Williams & Simons, 2000; Mitroff et al., 2002). The participant's response is the same in both trial types, but the presence of a change is different. This suggests that even when they fail to detect the change in a change trial, they take longer to respond. We therefore compared reaction times for no-change trials and *blind* trials. Out of the 20 participants, 15 were slower to respond when they were blind to the change, compared to no-change trials (75%), which is higher than the 68% reported by Williams & Simons (2000). Although no significant differences were found between all *blind* and no-change trials, *blind uncertain* trials were significantly slower. It is possible that in *blind certain* trials, no information about the change is registered by the participant, and therefore reaction times are similar to no-change trials. However, in *blind uncertain* trials, some information may be available to the participant, leading to slower reaction times, but not enough for them to be confident to report the change.

As the average accuracy for question 1 (yes/no) was roughly 50% across

participants, change trials were fairly equally divided into *see* (all trials where a change was correctly identified) and *blind* conditions. Within the *see* trials, accuracy for question 2 ('where did the change occur?') was roughly 70%, leaving more trials in the *localise* condition than the *sense* condition.

Unfortunately, the number of false alarm trials was low, meaning that a comparison of false alarms trials in the EEG data was not possible. Within the *sense* trials, there was also a low number of 'certain' trials, meaning that dividing the awareness conditions into certain/uncertain for EEG analysis was also not possible. Future experiments could focus on obtaining higher trial numbers, which would hopefully facilitate this analysis. However, the very nature of the *sense* condition means that participants are unlikely to be 'certain' during many of the trials.

We defined the difficulty of the task as the number of squares that were presented to the participant during each trial. Participants ranged in the difficulty within which they could perform the task with similar accuracy. The maximum difficulty ranged from 10 to 36, with only one participant reaching the highest possible level. The fact that the difficulty measures, such as maximum difficulty and mean difficulty, were not correlated with accuracy or d'prime, suggests that the difficulty modulation managed to control for individual differences in ability across participants. However, despite the difficulty modulation, the range of accuracy demonstrated by the participants was large (32% - 73%). Future studies could benefit from a more sophisticated measure of trial-by-trial adaptation, to further balance the number trials within each condition and participant.

3.5.3 Conclusions

Overall, the main aim of this experiment was to identify neural differences between full and partial awareness of colour changes, while controlling for individual differences in performance. Behaviourally, reaction time results allowed us to distinguish *sense* trials from *false alarm* and *blind* trials, when

taking participant certainty into account. For EEG data in the N2pc range, *localise* and *sense* were both significantly different to *blind* trials, but not significantly different from each other. In comparison, within the LP range, all conditions were significantly different, indicating that the difference between levels of awareness was represented in this late potential. Overall, our results suggest that the *sense* condition may be distinguishable from the traditional *see* condition, and that utilising participant confidence is a valuable method to distinguish between levels of awareness in change blindness.

Chapter 4 contains the manuscript for a combined EEG-fMRI experiment titled 'Simultaneous EEG, fMRI, and behavioural measures of detection without localisation in change blindness'. As far as we are aware, this is the first fMRI study investigating the *sense* condition in change blindness. It is therefore also the first EEG-fMRI study.

The majority of the analysis is reported within the main text. However, additional figures and reports can be found in the appendix (A.4).

Chapter 4

Simultaneous EEG, fMRI, and behavioural measures of detection without localisation in change blindness

4.1 Introduction

It is common for us to overestimate the amount of information that we can process and store about the world around us. Although we may assume that we would notice a cyclist entering the path of our car, or if a building on our street changed in colour, in reality we very often miss these occurrences (Simons, 2000; Simons & Levin, 1997). The failure to detect changes between visual scenes is known as change blindness, and is used as evidence to suggest that our internal representation of the outside world is not as complete as once thought (Rensink, 2004; Noe et al., 2000). When changes to an image are disrupted in some way, for example by a distractor image or a visual saccade, we cannot use visual transients (or motion) to detect them, and are often blind to the difference (Rensink et al., 1997; Kanai & Verstraten, 2004).

It was previously assumed that if we are blind to a change then we cannot provide any information about it, and that the change should not influence our behaviour in any way. Blindness to changes is thought to result from a lack of detailed representation about the pre- and post-change scenes, or an inability to successfully compare the two (Simons, 2000). If this is the case, then our knowledge when we are blind to changes should be equivalent to

that when there is no change at all. Anecdotally, this does not align with the experience of observers in a change blindness experiment; it is common for them to remark that they suspected something had changed, but that they were not sure about its nature or location. This experience appears to be phenomenologically different from complete change blindness, but whether this difference is represented in behavioural and neuroimaging data is yet to be confirmed.

In an early experiment, Rensink (2004) suggested the presence of a *sense* condition, in which observers could detect a change without fully identifying it. Observers were asked to indicate when they 'thought' that something had changed, and then again when they were certain of it. Trials in which the time between these two responses was greater than 1 second were labelled as trials with a significant duration of *sensing*. He argued that this condition is both phenomenologically and perceptually distinct from the traditionally reported *see* condition in which participants are fully aware of what change occurred. The presence of a *sense* condition has been suggested as evidence that change blindness may arise from a failure to compare two displays or images, rather than a failure to encode the visual information (Simons et al., 2005a; Hollingworth et al., 2001). Furthermore, *sense* trials may occur when features of a changing object only reach a pre-attentive stage, and are not fully integrated at later stages of visual processing (Galpin et al., 2008; Busch et al., 2009).

However, it may be that participants who *sense* a change are simply applying a more liberal response criterion when completing the task, and in fact are not really aware of the change (Simons et al., 2005a). Similarly, implicit awareness of changes could also be explained by explicit mechanisms such as guessing or a process of elimination (Mitroff et al., 2002). The feasibility of gauging an implicit level of awareness using explicit measures, such as participants' responses, has also been questioned. If we are capable of processing information that is not complete enough for explicit report, then it is possible that we

will underestimate the ability of the visual system to detect changes if we only rely on self-report measures (Lamme, 2004; Fernandez-Duque & Thornton, 2000). In response, several researchers turned to neuroimaging methods, such as electroencephalography (EEG), to identify neural correlates of the *sense* condition (Busch et al., 2010; Fernandez-Duque & Thornton, 2003; Kimura et al., 2008; Lyyra et al., 2012). Despite the current lack of a clear definition for the *sense* conditions, neuroimaging data provides a valuable contribution to the debate over whether this condition really is distinguishable from other levels of awareness.

In a previous EEG experiment, we found that *sense* trials had a larger N2pc over occipital electrodes than when observers were *blind* to the change (Scrivener et al., 2019). The presence of an N2pc reflects the allocation of attention towards an attended object (Luck & Ford, 1998), and the amplitude is increased when participants are aware of the stimuli (Schankin & Wascher, 2007). Although the N2pc is thought to reflect the allocation of attention towards an object (Luck & Ford, 1998), it has also been found for unseen stimuli in a masking paradigm (Woodman & Luck, 2003), and therefore does not necessarily represent explicit awareness of a change. We therefore concluded that the presence of an N2pc for both *sense* and *localise* conditions indicated a shift in attention towards the hemisphere of the change, but that this shift in attention was not sufficient to facilitate correct localisation in *sense* trials.

In contrast, the late positivity (LP) is thought to reflect conscious aspects of task processing (Railo et al., 2011), and has been shown to correlate with participants' confidence in their responses. Within the LP, we found that *sense* was greater in amplitude than both *blind* trials, and those in which participants were completely aware of a change (they could both detect and localise the change). These results, among others (Busch et al., 2010; Fernandez-Duque & Thornton, 2003; Kimura et al., 2008; Lyyra et al., 2012), suggest that *sensing* and *seeing* may rely on two separate mechanisms. However, it is not clear whether these mechanisms rely on distinct or overlapping networks of brain

activation (Rensink, 2004; Busch et al., 2009; Howe & Webb, 2014).

To our knowledge, the *sense* condition has not been explored using fMRI. However, several studies have compared the traditional *see* and *blind* conditions. In general, ventral areas of the brain are thought to provide the substance for visual awareness, whereas frontal and parietal activation facilitate its conscious experience (Kanwisher, 2001; Dehaene et al., 2006; Pins, 2003). More specifically, in a comparison between trials in which observers correctly detected changes and change blind trials, awareness was associated with activation in the bilateral superior parietal lobule, right middle frontal gyrus, and fusiform gyrus (Beck et al., 2001). This suggests the combined involvement of regions from both dorsal parietal regions and category-selective regions of the ventral stream, and therefore that an interaction between these two networks may facilitate conscious detection of changes.

However, the difference in attentional state driving the accuracy differences between correct versus incorrect trials could also modulate brain activation in this way (Pessoa, 2004; Ungerleider, 2000; Ress et al., 2000). An alternate comparison between correctly identified changes and no change trials found activation in a network of frontal and parietal regions, as well as the pulnivar, cerebellum, and inferior frontal gyrus (Pessoa, 2004). A similar pattern was identified for false alarm trials, where participants reported a change when no change occurred, suggesting that activity was related to the participants' perception of the change rather than properties of the visual stimulus. Few regions were specifically activated when participants were blind to the change.

Given the lack of fMRI evidence for the *sense* condition, the main aim of this study was to investigate the existence and nature of the *sense* condition in the change blindness paradigm, using combined EEG-fMRI and behavioural measures. While a range of evidence posits a distinction between *sense* and *blind* conditions in EEG data, not such distinction has been made for the *sense* condition using fMRI. Further, we aimed to improve the respective temporal and spatial resolution of EEG and fMRI by measuring them simultaneously.

We therefore aimed to identify brain regions with BOLD activity that co-varied with activity in the EEG data, to detect possible sources or networks associated with awareness of changes.

We hypothesised that we would replicate our results from a previous EEG experiment, identical in nature except for the recording environment and inter-trial intervals (Scrivener et al., 2019). Based on the EEG and behavioural evidence, we also hypothesised that the *sense* condition would result in greater visual and parietal activation than that resulting from the *blind* conditions, but not to the extent associated with full awareness (Dehaene et al., 2006; Koivisto & Revonsuo, 2003; Beck et al., 2001; Pins, 2003).

4.2 Materials and Methods

All materials and analysis methods were pre-registered in an open document on the Open Science Framework (uploaded 27/06/18), where the data and analysis scripts for this project can also be found (https://doi.org/10.17605/OSF.IO/W6BH3).

4.2.1 Participants

Twenty one right-handed subjects (mean ± SD, age = 21 ± 3.6, 6 male) with no history of psychiatric or neurological disorders participated in this EEG-fMRI study. All had corrected-to-normal vision and were not colour blind (based on self report). The experiment was approved by the University of Reading ethics committee (UREC: 16/120), and was conducted in accordance with the Declaration of Helsinki (as of 2008). All participants gave informed consent to take part, including consent to share their anonymised data. One participant was removed from the EEG analysis due to failure to remove MRI related artifacts, leaving N=20. An additional participant was removed from the behavioural analysis as the difficulty modulation did not function correctly, leaving N=19. A total of five participants were removed from the fMRI and EEG-fMRI analysis for having motion greater than one voxel size in the fMRI

data, leaving N=16.

4.2.2 Stimuli and presentation

A change blindness task was presented using Psychtoolbox (Kleiner et al., 2007), on a 1920 x 1080 LCD monitor with a 60 Hz refresh rate. The paradigm was displayed on a screen displayed approximately 47cm away from the centre of the scanner bore. This was viewed by the participant through a mirror mounted onto the coil, at approximately 12cm from the participant's eyes. In their left hand, the participant held an alarm ball, and in their right they held a 4 key button box. They had to use all of the 4 keys to respond to the task. Participants were asked to fixate on a central fixation cross and identify changes between consecutive displays of coloured squares. These were interrupted by a short fixation display to facilitate the change blindness phenomenon (see figure A.6 for details on display duration). On change trials, one of the squares changed colour from the first to the second display. On no-change trials, the displays were identical. This was followed by two or three questions, depending on the participant's response to the first question.

Question 1 asked 'Did you see a change?' to which participants could respond 'yes' or 'no'. Question 2 asked participants to localise the change, based on a 2x2 grid from top left to bottom right. Question 3 asked how certain participants were of their responses, ranging from '1: Very Uncertain' to '4: Very Certain'. If participants responded 'no' change to question 1, they were moved straight to question 3. This decision was made as our hypotheses did not relate to 'implicit' change detection, as reported in Fernandez-Duque & Thornton (2000), and removing this question allowed for a greater number of trials within the same period of time. Participants were asked to respond within a limit of two seconds for each question, and trials with any response missing were not included in further analysis.

This study had a within-subjects repeated measures design, and each participant completed 5 blocks of 50 trials, meaning a total of 250 trials. Of these

250 trials, 165 contained a change in coloured square, and the remaining trials contain no change. The ratio was not kept at 50/50, as the trials containing the change were of most interest for analysis. However, after the experiment participants were asked to report the percentage of trials that they believed contained a change. After each block of 50 trials, the participants were presented with a break screen, advising them to take a break. The participant was able to continue the experiment at their discretion by pressing any button on the button box. Before beginning the main task, participants were given a short block of 10 trials in which to practice responding to the paradigm with the button box. The data from this practice block was not analysed.

Difficulty was modulated in real time by adding and removing two squares from the display, based on the assumption that more distractors increases task difficulty (Vogel et al., 2005). This was to prevent floor and ceiling performance during the task as a result of individual differences (Luck & Vogel, 2013), and optimise for performance rather than to establish specific individual thresholds. Performance over the previous two trials was used to update the current trial; two consecutive correct answers added two squares, two incorrect deducted two squares, and one correct and one incorrect resulted in no change. The decision to increase or decrease the number of squares was made using responses to the localisation question (Q 2), as we were specifically interested in controlling the number of *sense* and *localise* trials. The number of squares always changed by two, to balance the number on the left right hemifields of the screen. The location of the change on each trial was random, but the change occurred an equal number of times on the left and right hemifield of the screen. The display was divided into 36 even sections, with 6 in each quadrant, within which the squares could appear. As the colour of the squares was not related to our main hypotheses, we used seven default MATLAB colours; blue, cyan, yellow, green, white, red, and magenta (MathWorks, Inc., version 2016b).

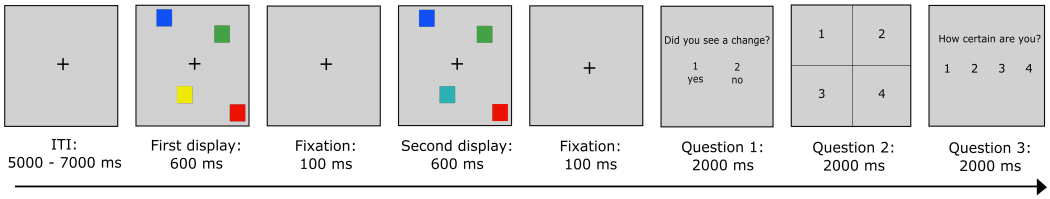

Figure 4.1: Illustration of the experimental paradigm. The number of squares presented varied from 2 to a maximum of 36. Question 1 asked 'Did you see a change?' to which participants could respond 'Yes' or 'No'. Question 2 asked participants to localise the change, based on a grid from top left to bottom right. Question 3 asked how certain participants were of their responses, ranging from '1: Very Uncertain' to '4: Very Certain'. If participants responded 'no change' to question 1, they were moved straight on to question 3.

4.2.3 Behavioural Analysis

The trials in which a change occurred were divided into three conditions: *blind* (no change detection), *localise* (change detection and localisation), and *sense* (change detection without localisation). Trials in which no change occurred were divided into *correct rejection* (no change reported) and *false alarm* (change incorrectly reported). The number of false alarm trials was low, with a mean of 10 trials (range $= 1 - 28$, $SD = 7.34$), and therefore EEG analysis comparing *false alarm* to *sense* trials was not possible. The percentage of *false alarm* trials was calculated in relation to the the total number of no-change trials, whereas the percentage of *sense* trials was calculated in relation to the total number of change trials.

Detection accuracy for each participant was calculated based on the percentage of change trials in which they correctly detected a change. Localisation accuracy was calculated as the percentage of correctly detected changes where the localisation was also correct. We also recorded each participant's mean and maximum difficulty scores, with the maximum referring to the highest number of squares that were displayed to them during the experiment. Behavioural analysis was completed in JASP 2018 (version 0.8.2.0).

D'prime was calculated as a measure of participant response bias. This was calculated using the equation $d = z(\text{hit rate}) - z(\text{false alarm rate})$ (Stanislaw

& Todorov, 1999), and is defined as the difference between the means of signal and noise distributions, normalised by the variance. Response bias, or criterion, was also calculated, where $c = -0.5 * (z(\text{hit rate}) + z(\text{false alarm rate}))$ (Stanislaw & Todorov, 1999) . $c = 0$ indicates no response bias to either 'yes' or 'no' responses. $c > 0$ indicates a bias towards 'no' responses, with fewer hits and fewer false alarms. $c < 0$ indicates bias towards 'yes', with more hits but also more false alarms. We expected that participants would display a range of response strategies.

One problem faced in identifying a *sense* condition is that it is difficult to distinguish these trials from *false alarm* trials, or those where participants press the wrong response key (Simons & Ambinder, 2005; Mitroff et al., 2002). Rensink et al. (2004) found that reaction times for *sense* trials were shorter for change trials than no-change trials, meaning that participants were slower when they were simply making a false alarm. Galpin et al. (2008) also found greater certainty associated with *sensing* during change trials, compared to *false alarms*. We therefore compared reaction times across awareness conditions, as well as between levels of confidence. As trial numbers were low, 'very uncertain' and 'uncertain' responses were combined, and 'certain' and 'very certain' were combined. Each awareness condition therefore had two levels of certainty; for example, *localise certain* and *localise uncertain*.

To establish if the location of the change influenced the likelihood that it was detected, we conducted two chi-square analyses. The first analysis divided the 6 x 6 grid of possible change locations into two conditions, outside and central. Changes occurring in any of the 20 outermost locations were considered to be outside changes, and the 16 central locations were considered to be central. We ran a 2 x 3 chi-square with the independent variables of location (outside/central) and awareness (*blind*/*localise*/*sense*), and the dependent variable as the frequency of trials within each condition, across participants. In the second analysis, we instead compared the side of the display in which the change occurred, resulting in a 2 x 3 chi-square for hemisphere (left/right) and

awareness (*blind/localise/sense*). Additional behavioural analysis and results, in reference to the pre-registration, can be found in appendix A.4.

4.2.4 EEG data acquisition

EEG data was recorded with an MRI-compatible cap equipped with carbon-wired Ag/AgCL electrodes (Braincap MR) from 64 scalp positions according to the international 10-10 system. The reference electrode was placed at FCz and the ground at AFz. An additional ECG electrode was positioned on the back to measure heart rate. An MRI-compatible EEG amplifier was used (BrainAmp MR, Brain Products) with a sampling rate of 5000Hz. Impedance was kept below 10kΩ for EEG channels and 5kΩ for the ECG. EEG recordings were performed with Brain Vision Recorder Software (Brain Products) and timings kept constant using a BrainProducts SyncBox to synchronise EEG with the MRI system clock.

4.2.5 EEG pre-processing

Raw EEG data was pre-processed using Brain Vision Analyzer version 2.1 (Brain Products). Correction for the MR gradient artifact was performed using a baseline corrected sliding average of MR volumes (Allen et al., 2000). Removal of cardioballistic artfacts involved the subtraction of heartbeat artifacts on a second by second basis, using a sliding average of 21 (Allen et al., 1998). The delay was detected using the CBC detection solution, individually for each subject. Peaks were detected semi-automatically, with a manual check of the algorithm's estimations. ICA was then used to remove further BCG residual artifacts (range: 1 - 4 additional ICs removed per participant). As outlined in (Debener, 2005), the presence of visual P1 and N1 peaks in the averaged data after pre-processing was used as an indication of the successful removal of artifacts.

The data was downsampled to 500 Hz to reduce computation time and then filtered with a high-pass filter of 0.1 Hz to remove low frequency drift

(Butterworth, 2nd order). A low-pass filter of 50 Hz and a notch filter of 50 Hz were chosen to remove line noise. Independent component analysis (ICA) was used to remove eye movement artifacts (FastICA). Two components were removed for each participant; one corresponding to eye-blinks and the other to lateralised eye-movements.

Further analysis was completed using EEGLab (Delorme & Makeig, 2004). Trials were marked as outliers if any ERP value was greater than 3 standard deviations from the mean value of that ERP across all trials (using the MATLAB function 'isoutlier'). Note that we only searched for outliers in the electrodes used for analysis (P07, P08, Cz, Pz, and CPz). Trials marked as containing outliers were excluded from further analysis ($M = 7$ trials, $SD = 12.98$), as well as those where a response to any question was not made within the response time ($M = 2$ trials, $SD = 2.79$).

Segments were then taken from -200 to 7000 ms to include the whole trial, and baseline corrected using a mean of the data within -200ms to 0ms, where 0ms was the start of the first display of coloured squares (see figure A.6). We chose the baseline period to be before the first display onset, rather than the second, as we were interested in visual ERPs that occurred in response to the both displays. It has also been suggested that ERPs in response to the first presentation of stimuli are related to the subsequent perception of change (Pourtois et al., 2006).

4.2.6 EEG Analysis

To identify the peaks of the visually evoked potentials (P1 and N1), a grand average ERP was calculated across all conditions and participants, as advised in Luck & Gaspelin (2017), from electrodes P07 and P08. From here, the peaks of interest were determined by identifying the local maxima/minima of the expected peaks, using the peak detection function in BrainVision Analyzer. The mean value within a window around the peak was used instead of the peak value, as the mean is more robust against noise (Luck, 2014). A window of

40ms around the mean was chosen as the appropriate window for visual ERPs P1 and N1. In relation to the first display onset, the first P1 was identified at 124ms, and the first N1 at 142ms. In relation to the second display onset, the second P1 was identified at 108ms, and the second N1 at 168ms.

Based on previous literature (Busch et al., 2010; Tseng et al., 2012; Fernandez-Duque et al., 2003), the N2pc was defined as the mean within 200-400 ms after the second display at occipital electrodes PO7 and PO8. Over central parietal electrodes Cz, CPz and Pz, the VAN was defined within a window of 130-330 ms after the second display, and the LP within a window of 400-600ms. We used window sizes of 200 ms, defined a-priori, in an attempt to be conservative given the large variation within the literature.

To assess how differences between early visual components across detection conditions were reflected at each stimulus presentation, P1 and N1 amplitudes were compared in two separate 2x3 repeated measures ANOVAs, with display (first/second) and awareness (*blind*/*localise*/*sense*) as the independent variables. Differences across hemispheres in the N2pc were analysed with another 2x3 repeated measures ANOVA, with the independent variables of hemisphere (contralateral/ipsilateral) and awareness (*blind*/*localise*/*sense*). Amplitudes of the VAN and the LP were compared in two separate repeated measures ANOVAs with awareness (*blind*/*localise*/*sense*) as the independent variable. Where Mauchly's Test of Sphericity indicated that the assumption had been violated, Greenhouse-Geisser correction was used. All post-hoc comparisons were two-tailed, and corrected for multiple comparisons using false discovery rate where $q = .05$ (Benjamini & Hochberg, 1995). Effect sizes are reported as partial eta squared for ANOVA, and repeated measures Hedge's g for t-tests (Lakens, 2013).

4.2.7 Single-trial EEG Analysis

For each ERP time window, single-trial values were calculated as the mean amplitude within the predefined window for that peak. These values were

then baseline corrected by subtracting the mean amplitude across the trial from which they were taken. Outliers were identified as trials where the amplitude was more than 3 standard deviations away from the mean amplitude for that ERP. As large artifacts can raise the mean amplitude, we added the additional classification of outliers at values +/- 30 μV. These outlier values were replaced by the mean value across all other trials, as outlined in Bénar et al. (2007). The subsequent parametric regressors created for the ERPs were mean-centred, meaning that outlier trials had no influence on the fit of the model.

4.2.8 fMRI recording

MRI data was acquired using a 3.0-T whole-body MRI scanner (Prisma, Siemens) and a 64 channel coil for functional imaging. Interleaved slices were recorded using a 2D echo planar imaging (EPI) sequence [repetition time (TR) 1630ms; echo time (TE) 30ms; flip angle 90°; voxel size 3mm x 3mm; thickness 3mm; encoding direction A to P; distance factor 20%; FOV read 192mm; number of slices 30; transveral orientation]. Three dummy scans were acquired at the beginning of each block. As well as the functional scans, an anatomical scan of the entire brain was acquired [3D MPRAGE; saggital; TE 2.37ms; TR 1800ms; flip angle 8°; voxel size 0.98mm x 0.98mm; FOV read 250mm; slice thickness 0.85mm; slices per slab 208; ascending acquisition; phase encoding direction A to P].

4.2.9 fMRI Pre-processing

MRI images were pre-processed using the procedure recommended in SPM12 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK). Functional images were first re-aligned per experimental block. These were registered to the mean image with a 6th degree spline interpolation. Following this was co-registration of the structural image to aligned functional images, segmentation of white and gray matter, normalisation of functional images using the deformation field created during segmentation, and normali-

sation of the functional to structural. The resulting data was smoothed with a 4-mm full-width-half-maximum Gaussian Kernel, and a high-pass filter with a cut off period of 128 s was applied. The registration of images was checked visually at each stage.

4.2.10 fMRI Analysis

At first level analysis, general linear models (GLM) with event-related designs were conducted in SPM12, to identify voxels activated in response to trial type (*blind/ localise/ sense/ false alarm/ correct rejection*). Regressors were created for each trial type by convolving the stimulus onset times with the hemodynamic response function (HRF) across all blocks (Friston et al., 1994). Each block was modelled with a separate set of regressors, with 6 motion regressors added as nuisance variables. For each participant we ran the following contrasts during first-level analysis; *sense* > *blind*, *localise* > *blind*, *localise* > *sense*, *blind* > no-change, *sense* > *false alarm*, *false alarm* > *sense*. We then compared awareness conditions at the second-level using one-sample t-tests. An additional paired-samples t-test was used to identify voxels with activation that was significantly different between the pair of contrasts *localise* > *blind* and *sense* > *blind*.

To identify voxels with activation that correlated with the change in task difficulty over time, a GLM model was constructed with one regressor for the onsets of all trials, and a parametric regressor using the difficulty (or number of squares presented) at each trial.

To identify voxels with activation that correlated with the change in participant certainty over time, a GLM model was constructed with one regressor for the onsets of all trials, and a parametric regressor using the certainty value reported by the participant at each trial.

Across all fMRI analyses, we report clusters with a minimum size of 20 voxels and a cluster-level FWE corrected $p < .001$. Extended local maxima were labelled using the automated anatomical labeling (AAL) toolbox, with

a local maximum radius of 5mm. MNI co-ordinates were used to label voxels according to Broadmann areas. We used the SPM render function to plot our results on the cortex of an MNI brain.

4.2.11 ERP-informed fMRI Analysis

For ERP-informed fMRI analysis, one regressor was constructed for the onset of all change trials (*blind/localise/sense*), with single-trial ERP values included as a parametric regressor. The LP ERP was chosen a-priori for this analysis, as significant differences have previously been identified between awareness conditions within this late parietal potential (Scrivener et al., 2019; Fernandez-Duque & Thornton, 2003; Busch et al., 2010). A second regressor was added for the onset of all no change trials. Motion parameters were also included as nuisance variables.

4.3 Behavioural Results

4.3.1 Accuracy and Difficulty

Accuracy for question 1, in which participants had to identify a change, had a mean of 54% (range $= 39 - 69\%, SD = 9$). Accuracy for question 2, in which participants had to localise the change, had a mean of 72% (range $= 61 - 86\%, SD = 8$). The mean difficulty level given to each participant ranged from 6 to 23 ($M = 16, SD = 4$), with the maximum difficulty experienced by each participant ranging from 18 to 36 ($M = 27, SD = 5$). D'prime scores ranged from .940 to 2.30 ($M = 1.38, SD = .38$). One person had a negative criterion, meaning that they had a response bias towards false alarms. All other participants had positive criterion, indicating a conservative response strategy ($M = .61, SD = .33$). D'prime scores were significantly different from 0 in a one-sampled t-test, indicating that participants could discriminate between change and no change trials, $t(19) = 16.263, p < .001$.

Mean difficulty correlated with mean location accuracy ($r = .590, p = .008$)

and d'prime ($r = -.601p = .005$), but not with mean detection accuracy ($r = -.371, p = .107$). Maximum difficulty also correlated with mean location accuracy ($r = .537, p = .015$) and d'prime ($r = -.482p = .031$), but not with mean detection accuracy ($r = -.349, p = .131$).

4.3.2 Comparison of *sense* and *false alarm* trials

The percentage of *false alarm* trials ($12.23\% \pm 8.64$) was lower than the percentage of *sense* trials ($28.07\% \pm 7.73$) $t(19) = -6.815, p < .001, g_{rm} = 1.85$, suggesting that *sense* trials occurred more often than participants made false alarms. Additionally, the percentage of false alarms was not significantly correlated with the percentage of *sense* trials ($r = .198, p = .403$).

Reaction times for *sense* and *false alarm* trials were compared, to determine if *sense* trials were different to trials where the participant incorrectly reported a change during a no change trial. Reaction times for all *sense* trials (0.636 ± 0.167 s), regardless of certainty, were not significantly different to *false alarm* trials (0.662 ± 0.148 s), $t(19) = -0.974, p = .343, g_{rm} = 0.17$. However, *sense certain* trials (0.543 ± 0.139 s) were significantly faster than *false alarm* trials, $t(19) = -4.500, p < .001, g_{rm} = 0.79$. Therefore, when participants were certain that a change occurred, they responded more quickly than when they were simply making a false alarm.

Reaction times for *sense certain* trials (0.543 ± 0.139 s) were also significantly faster than *false alarm uncertain* trials (0.739 ± 0.213 s), $t(19) = -4.535, p < 001, g_{rm} = 1.01$. However, this may be explained by the general finding that, across all conditions, certain trials (0.546 ± 0.134) were faster than uncertain trials (0.706 ± 0.181 s), ($t(19) = -7.917, p < .001, g_{rm} = 3.43$)

4.3.3 Comparison of *sense* and *blind* trials

Reaction times for *sense* trials (0.636 ± 0.167 s) were not significantly different to *blind* trials (0.665 ± 0.184 s), $t(19) = -0.903, p = .378, g_{rm} = 0.15$. However, reaction times for *sense certain* trials (0.543 ± 0.139 s) were significantly faster

than *blind* trials, $t(19) = -3.499, p = .002, g_{rm} = 0.70$. Therefore, on trials where the participant did not see the change (*blind*), they responded more slowly than when they suspected a change but could not provide additional information about it (*sense*).

Reaction times for *sense certain* trials (0.543 ± 0.139 s) were also significantly faster than *blind uncertain* trials (0.747 ± 0.205 s), $t(19) = -5.121, p < .001, g_{rm} = 1.07$, which again may be explained by the fact that uncertain trials were slower over all conditions.

4.3.4 Comparison of *blind* trials and no-change trials

Out of the 20 participants included in the analysis, 15 were slower to respond when they were *blind* to the change, compared to no-change trials (75%). Reaction times for *blind* trials were significantly slower than no-change trials (0.617 ± 0.176 s), $t(19) = -3.613, p = .002, g_{rm} = 0.25$. Similarly, *blind uncertain* trials (0.747 ± 0.205 s) were significantly slower than no-change trials, $t(19) = 5.328, p < .001, g_{rm} = 0.63$. Therefore, despite being *blind* to the change, the presence of a change in the display increased reaction times, particularly for trials where the participant was uncertain.

4.3.5 Influence of change location

We found a significant effect of location of the changed item (outside/central) on awareness (*blind*/*localise*/*sense*), $\chi^2(2) = 26.68$, p ¡ ,001, as participants were more likely to be *blind* to the change when it occurred on the outside of the display (see table 4.1). There were also a greater number of *sense* trials for outside changes, suggesting that these changes may be harder to localise than central changes. The hemisphere of the display in which the change occurred (left/right) had no significant effect on participant awareness (*blind*/*localise*/*sense*), $\chi^2(2) = 4.941$, $p = .085$ (see table 4.2).

Location	Blind	Localise	Sense
Outside	911	627	290
Central	619	631	220

Table 4.1: Frequency of trials in each level of awareness across different change locations. Changes occurring in any of the 20 outermost positions on a 6 x 6 grid were considered outside changes. Those occurring in any of the 16 central positions on the grid were considered central.

Location	Blind	Localise	Sense
Left	781	651	236
Right	749	607	276

Table 4.2: Frequency of trials in each level of awareness across left and right change locations, based on the hemisphere in which the change occured on a 6 x 6 grid of possible locations.

4.4 EEG Results

4.4.1 P1 and N1

For P1 amplitudes, the main effect of awareness was not significant, $F(2, 38) = .568, p = .572., \eta^2 = .029$. Display was also not significant, $F(1, 19) = .143, p = .709, \eta^2 = .007$. The interaction between awareness and display was not significant, $F(2, 38) = 3.250, p = .050$, $\eta^2 = .146$ (figure 4.2).

For the N1, the main effect of awareness was not significant, $F(2, 38) = 2.008, p = .148$, $\eta^2 = .096$. Display was also not significant, $F(1, 19) = .68., p = .797, \eta^2 = .004$, nor was the interaction between awareness and display, $F(2, 38) = 2.046p = .143$, $\eta^2 = .097$ (figure 4.2).

4.4.2 N2pc

The main effect of hemisphere on N2pc amplitudes was not significant, $F(1, 19) = .338, p = .568$, $\eta^2 = .018$, nor was the main effect of awareness, $F(2, 38) = .878, p = .424$, $\eta^2 = .044$. The interaction was not significant, $F(2, 38) = .572, p = .569$, $\eta^2 = .029$.

As we had strong hypotheses about the presence of an N2pc for *localise*

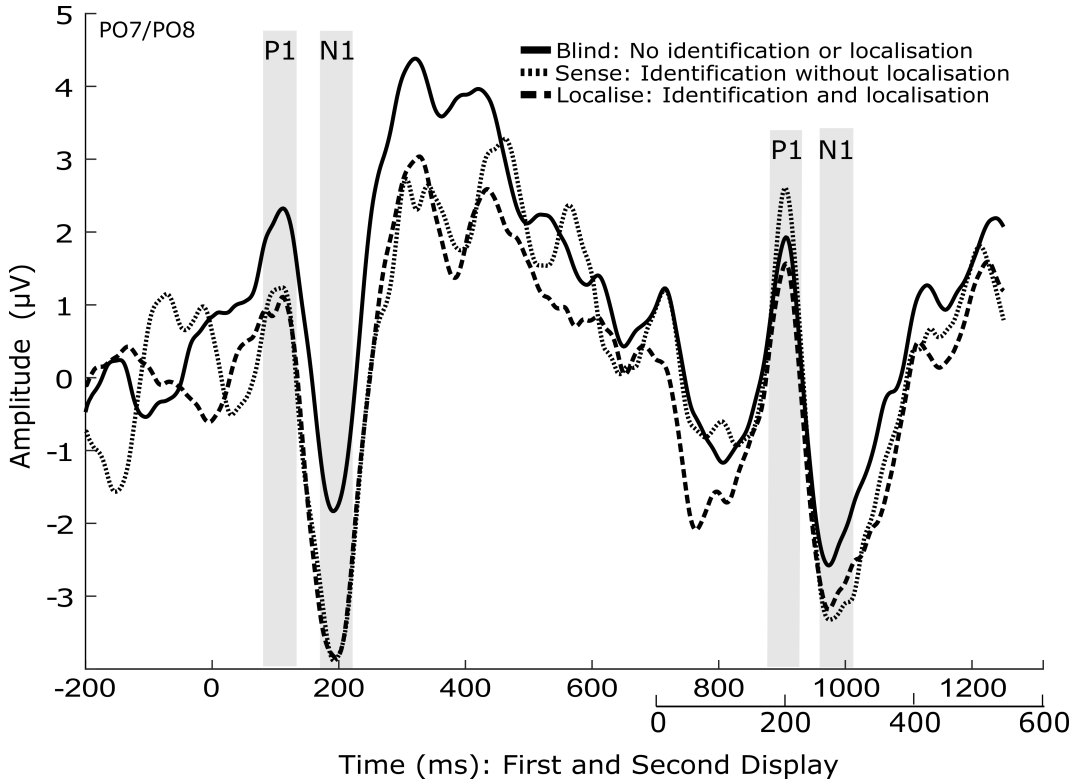

Figure 4.2: ERP plot showing the mean of electrodes PO7 and PO8, for each awareness condition. Condition means for the values within the shaded time windows were used for ERP analysis.

trials, we also ran corrected post-hoc pairwise comparisons across awareness levels. A significantly increased negativity across both hemispheres was found for *localise* trials ($M = -1.573$) compared to *blind* ($M = -.810$) $p = .038$. *Blind* and *sense* ($M = -1.720$) were not significantly different, $p = .259$, nor were *sense* and *localise*, $p = .862$.

4.4.3 Visual Awareness Negativity (VAN)

The main effect of awareness on the VAN was not significant $F(2, 38) = .029, p = .971$, $\eta^2 = .002$.

4.4.4 Late Positivity (LP)

There was a main effect of awareness on LP amplitudes $F(2, 38) = 3.776, p = .032$, $\eta^2 = .166$. In corrected post-hoc comparisons, *localise* trials ($M = 2.270$) had a significantly greater LP amplitude than *blind* ($M = .032$), $p = .024$.

However, *sense* ($M = 1.069$) was not significantly different to *blind*, $p = .130$, or *localise* trials, $p = .174$ (figure 4.3).

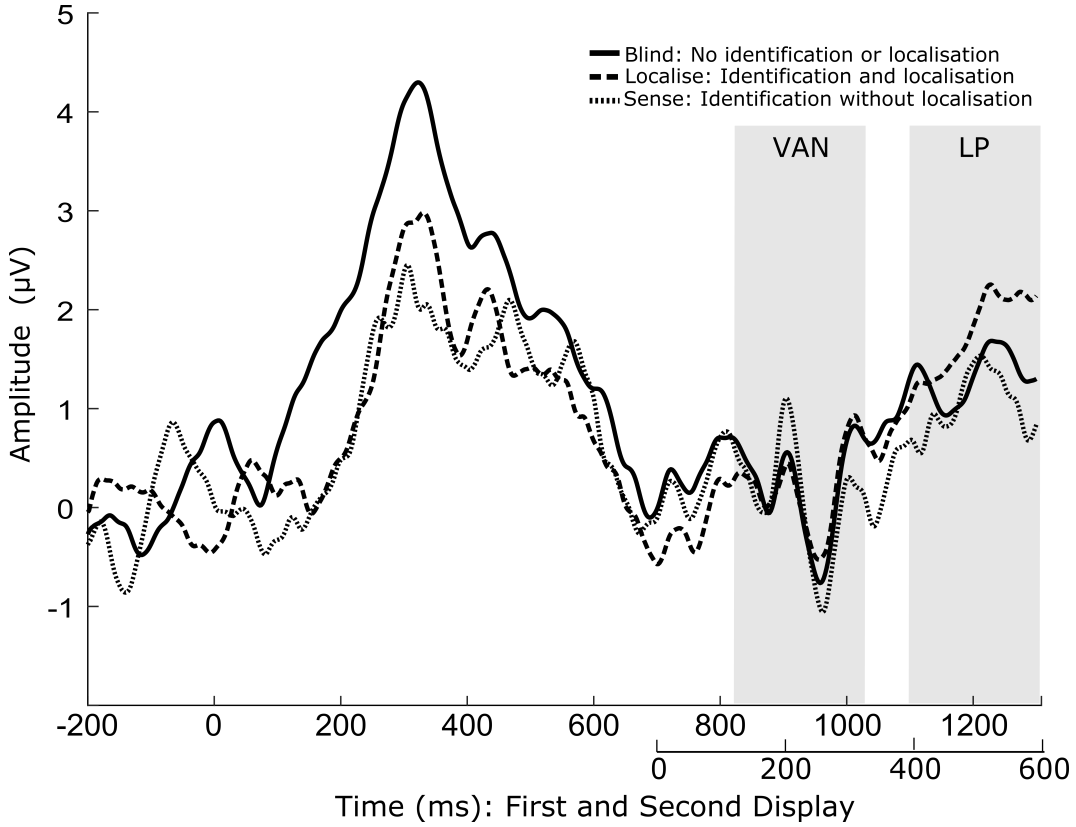

Figure 4.3: ERP plot showing a mean of electrodes Cz, CPz, and Pz, for each awareness condition. Condition means for the values within the shaded time window were used for ERP analysis.

4.5 fMRI Results

4.5.1 Awareness

For the contrast *localise>blind*, increased BOLD activation was found in the bilateral occipital cortex (BA18, V2), the left supramarginal gyrus (BA40), bilateral putamen (BA49), left insula (BA13), left angular gyrus (BA39), left pre-motor cortex (BA6), and right primary sensory cotext (BA1) (see table 4.3).

For the contrast *sense>blind*, increased activation was found in the left pre-motor cortex (BA6), left occipital cortex (BA18, V2), left anterior cingulate

cortex (BA24), and the left supramarginal gyrus (BA40) (see table 4.4).

We also looked for any activation that was significantly greater in one contrast than the other (Loc>blind vs. sense>blind). However, no significant activations remained after correction for multiple comparisons.

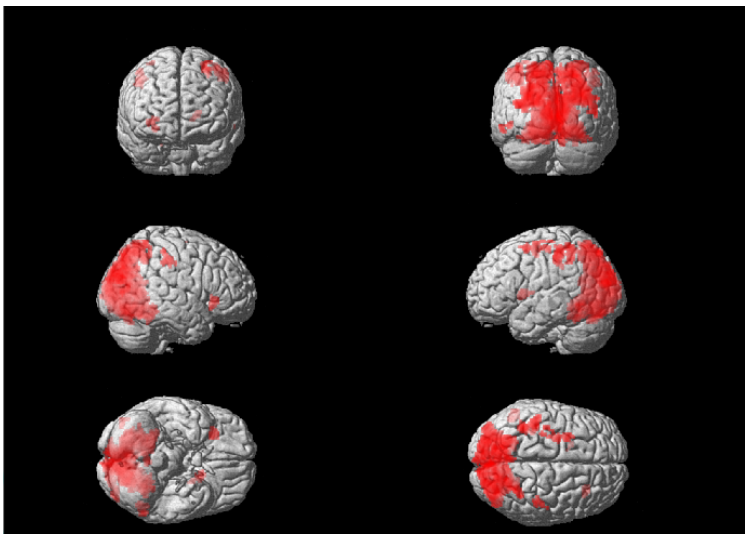

Figure 4.4: Voxels activated for the contrast localise > blind trials. Multiple comparisons were controlled using a cluster level family wise error correction where p < .001, as well as a minimum cluster size of 20 voxels.

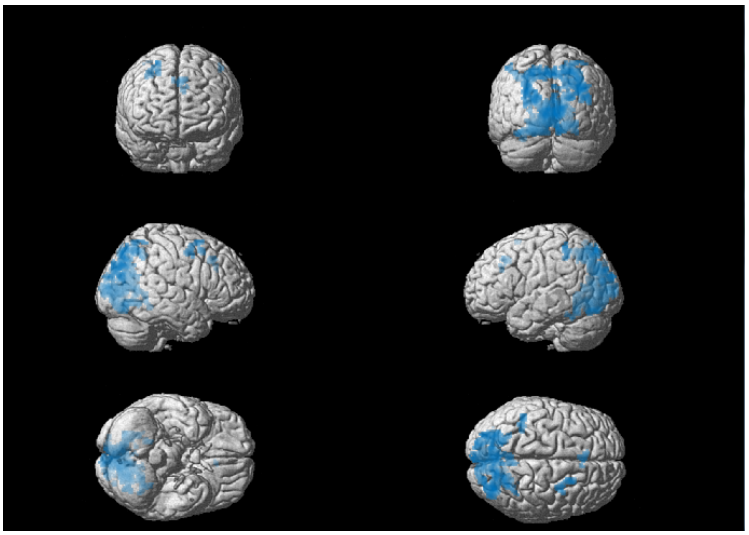

Figure 4.5: Voxels activated for the contrast sense > blind trials. Multiple comparisons were controlled using a cluster level family wise error correction where p < .001, as well as a minimum cluster size of 20 voxels.

No voxels survived for the following contrasts; *localise*>*sense*, *blind*>no change, *sense*>*false alarm*, or *false alarm*>*sense*.

Post-hoc conjunction analysis

Given that the contrasts *localise* vs *blind* and *sense* vs *blind* revealed similar networks of activation, we ran a conjunction analysis to determine which voxels were significantly activated in both contrasts. To do this, we entered the two first-level contrasts for each participant into a one way ANOVA at the second-level (independence not assumed). We then ran a conjunction analysis across both contrasts (1 0 and 0 1) to identify common voxels, using the conjunction null hypothesis as suggested in Nichols et al. (2005). Significant activation was identified in the primary visual cortex (BA18) and parietal cortex (BA40) (see table 4.7 and figure 4.6).

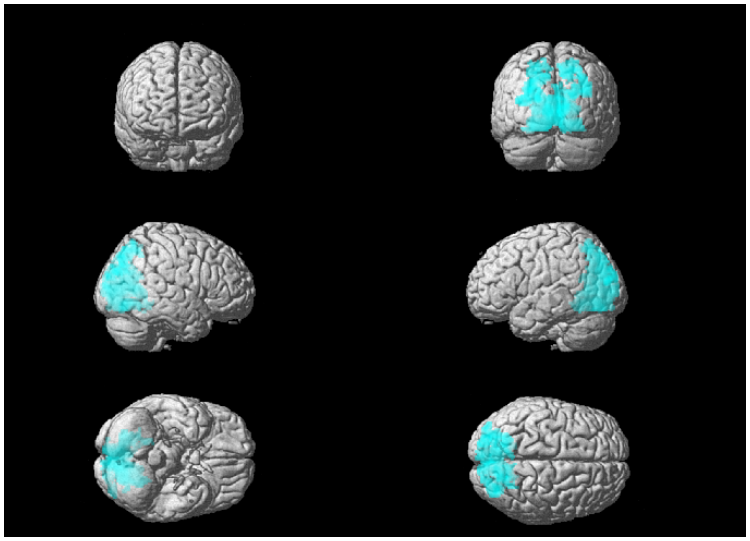

Figure 4.6: Conjunction analysis: voxels significantly activated for both *localise>blind* and *sense>blind* contrasts. Multiple comparisons were controlled using a cluster level family wise error correction where $p < .001$, as well as a minimum cluster size of 20 voxels.

4.5.2 Difficulty and Certainty

The parametric regressor of participant certainty revealed significant activation in the right visual cortex (BA18, V2) and right supramarginal gyrus (BA40). The parametric regressor of task difficulty (the number of squares presented per trial) revealed significant activation in the visual cortex (BA18, V2).

4.5.3 ERP-informed fMRI

No significant voxels were identified for the LP-informed fMRI analysis.

4.6 Discussion

The main aim of this change blindness experiment was to distinguish between trials in which participants could both detect and localise a change in coloured square (*localise*), versus those in which they could only detect it (*sense*), or not detect it at all (*blind*), using combined EEG-fMRI. For the late parietal positivity ERP, *localise* trials were significantly higher in amplitude than *blind* trials as previously found (Scrivener et al., 2019), but *sense* trials were not distinguishable from those where participants were *blind* to the change. Similarly, no differences were found between *sense* and *blind* trials in the N2pc or VAN. It is not clear whether this is due to false positive findings in the previous study, or the smaller signal to noise ratio in the combined EEG-fMRI data. The fMRI results revealed significant differences in BOLD activation for both *localise* and *sense* trials when compared to *blind*, suggesting that they are separable to trials where participants were completely unaware of the change. These results suggest that the *sense* condition may be distinguishable from the traditional *blind* condition, meaning that subjects may have access to more information when they are able to *sense* a change. However, the contrast between *localise* and *sense* conditions revealed no significant activations, meaning that activation specific to full awareness of changes in a change blindness paradigm was not identified.

4.6.1 Behavioural

One explanation for the presence of a *sense* condition in change blindness is that it reflects a liberal response criteria, such that participants report seeing a change even though they were not certain that it occurred (Simons & Ambinder, 2005). In other words, they make a *'false alarm'* during change trials.

If this is the case, then these trials may be similar in number to *false alarm* trials, where participants incorrectly report a change for identical displays where they could not have seen a change. We found that participants had fewer *false alarms* than *sense* trial, and the percentage of these trials across participants was not correlated. This suggests that *sense* trials cannot be attributed to a liberal response criterion of the participants, as the tendency of participants to make a *false alarm* did not influence the number of times they could *sense* a change. However, this differs from previous results, where a significant correlation was found in the percentage of the two trial types (Scrivener et al., 2019). Further behavioural data may therefore be needed to confirm this relationship.

In the reaction time data, *sense certain* trials were significantly faster than both *false alarm* and *blind* trials, which was also found in our previous experiment (Scrivener et al., 2019). This provides evidence that *sense* trials may not be explained entirely by explicit mechanisms, such as participants pressing the wrong response when they have not detected a change. If this were the case, it could be expected that *blind* and *sense* trials would be similar in response time.

Previous studies have also reported that participants responded 'no change' more quickly for no-change trials, compared to change trials (Williams & Simons, 2000; Mitroff et al., 2004). The participant's response is the same in both trial types, but the presence of a change is different. This suggests that even when they fail to detect the change in a change trial, they take longer to respond. We therefore compared reaction times for no-change trials and *blind* trials. Out of the 20 participants, 15 were slower to respond when they were blind to the change, compared to no-change trials (75%), which is higher than the 68% reported by Williams & Simons (2000). Reaction times for *blind* trials were also significantly slower than no-change trials, meaning that even when participants did not notice the change, its presence increased their reaction times. It is possible that in *blind* trials, some information may be available to the participant, leading to slower reaction times, but not enough for them to

be confident to report the change.

The location of the square that changed in colour during the experiment had a significant influence on the likelihood that the change was detected; changes closer to the central fixation were detected at a higher frequency across participants than those further away. One explanation is that the participants were asked to fixate at the centre of the screen, and therefore their overt attention was directed here during the trial. As attention has been found to correlate with change detection, this finding is not surprising (Rensink et al., 1997).

4.6.2 EEG

For the late parietal positivity ERP, *localise* trials were significantly higher in amplitude than *blind* trials. Other studies have also reported increased LP amplitudes for detected versus undetected changes (Fernandez-Duque & Thornton, 2003; Busch et al., 2010), which has been suggested to reflect conscious awareness of changes (Railo et al., 2011) and participant confidence (Eimer & Mazza, 2005). However, *sense* trials were not distinguishable from trials where participants were *blind* to the change. This contradicts our own results from a previous study where all three awareness conditions were distinguishable within the LP (Scrivener et al., 2019). Note that the presence of a significant LP for *localise*, but not for *sense*, should not be used as evidence that the two are different, as the post-hoc comparison was not significant.

For the N2pc results, it should be emphasised that the main effect of hemisphere was not significant. Therefore, the post-hoc comparison in amplitude between *localise* and *blind* trials does not reflect the traditional asymmetry of the N2pc component, with a greater negativity in the contralateral hemisphere. It can only be concluded that there was an increased negativity for *localise* trials across both hemispheres, and may be better interpreted as an N2 component. This is a common finding, and in a review of the ERP correlates of visual awareness (Koivisto & Revonsuo, 2010) the majority of change

blindness paper reported enhanced negativity in the N1-N2 range for detected changes (with the exception of Fernandez-Duque et al.,2003 and Neideggen et al., (2001)).

In a previous EEG study we did find a significant N2pc for both *localise* and *sense* conditions, including a significant main effect of hemisphere (Scrivener et al., 2019). We concluded that the presence of an N2pc for both awareness conditions indicated a shift in attention towards the hemisphere of the change (Luck & Ford, 1998), but that this shift in attention was not sufficient to facilitate correct localisation in *sense* trials. In this experiment, we failed to find any evidence for this shift in either awareness condition, as characterised by the N2pc.

Unlike previous findings from Pourtois et al., the amplitude of the P1 during the first stimuli display was not influenced by the level of awareness (Pourtois et al., 2006). In fact, no significant modulations of awareness were identified within either of the visual ERPs, P1 and N1, across either display, which fails to support previous findings that P1 amplitude during a visual display varies with attention (Wilenius & Revonsuo, 2007) and identification of changes (Mathewson et al., 2009). One possible reason for this could be that the number of squares varied across trials, unlike other experiments where the number was fixed (Pourtois et al., 2006), and therefore possibly driven by inter-individual differences in performance. In a review of the ERP correlates of visual awareness, Koivisto & Revonsuo (2010) list a number of change blindness EEG studies that also failed to detect modulation of an early P1 peak (Eimer, 2000; Koivisto & Revonsuo, 2003; Fernandez-Duque et al., 2003; Schankin & Wascher, 2007; Turatto, 2002; Niedeggen et al., 2001), compared to two studies which did (Busch et al., 2010; Pourtois et al., 2006).

In the visual awareness negativity time window, we found no main effect of awareness. In comparison, (Busch et al., 2010) were able to identify a VAN for their *sense* condition, compared to *blind*. The VAN is thought to be dependent on spatial attention, and requires both the location and identity of an object

to be stored such that it is available for conscious report (Koivisto et al., 2008). As participants were not able to identify the location of change in our *sense* condition, this may explain the lack of significant VAN ERP. In another study (Koivisto et al., 2008), VAN was found to be reduced when participants were asked to keep their eyes fixated at the centre of the screen. This was the case in this experiment, which may also have contributed to the lack of a significant finding withing the VAN window.

4.6.3 fMRI

Awareness

One aim of this experiment was to improve our knowledge of the neurological basis of the *sense* condition with the addition of fMRI results. We found largely overlapping activation for both *localise* and *sense* conditions when contrasted with trials where participants were *blind* to the change in coloured square. Both awareness conditions had significantly greater activation in the early visual cortex (B18, V2), the left supramarginal gyrus in the inferior parietal lobe (BA40), and the left pre-motor cortex (BA6).

The posterior parietal cortex and early visual cortex are commonly implicated as storage sites for the contents of visual working memory (Todd and Marois, 2004; Edin 2009; D'Esposito 2015), and previous fMRI studies of change detection also found activations in these areas (Beck et al., 2001; Pessoa, 2004). Using MVPA, Christophel et al (2012) identified stimuli-specific information contained in both early visual and posterior parietal areas (around the intraparietal sulcus), further implicating these regions as storage sites for visual representations. The activation of these visual and parietal regions in both *localise* and *sense* conditions suggests the presence of visual representations of the stimuli for both levels of awareness. This supports the hypothesis that change blindness may arise from a failure to compare two displays or images, rather than a failure to encode the visual information (Simons et al.,

2005a; Hollingworth et al., 2001). Therefore, the inability of participants to localise the change during *sense* trials may not be explained by a lack of parietal representation, as activity in the dorsal stream (BA18 and BA40) was greater than during *blind* trials.

Activation found only in the *localise* contrast (but not for *sense*) were located in the primary somatosensory cortex (BA1), putamen (BA49), insula (BA13), and angular gyrus in the inferior parietal lobe (BA39). This forms a wider network of activation than the *sense* versus *blind* contrast, including mid-brain structures. The insula and putamen are both hypothesised to act as hubs in key brain networks relating to cognitive control, and their activation specific to *localise* trials may indicate their role in facilitating full awareness of the change. More specifically, the insula forms an integrative hub between attention and salience networks (Menon & Uddin, 2010; Eckert et al., 2009), balancing external attentional cues with internal performance monitoring (Uddin et al., 2017). In contrast, the putamen is suggested to be a central component of a frontal-subcortical network (including the superior parietal and premotor cortex) related to cognitive control (van Belle et al., 2014), and has anatomical connections with rostral parietal areas (Jarbo & Verstynen, 2015). Further, patients with putamen lesions show symptoms of left-sided neglect (Karnath & Rorden, 2012), which is a disorder of attention.

Overall, the pattern of findings indicates both anatomical and functional links between the putamen/insula and parietal cortex, which may explain their increased activation during *localise* trials. However, it should be noted that our fMRI sequence parameters were not specifically designed for accurate recording of mid-brain structures, which may influence the reliability of these results (Eapen et al., 2011).

Activation in the anterior cingulate cortex (ACC) was found in the *sense* versus *blind* contrast. The ACC is commonly linked to functional networks underlying attention (Ungerleider, 2000), and more specifically in boosting attention towards task-relevant stimuli (Orr & Weissman, 2009; Kim et al.,

2016). Further, Mitchell and Cusack (2008) found ACC activation that correlated with estimates of the number of items stored by each participant during a working memory task. If this activation reflects increased attention towards the changed stimuli, then it would be expected to occur in both awareness conditions, as attention facilitates change detection (Rensink et al., 1997). However, ACC activation was not found in the *localise* condition, and therefore cannot be necessary for full awareness of the change.

A more fitting explanation of the ACC activation specific to the the *sense* condition is that it reflects error processing during the task. This is because *sense* trials contained a response error, as participants incorrectly localised the change. Using combined EEG-fMRI, ACC activation has been linked to error processing and is correlated with the error related negativity (ERN) in EEG (Iannaccone et al., 2015; Debener, 2005). Activity in this area could therefore relate to the incorrect responses of the participants during *sense* trials. However, it should be noted that activation in the ACC is found for a wide range of tasks and the specificity of this activation is debated (Dehaene, 2018).

It could be argued that *blind* trials also contain a response error, as the participant failed to report a change that did occur. This should therefore also activate the ACC, if ACC activation reflects error monitoring (and that this error monitoring need not be conscious). Compared to *blind* trials, *sense* trials contained activation in visual (BA18) and parietal (BA40) areas, and the participant correctly reported the change. Dehaene (2018) suggested that there are three distinct types of neural representation: the actual motor response, the intended response, and the accuracy of that response. Using machine learning decoders, they identified the coexistence of representations for both the intended and actual response during a trial. Even when participants made the wrong response, the correct response was still encoded in their brain activity.

It is possible that in *sense* trials, the ACC activation reflects a mismatch between the intended response and the actual response. Although partici-

pants had represented the stimuli in visual working memory (indexed by the increased visual and parietal activation that was similar to *localise* trials), and planned the correct response, their actual response did not match their intended one leading to ACC activation. In *blind* trials, participants had significantly reduced visual and parietal activation, and may not have known which response was correct. Therefore, this mismatch between intended correct response and actual response did not occur. While this may explain our results, this is currently a working theory that should be explored in further research.

Difficulty and certainty

Using participant certainty at each trial as a parametric regressor, we found significant activations in the right visual cortex (BA18, V2) and bilateral supramarginal gyrus (BA40). These regions were also found to increase with awareness of the change (*localise* and *sense* trials), possibly due to the relationship between awareness and certainty. Specifically, when participants were aware of the change and could localise it correctly, they were likely to report higher certainty in their responses. The parametric regressor of task difficulty (the number of squares presented per trial) revealed significant activation in the visual cortex (BA18, V2). This finding likely reflects the greater visual stimulation associated with a more complex visual array. In previous literature, parietal activity has also been correlated with set size and the number of objects stored in visual working memory (Mitchell & Cusack, 2008). Activity also predicts individual differences in working memory capacity (Vogel & Machizawa, 2004). We failed to find this effect, which may be explained by the variation in set sizes that were presented across participants. Instead of presenting a number of blocks with a number of difficulty levels, the difficulty was modulated in real time depending on participant performance (for histograms of participant difficulty levels, see appendix A.4).

4.6.4 ERP-informed fMRI

Our pre-registered analysis method of LP-informed fMRI revealed no significant results. We therefore failed to identify voxels with activation that significantly co-varied with fluctuations in the EEG. It is acknowledged that EEG-BOLD couplings are weak, as they measure the effects remaining after the mean evoked BOLD responses are explained (Liu et al., 2016). However, previous combined EEG-fMRI experiments have managed to identify correlates of EEG using ERP-informed fMRI (Debener, 2005; Eimer & Mazza, 2005), even if at liberal correction thresholds.

One possible reason for the failure to find significant ERP-informed BOLD effects is the reduced signal to noise in EEG signals recorded inside the MRI environment. A second possibility is the method that we used to quantify single-trial ERPs. There is no single method for ERP-informed fMRI analysis, and we therefore chose to run the method that required the least manipulation of the data. In a similar way, both Benar (2007) and Iannaccone et al. (2015) extracted the maximum value from pre-defined windows within each trial in order to construct their single trial regressor, resulting in significant activation patterns. However, there are many more variations on this method. For example, Debener er al. (2005) selected an independent component (derived using ICA) that best represented the error related negativity ERP that they were measuring, filtered the time-series with a 2-10 Hz bandpass filter, and then took a peak-to-peak value within windows chosen based on the grand-averaged ERP. Wirsich et al. (Wirsich et al., 2014) also used single trial values from ICA components chosen to match the N170 ERP of interest. Other processing steps used in ERP-informed fMRI include linear classifiers (Walz et al., 2015; Goldman et al., 2009), autoregressive models (Nguyen et al., 2014), and spatial laplacian filters (Liu et al., 2016), to name only a few.

The majority of ERP-informed fMRI experiments do not report analysis using the simplest single-trial amplitude extraction method that we chose, and

therefore it is not clear if they also ran this but moved on to more complex processing techniques to improve their results. It is our intention to explore further analysis techniques on this data set, and consider the implications of all processing methods in single-trial ERP extraction for combined EEG-fMRI. Further, the decision to run ERP-informed fMRI analysis using the LP ERP was informed by previous results indicating a distinction between *blind*, *localise* and *sense* conditions within this time window (Scrivener et al., 2019). However, we failed to distinguish between *blind* and *sense* in the LP in this data set, and therefore it may not be the most appropriate ERP to identify BOLD correlates of explicit and implicit awareness.

4.6.5 Conclusions

Overall, one of the main aims of this experiment was to establish if the *sense* condition is separable from other awareness conditions in neural signals, as measured using EEG and fMRI. While the phenomenological experience of *sensing* differs from full awareness, it remains unclear whether this arises from a distinct state of neural activation, or whether these trials can be explained by explicit behavioural mechanisms such as participant response errors or lack of confidence. The strongest evidence presented here is the difference in fMRI activation for *blind* trials compared to *sense* trials. Across our sample, there was a greater spread of activation within areas such as the early visual cortex and inferior parietal sulcus when participants suspected a change, compared to when they missed it completely. This suggests that *sense* trials were measurably different to *blind* trials, and that participants did have access to more information regarding the change.

However, the contrast between *sense* and *localise* trials, where participants had full awareness, revealed no significant differences in activation. Additionally, we found no significant activation in a paired-samples t-test comparing the contrasts *sense* > *blind* and *localise* > *blind*. It is therefore difficult to draw any conclusions about the difference in activation between these two condi-

tions, and what activation, if any, is necessary for full awareness of changes. The contrasts for both *localise* and *sense* conditions compared to *blind* trials also revealed an overlapping network of activity. As we failed to distinguish between *sense* and *localise* trials within the EEG data, it is possible that some or all *sense* trials are simply *localise* trials with a response error. Our current definition of *sensing* makes it difficult to eliminate this possibility, and therefore future work could attempt to distinguish between these two trial types with a different experimental paradigm.

While we attempted to distinguish between true *sense* trials and *localise* trials with an error using participant certainty, the number of *sense certain* responses was low. This meant that dividing the awareness conditions into certain/uncertain for EEG or fMRI analysis was not possible. Future experiments could focus on obtaining higher trial numbers, which would hopefully facilitate this analysis. However, the very nature of the *sense* condition means that participants are unlikely to be 'certain' during many of the trials. One way around this would be to include a response option for participants to indicate if they think that they made a response error, although this would only identify trials where the participants were aware of their mistake.

In summary, our data suggests that the phenomenological experience of *sensing* a change is associated with increased activity in visual, parietal, and anterior circulate cortices, when compared to change blind trials. Given this increased activation including areas that are commonly implicated as the storage sites of visual working memory, we argue that *sensing* was not caused by a lack of representation of the visual display. Instead, *sensing* may reflect unsuccessful comparison of the two displays, that prevented localisation of the change in space (Simons et al., 2005a; Hollingworth et al., 2001). Furthermore, *sense* trials may occur when features of a changing object only reach a pre-attentive stage, and are not fully integrated at later stages of visual processing (Galpin et al., 2008; Busch et al., 2009).

Brain Region	Cluster size	% Cluster	MNI Coordinates x	y	z	Z score	T (peak level)	Cluster level FWE (p)
Lingual R (BA18, V2)	12408	42	4	-82	-2	5.36	9.64	<.001
Calcarine L		28						
Lingual L		16						
Calcarine R		14						
Postcentral L (BA40, supramarginal)	421	75	-38	-30	42	4.71	7.33	<.001
Parietal inf L		25						
Putamen L (BA49)	90	71	-18	8	8	4.25	6.08	<.001
Caudate L		19						
Pallidum L		10						
Frontal Mid L (BA6, pre-motor)	182	99	-28	4	56	4.18	5.89	<.001
Precentral L		1						
Insula R (BA13)	96	51	28	22	-10	4.08	5.66	<.001
Inferior orbitofrontal R		33						
Putamen R		16						
Angular gyrus L (BA39)	84	100	-56	-56	12	3.71	4.87	<.001
Postcentral R (BA1, primary sensory)	110	91	44	-24	38	3.62	4.69	<.001
Supramarginal R		9						

Table 4.3: Voxels significantly activated for the contrast *localise* > *blind*, cluster FWE $p < .001$, minimum 20 voxels.

Brain Region	Cluster size	% Cluster	MNI Coordinates x	y	z	Z score	T (peak level)	Cluster level FWE (p)
Precentral R (BA6, pre-motor)	145	81	-32	-4	48	4.94	8.07	<.001
Frontal Mid R		19						
Calcarine L (BA18, V2)	5526	91	-4	-96	4	4.92	8.00	<.001
Occipital superior L		6						
Cuneus L		3						
Frontal medial superior L (BA24, anterior cingulate)	88	91	-2	26	0	3.96	5.38	<.001
Cingulum mid L		9						
Parietal inf L (BA40, supramarginal)	87	79	-48	-42	56	3.67	4.78	<.001
Postcentral L		21						

Table 4.4: Voxels significantly activated for the contrast *sense* > *blind*, cluster FWE $p < .001$, minimum 20 voxels.

Brain Region	Cluster size	% Cluster	MNI Coordinates x	y	z	Z score	T (peak level)	Cluster level FWE (p)
Lingual gyus R (BA18, V2)	160	83	14	-74	0	4.11	5.74	<.001
Calcarine R		17						
Inferior parietal R (BA40, supramarginal)	104	57	50	-34	48	3.84	5.12	<.001
Supramarginal gyrus R		27						
Postcentral gyrus R		16						

Table 4.5: Parametric regressor: participant certainty, cluster FWE $p < .001$, minimum 20 voxels.

Brain Region	Cluster size	% Cluster	MNI Coordinates x	y	z	Z score	T (peak level)	Cluster level FWE (p)
Mid occipital L (BA18, V2)	69	30	-16	-92	-6	4.23	6.02	0.003
Inferior occipital L		27						
Calcarine L		26						
Lingual gyrus L		14						

Table 4.6: Parametric regressor: task difficulty, cluster FWE $p < .001$, minimum 20 voxels.

Brain Region	Cluster size	% Cluster	MNI Coordinates x	y	z	Z score	T (peak level)	Cluster level FWE (p)
Lingual L (BA18)	6116	73	-4	-70	6	5.51	7.35	¡.001
Calcarine L		27						
Supramarginal gyrus (BA40)	445	89	28	-76	44	4.91	6.16	¡.001
Parietal Sup R		11						
Precuneus R (BA40)	421	73	18	-70	42	4.73	5.84	¡.001
Occipital Sup R		16						
Cuneus R		11						

Table 4.7: Conjunction analysis: voxels activated for both *localise* versus *blind* and *sense* versus *blind* contrasts. Cluster FWE $p < .001$, minimum 20 voxels.

Chapter 5 contains the manuscript for a short paper comparing the EEG and behavioural results from the data recorded inside and outside of the MRI environment. This utilises the data from the EEG only experiment reported in Chapter 3, and the combined EEG-fMRI data reported in Chapter 4. The aim of this Chapter is to assess the influence of the scanner environment on the ERP and behavioural results.

The majority of the analysis is reported within the main text. However, additional figures and reports can be found in the appendix (A.5).

Chapter 5

Behavioural and electrophysiological differences between EEG data recorded inside and outside of the MRI environment

5.1 Introduction

The combined recording of EEG and fMRI has the ambitious aim of overcoming the respectively poor spatial and temporal resolution of the two methods (Ullsperger & Debener, 2010). However, by recording EEG inside the magnetic environment of an MRI scanner, several additional artifacts are added to those typically found in EEG data. On top of the typical eye movement and muscle related artifacts, EEG recorded in the scanner suffers from artifacts related to the MR gradient, the participant's heart beat (Debener et al., 2008), and to the helium pump (Garreffa et al., 2004). The MR artifact can be successfully removed, provided that the MR and EEG clocks are synced (Allen et al., 2000), as the artifact is stationary over time. However, the other two artifacts are difficult to remove, and can reduce the quality of EEG data recorded in the MRI environment (Allen et al., 1998).

One way to assess the influence of MRI artifacts on EEG data is to measure the EEG during quiet scanning periods, when the scanner is not acquiring functional data. This can then be compared to EEG recorded during simultaneous

scanning to assess the feasibly of the method and quality of the data. Overall, results from experiments using visual stimuli, such as a flashing checkerboard, indicate consistent visual ERPs with and without scanning. For example, Kruggel et al. (2000) found no significant differences in visual P2 or P3 latencies in response to a flashing checkerboard. Similar results were found by Becker et al. (2005), who also report consistent amplitudes of P2 and N3 peaks during scan periods. In addition, ERPs recorded with and without simultaneous scanning across participants were highly correlated. Further, Sommer et al. (2002) reported no differences in P1 or N145 amplitudes. Visual P1 and N1 EPRs elicited by the onset of coloured words were also found to be comparable by Comi et al. (2005). A later study by Warbrick & Bagshaw (2008) reported slightly larger P1 and N140 amplitudes during a visual checkerboard with continuous scanning, however, the differences were not significant.

Auditory ERPs have also been compared. Novitski et al. (2001) reported increased P1, N1, and P2 latencies with simultaneous scanning, but unaffected mismatch negativity (MMN) and P3a. However, a later study found no increase in latency for the auditory N1 or P2 (Mayhew et al., 2010). The increased auditory stimulation caused by simultaneous scanning could provide one explanation for these differences.

An alternative comparison is between EEG recorded inside the MRI environment, to EEG recorded in a separate room. During an auditory oddball task, Mulert et al. (2004) found reduced N1 amplitudes and latencies during scanning compared to an equivalent experiment outside. However, the P3 was unaffected. Bregadze & Lavric (2006) also found no differences in visual P3 amplitude during both no-go and N-back tasks. Sammer (2005) compared a number of commonly recorded EEG features, including steady-state visual evoked potentials (SSVEP), the lateral readiness potential (LRP), and theta oscillations. They found a reduction in the frequency of interest for the SSVEP paradigm, suggesting that artifact-subtraction led to an attenuation in the frequency domain. However, the averaged peaks in the time domain were higher

inside the MRI scanner for both SSVEP and LRP experiments. To complicate this further, the theta frequency was increased during the mental arithmetic experiment, but only for lateral electrodes. While these differences did not confound or destroy the effects of interest, the influence of the MRI environment requires further investigation given the conflicting results.

The aim of this experiment was to compare behavioural and ERP results acquired inside and outside of the scanner environment during a change blindness task, using a between groups design; group 1 completed a change blindness paradigm with EEG recording in a shielded room, and group 2 completed the same paradigm with simultaneous EEG-fMRI recording. This is in contrast to previous literature mentioned above that focus on within-group variation.

It is important to note the differences between the experimental sessions, including: display presentation (LCD monitor versus mirror display), response methods (keyboard versus response box), and EEG caps (BrainProducts EasyCap versus BrainProducts EasyCap MR). While these differences have the potential to add variance to the recorded data, we believe that it is still beneficial to examine the concordance between EEG recorded inside and outside of the MRI environment. Many studies report similar findings across recording environments, suggesting that these differences may not disrupt the effects of interest. However, given the relatively small number of previous studies, we believe any comparison to be a useful addition to the literature.

5.2 Materials and Methods

For full details of the methods and behavioural analysis for the EEG only experiment, please refer to Chapter 3. For the EEG-fMRI experiment, please refer to Chapter 4.

5.2.1 EEG Data Acquisition

EEG only: EEG data was recorded with a BrainVision EasyCap (Brain Products), with 64 passive electrodes including an IO channel, arranged according to the 10-10 layout. The reference electrode was placed at FCz and the ground at AFz. Impedance was kept below 10kΩ for all the EEG channels, and 5kΩ for the IO channel. EEG signals were recorded using BrainVision Recorder (Brain Products, version 1.20) at a sampling rate of 5000 Hz.

EEG-fMRI: EEG data was recorded with an MRI-compatible cap equipped with carbon-wired Ag/AgCL electrodes (Braincap MR) from 64 scalp positions according to the international 10-10 system. The reference electrode was placed at FCz and the ground at AFz. An additional ECG electrode was positioned on the back to measure heart rate. An MRI-compatible EEG amplifier was used (Brain-Amp MR, Brain Products) with a sampling rate of 5000Hz. Impedance was kept below 10kΩ for EEG channels and 5kΩ for the ECG. EEG recordings were performed with Brain Vision Recorder Software (Brain Products) and timings kept constant using a BrainProducts SyncBox to synchronise EEG with the MRI system clock.

5.2.2 EEG pre-processing

EEG only: Raw EEG data was pre-processed using BrainVision Analyzer (Brain Products, version 2.1). The data was first downsampled to 500 Hz to reduce computation time, then filtered with a high-pass filter of 0.01 Hz to remove low frequency drift (Butterworth, 2nd order). A low-pass filter of 50 Hz and a notch filter of 50 Hz were chosen to remove line noise. Independent component analysis (ICA) was used to remove eye movement artifacts (FastICA). Two components were removed for each participant; one corresponding to eye-blinks and the other to lateralised eye-movements.

EEG-fMRI: Raw EEG data was pre-processed using Brain Vision Analyzer version 2.1 (Brain Products). Correction for the MR gradient artifact

was performed using a baseline corrected sliding average of MR volumes (Allen et al., 2000). Removal of cardioballistic artfacts involved the subtraction of heartbeat artifacts on a second by second basis, using a sliding average of 21 (Allen et al., 1998). The delay was detected using the CBC detection solution, individually for each subject. Peaks were detected semi-automatically, with a manual check of the algorithm's estimations. ICA was then used to remove further BCG residual artifacts (range: 1 - 4 additional ICs removed per participant). As outlined in (Debener, 2005), the presence of visual P1 and N1 peaks in the averaged data after pre-processing was used as an indication of the successful removal of artifacts.

The data was downsampled to 500 Hz to reduce computation time, then filtered with a high-pass filter of 0.01 Hz to remove low frequency drift (Butterworth, 2nd order). A low-pass filter of 50 Hz and a notch filter of 50 Hz were chosen to remove line noise. Independent component analysis (ICA) was used to remove eye movement artifacts (FastICA). Two components were removed for each participant; one corresponding to eye-blinks and the other to lateralised eye-movements.

5.2.3 EEG Analysis

The trials in which a change occurred were divided into three conditions: *blind* (no change detection), *localise* (change detection and localisation), and *sense* (change detection without localisation). Trials in which no change occurred were divided into *correct rejection* (no change reported) and *false alarm* (change incorrectly reported).

To identify the peaks of the visually evoked potentials (P1 and N1), a grand average ERP was calculated across all conditions and participants, as advised in Luck & Gaspelin (2017), from electrodes P07 and P08. From here, the peaks of interest were determined by identifying the local maxima/minima of the expected peaks, using the peak detection function in BrainVision Analyzer. The mean value within a window around the peak was used instead of the peak

value, as the mean is more robust against noise (Luck, 2014). For the EEG only experiment: In relation to the first display onset, the first P1 was identified at 122ms, and the first N1 at 212ms. In relation to the second display onset, the second P1 was identified at 114ms, and the second N1 at 222ms. For the EEG-fMRI experiment: the first P1 was defined at 124 ms, the second P1 at 108 ms, the first N1 at 184 ms, and the second N1 at 168 ms.

Based on previous literature (Busch et al., 2010; Tseng et al., 2012; Fernandez-Duque et al., 2003), the N2pc was defined as the mean within 200-400 ms after the second display at occipital electrodes PO7 and PO8. Over central parietal electrodes Cz, CPz and Pz, the VAN was defined within a window of 130-330 ms after the second display, and the LPP within a window of 400-600ms. We used window sizes of 200 ms, defined a-priori, in an attempt to be conservative given the large variation within the literature.

P1 and N1 amplitudes were compared in two separate repeated measures ANOVAs, with display (first/second) and awareness (*blind*/*localise*/*sense*) as the independent variables, and a between subject factor of group (EEG only/EEG-fMRI). LPP, and VAN amplitudes were analysed using separate repeated measures ANOVAs for awareness condition (*blind*/*localise*/*sense*), with a between subject factor of group (EEG only/EEG-fMRI). N2pc amplitudes were analysed using a repeated measures ANOVA for awareness condition (*blind*/*localise*/*sense*) and hemisphere (ipsilateral/contralatel), with a between subject factor of group (EEG only/EEG-fMRI). Latency values were analysed in the same way.

Where Mauchly's Test of Sphericity indicated that the assumption had been violated, Greenhouse-Geisser correction was used. All post-hoc comparisons were two-tailed, and corrected for multiple comparisons using false discovery rate where $q = .05$ (Benjamini & Hochberg, 1995). Effect sizes are reported as partial eta squared for ANOVA, and repeated measures Hedge's g for t-tests (Lakens, 2013).

For single-trial analysis, time courses were constructed for each participant

	Experimental Session		
DV	EEG (SD)	EEG-fMRI (SD)	Difference (p)
Mean difficulty	14 (3)	16 (4)	0.095
Maximum difficulty	26 (4)	27 (5)	0.349
Detection accuracy (%)	49 (13)	54 (9)	0.217
Localisation accuracy (%)	70 (8)	72 (8)	0.375
D'prime	0.74 (.27)	1.38 (.38)	0.059
Criterion	0.60 (.42)	0.61 (.33)	0.942

Table 5.1: Behavioural differences between EEG and EEG-fMRI experimental groups. Mean (SD).

from the single-trial values of each ERP, at each channel (7 ERPs, 64 channels, 20 participants). Each single-trial value was calculated as the mean amplitude within the pre-defined ERP window at each trial. These values were baseline corrected by subtracting the mean of the trial from which they were selected. Outliers were defined as single-trial values with an amplitude greater than 3 standard deviations from the mean amplitude at each electrode. As large artifacts can raise the mean amplitude, we added the additional classification of outliers at values +/- 30 μV. The number of outliers present in the electrodes used for analysis was then compared across both experiments. Additionally, we calculated the variance of the single-trial values for each ERP, in relation to grand averaged waveform. This was calculated for all ERP except the N2pc, as this is a difference wave rather than a mean amplitude.

5.3 Behavioural Results

No differences were found across experimental groups for any of the main behavioural measures, including task difficulty and accuracy (see table 5.1). Additional behavioural results relating to the presence of the *sense* condition can be found in the appendix (A2).

5.3.1 Reaction times

Average reaction were compared using a repeated measures ANOVA calculated for question (1/2/3) and accuracy (correct/incorrect), with a between group factor of experimental condition (EEG only/EEG-fMRI). The main effect of question was significant, $F(2, 76) = 7.117, p = .001, \eta^2 = .158$. (Q1: $M = .633$, Q2: $M = .710$, Q3: $M = .718$). The detection question (Q1) was significantly faster than the localisation question (Q3), $p = .007$, and the certainty question (Q3), $p = .001$. However, the localisation and certainty questions were not significantly different, $p = .720$.

There was a significant main effect of accuracy, $F(1, 38) = 81.612, p < .001, \eta^2 = .682$, as reaction times for correct answers were faster ($M = .637$) than incorrect answers ($M = .650$). The interaction between question and accuracy was not significant, $F(2, 76) = 2.830, p = .030, \eta^2 = .068$. This is because, for all questions, correct responses were faster than incorrect responses.

There was a significant main effect of group, $F(1, 38) = 15.681, p < .001, \eta^2 = .292$, as reaction times for the EEG-fMRI group were faster ($M = .605$) than the EEG only group ($M = .769$). The interaction between question and group was significant, $F(2, 76) = 6.339, p = .003, \eta^2 = .143$, as the differences in reaction times across the three questions was not the same in both groups. In the EEG-fMRI experiment, the certainty question (Q3) had the slowest reaction times, while the detection (Q1) and localisation questions (Q2) were similar in their average reaction times (Q1: $M = .585$, Q2: $M = .578$, Q3: $M = .652$). In the EEG only experiment, the detection question (Q1) was the fastest, followed by the certainty question (Q3) followed by the location question which was the slowest (Q2) (Q1: $M = .680$, Q2: $M = .842$, Q3: $M = .784$).

The interaction between accuracy and group was also significant, $F(1, 38) = 5.450, p = .025, \eta^2 = .125$. In both groups, correct responses had a quicker reaction time. However, all responses were faster in the EEG-fMRI group (EEG: correct $M = .706$, incorrect $M = .831$, EEG-fMRI: correct $M = .568$,

incorrect $M = .642$).

5.4 EEG Results

5.4.1 P1

For P1 amplitudes, the main effect of awareness was not significant, $F(2, 76) = 1.148, p = .323$, $\eta^2 = .029$. Display was also not significant, $F(1, 38) = .367, p = .548, \eta^2 = .010$, However, the interaction between awareness and display was significant, $F(2, 76) = 33.785, p = .027$, $\eta^2 = .091$. In the *blind* condition, the P1 was larger in amplitude in the first display ($M = -2.304$) than the second ($M = 2,082$). For the *localise* and *sense* conditions, the P1 was smaller in the first display (*localise* $M = 1.289$, *sense* $M = 1.439$) than in the second (*localise* $M = 1.870$, *sense* $M = 2.406$). The main effect of group was not significant, $F(1, 38) = 1.662, p = .205$, $\eta^2 = .042$.

For P1 latency, the main effect of awareness was not significant, $F(2, 76) = .072, p = .930$, $\eta^2 = .002$. Display was significant, $F(1, 38) = 247.696, p < .001, \eta^2 = .867$. The interaction between awareness and display was not significant, $F(2, 76) = .024, p = .977$, $\eta^2 = .001$. The main effect of group was significant, $F(1, 38) = 320.883, p < .001$, $\eta^2 = .895$, with an earlier P1 found in the EEG-fMRI group ($M = 119.774$) compared to the EEG only ($M = 165.002$).

5.4.2 N1

For the N1, the main effect of awareness was significant, $F(2, 76) = 3.688, p = .030$, $\eta^2 = .088$. Display was not significant, $F(1, 38) = .263, p = .611, \eta^2 = .007$, nor was the interaction between awareness and display, $F(2, 76) = .970, p = .384$, $\eta^2 = .025$. The main effect of group was significant, $F(1, 38) = 28.269, p < .001$, $\eta^2 = .427$, with a larger N1 found in the EEG-fMRI group ($M = -8.984$) compared to the EEG only ($M = -2.746$).

Corrected post-hoc pairwise comparisons across awareness levels showed

significant differences between *blind* ($M = -5.229$) and *localise* ($M = -6.198$), $p = .042$, and between *blind* and *sense* ($M = -6.167$), $p = .019$. However, *sense* and *localise* were not significantly different, $p = .933$.

For N1 latency, the main effect of awareness was not significant, $F(2, 76) = .816, p = .446$, $\eta^2 = .021$. Display was significant, $F(1, 38) = 613.650, p < .001, \eta^2 = .942$. However, the interaction between awareness and display was not significant, $F(2, 76) = .320, p = .727$, $\eta^2 = .008$. The main effect of group was significant, $F(1, 38) = 247.682, p < .001$, $\eta^2 = .867$, with an earlier N1 found in the EEG-fMRI group ($M = 218.415$) compared to the EEG only ($M = 227.796$).

5.4.3 N2pc

There was a main effect of awareness on N2pc amplitudes, $F(2, 76) = 3.178, p = .047$, $\eta^2 = .077$. The main effect of hemisphere was not significant, $F(2, 38) = 3.981, p = .053$, $\eta^2 = .095$. The interaction was also not significant, $F(2, 76) = .926, p = .401$, $\eta^2 = .024$ nor was the main effect of group, $F(1, 38) = .922, p = .343$, $\eta^2 = .024$

Corrected post-hoc pairwise comparisons across awareness levels showed significant differences between *blind* ($M = -1.432$) and *localise* ($M = -2.257$), $p = .003$. However, *blind* and *sense* ($M = -2.284$), $p = .050$, and *sense* and *localise* were not significantly different, $p = .954$.

For N2pc latency, the main effect of awareness was not significant $F(2, 76) = .104, p = .901$, $\eta^2 = .003$. However, the main effect of group was significant $F(1, 38) = 32.317, p < .001$, $\eta^2 = .460$. The mean latency for EEG only group was 309.449 ms, in relation to the second display onset, and 289.185 ms for the EEG-fMRI group.

5.4.4 Visual Awareness Negativity (VAN)

The main effect of awareness on the VAN was not significant, $F(2, 76) = .520, p = .596$, $\eta^2 = .014$. The main effect of group was significant, $F(1, 38) =$

$4.729, p = .036$, $\eta^2 = .111$, with a larger VAN found in the EEG-fMRI group ($M = -1.867$) compared to the EEG only ($M = .116$).

For VAN latency, the main effect of awareness was not significant $F(2, 76) = .074, p = .929$, $\eta^2 = .002$. However, the main effect of group was significant $F(1, 38) = 24.012, p < .001$, $\eta^2 = .387$. The mean latency for EEG only group was 229.036 ms, in relation to the second display onset, and 214.446 ms for the EEG-fMRI group.

5.4.5 Late Positive Potential (LPP)

There was a main effect of awareness on LPP amplitudes $F(2, 76) = 6.760, p = .002$, $\eta^2 = .151$. The main effect of group was significant $F(1, 38) = 8.953, p = .005$, $\eta^2 = .191$, with a larger LPP found in the EEG-fMRI group ($M = 3.476$) compared to the EEG only ($M = .1.107$).

In corrected post-hoc comparisons, *blind* ($M = 1.482$) was significantly different to *localise* ($M = 3.063$), $p = .002$, and *sense* ($M = 2.330$), $p = .024$. *localise* and *sense* were not significantly different to one another, $p = .099$.

For LPP latency, the main effect of awareness was not significant $F(2, 76) = .909, p = .407$, $\eta^2 = .023$. The main effect of group was also not significant $F(1, 38) = 2.801, p = .102$, $\eta^2 = .069$

5.4.6 Single-Trial Analysis

In an independent samples t-test between experimental session, there was no significant difference between the mean single trial variance across ERPs, $t(10) = -.087, p = .932$. However, there was a significant difference between the standard deviations of the single trial variances, $t(10) = -2.896, p = .016$. The average SD for the EEG only experiment was 16.08, compared to an average SD of 33.61 in the EEG-fMRI experiment.

The mean number of outliers excluded per participant was higher in the EEG-fMRI experiment, with a higher standard deviation across participants ($M = 7$ trials, $SD = 12.98$) than in the EEG only experiment ($M = 3$ trials,

$SD = 2.46$).

5.5 Discussion

With the exception of the LP, all ERP latencies were significantly influenced by experimental group (EEG only/EEG-fMRI). The first and second P1, second N1, VAN, and N2pc occurred at earlier latencies in the EEG-fMRI group, whereas the first N1 latency was increased. In general, ERP amplitudes were increased in the EEG-fMRI experiment, with significantly greater amplitudes for the first and second N1. The differences in amplitude for the LP and VAN, however, did not remain significant when corrected for multiple comparisons.

While the mean variance of the ERP amplitudes did not vary significantly across experimental sessions, the standard deviation of the variance across participants was higher in the EEG-fMRI experiment. This means that participants in the EEG-fMRI group had a wider range of single trial values, and also explains the higher number of outliers identified. One explanation for this could be the smaller signal to noise ratio of EEG recordings in the MRI environment, and the larger number of possible artifacts (Ullsperger & Debener, 2010).

Another explanation for the differences between groups is the amount and extent of pre-processing that was used to prepare the data for analysis. In the EEG only experiment, standard pre-processing steps were used, such as down-sampling (from 5,000 to 500 Hz), filtering (low pass: 50 Hz, high pass, 0.01, notch: 50 Hz) and artifact removal using ICA (2 eye movement components per participant). However, the EEG data recorded inside the scanner required a larger amount of processing, given the addition of several artifacts from the scanner.

The MR and BCG artifact removal algorithms implemented in BrainVision Analyzer are both based the identification and removal of a template of the artifact that it detects. If this template is inaccurate, then additional data

may also be removed, or artifactual data left behind. While the MR artifact is stable over time, and can be removed fairly easily when the MR clock is synchronised with the EEG amplifier, the BCG artifact varies greatly, even within a single participant. Any change in heart-rate, breathing, or movement of the participant, can change the shape of the BCG artifact, and therefore impede successful removal.

There is some evidence for the influence of pre-processing on latency in ERP analysis. However, if the pre-processing used in the EEG-fMRI experiment produced a widespread shift in the data, then all ERPs should shift in the same direction. This was not the case; most ERPs had shorter latencies, except for the N1 which had a significantly later peak. Unfortunately, it is not possible to run identical pre-processing steps on both data sets, given the lack of MR and BCG artifacts present in the EEG only data.

However, differences in ERP amplitude compared across separate participant groups should be interpreted with caution. Raw voltage values recorded using EEG may vary across individuals for a number of different reasons, all independent of the experimental design. For example, differences in skull shape and thickness, anatomical arrangement of the cortex, the orientation of the contributing dipoles, the quality of the EEG electrodes and amplifier, the preparation of the electrodes, the position of the electrode cap, and the cleanliness of the participant's hair (Cohen, 2014; Ullsperger & Debener, 2010). Also, statistical power for between-group designs is reduced in comparison to within-group analysis (Field, 2013).

Behaviourally, participants responded more quickly in the EEG-fMRI experiment. Comparing the three questions, participants in the EEG-fMRI group had the slowest reaction times for the certainty question (Q3), whereas participants in the EEG only group were slowest for the localisation question (Q2). This may be a reflection of the differences in response method; inside the MRI scanner, participants had to respond using a four-button box and their right hand only, whereas the EEG participants used a keyboard and two fingers

		Experimental Session		
DV	ERP	EEG Only (SD)	EEG-fMRI (SD)	Difference (p)
Amplitude (μV)	First P1	0.851 (2.295)	2.310 (3.732)	0.145
	First N1	-2.698 (3.344)	-9.421 (5.164)	< .001*
	Second P1	2.1196 (4.293)	2.694 (4.674)	0.688
	Second N1	-1.870 (4.176)	-8.415 (4.559)	< .001*
	VAN	0.284 (2.761)	-1.570 (2.755)	0.040
	LPP	1.291 (3.531)	3.306 (2.103)	0.035
	N2pc	-0.001 (1.682)	0.557 (1.161)	0.230
Latency (ms)	First P1	281.112 (4.860)	262.029 (1.585)	< .001*
	First N1	187.7699 (6.316)	213.527 (1.989)	< .001*
	Second P1	205.4269 (2.491)	116.019 (1.947)	< .001*
	Second N1	267.952 (9.897)	232.05 (1.737)	< .001*
	VAN	227.608 (8.469)	213.497 (9.801)	< .001*
	LPP	504.691 (5.872)	501.611 (8.309)	0.184
	N2pc	309.449 (7.690)	289.185 (13.964)	< .001*

Table 5.2: Amplitude and latency differences between EEG and EEG-fMRI experimental groups. Latency values are calculated in relation to the onset of the relevant display. P-values with * indicate effects that remain significant when corrected for multiple comparisons using false discovery rate, where $q = .05$. Mean (SD).

from each hand.

Overall, we found significant differences in the amplitude and latency of ERPs recorded inside and outside the MRI environment, particularly within the visual N1 peak. However, these results should be interpreted with caution, given the between-subject design of this experiment and the different pre-processing steps performed on both data sets. Further investigation should focus on quantifying the effects of pre-processing steps on EEG data recorded with simultaneous fMRI acquisition, and the amount of variation that can be expected between participant groups.

Chapter 6 contains the manuscript for a review paper entitled 'When is simultaneous recording necessary? A guide for researchers considering combined EEG-fMRI'. This is the resource that I wish I had read before beginning my PhD.

Chapter 6

When is simultaneous recording necessary? A guide for researchers considering combined EEG-fMRI

6.1 Introduction

To obtain a complete and accurate map of neural function in a living brain, we would need a method that could provide a minimum of millimetre spatial resolution at millisecond temporal resolution. Currently, no such method exists, so we rely on the synthesis of information from a mixture of methods, each with their own strengths and weaknesses (Jorge et al., 2014). Electroencaphalogram (EEG) and functional magnetic resonance imaging (fMRI) are two of the most common methods in neuroscience, providing non-invasive measures of brain activity.

EEG has high temporal resolution, and can record upwards of 5000 samples every second. This rich data set provides detail on when activity occurs in the brain, and how task conditions, experimental groups, or a number of other factors influence the timing of this activation. However, EEG signals have poor spatial resolution as they capture the summation of activity that reaches scalp level (Nunez & Silberstein, 2000). Signals travel within the brain and through the scalp, meaning that they experience volume conduction and the signal is dispersed across wide areas (Michel et al., 2004). Also, the contribution of brain regions below the cortex is unknown, with the most likely sources of the

EEG signal lying only millimetres below the surface of the skull. The greatest contribution to the EEG signal is thought to come from pyramidal neurons oriented perpendicularly to the cortical surface (Niedermeyer & Silva, 2005), and therefore the ability to reliably measure activity from regions below the cortex is debated (Cohen, 2014; Seeber et al., 2019).

Conversely, fMRI has good spatial resolution, often recording activity within sections of the brain around 2 mm cubed. This resolution can also be improved if specific areas of the brain are chosen as a focus for data collection. The ability to localise activity is valuable for determining the contribution of brain regions to specific tasks, and the identification of networks with similar activation (Ogawa et al., 1990). However, the temporal resolution of fMRI is low, with a typical experiment measuring activation every 1 or 2 seconds (although this varies across experiment and scanning sequence). As action potentials take only milliseconds to occur in the neuron, this relatively long period of a time fails to differentiate individual activity.

EEG is an indirect measure of neural firing as it measures the summation of this activity at the level of the scalp. It does, however, record electrical activity, unlike fMRI. The blood oxygenation level dependent (BOLD) that is recorded using MRI does not measure electrical activity directly, but instead reflects the amount of oxygen present in the blood. This is based on the assumption that if a region of the brain is active, it will use and require more energy, and that more oxygenated blood will be supplied to it as a result (Ogawa et al., 1990). This means that active areas of the brain contain more oxygenated haemoglobin, and less de-oxygenated. It is assumed that this process is linked to neuronal firing through a process known as haemodynamic coupling (Logothetis, 2008). However, exactly how activity in one neuron relates to the amount of oxygenated blood supplied to that area is not completely known (Rosa et al., 2010).

Given that EEG has poor spatial resolution and high temporal resolution, whereas fMRI has high spatial resolution but low temporal resolution, it is un-

derstandable why researchers decided to combine the two modalities. The concurrent acquisition of EEG and fMRI has the ambitious aim of improving the spatial and temporal limitations of respective measures, promising increased understanding of brain function. Perhaps the most useful application has been the improved localisation of epileptic seizures in epilepsy patients, where increased spiking in the EEG can be correlated with activation in contributing brain areas. However, simultaneous EEG-fMRI has become increasingly used to investigate brain activation in healthy subjects and a range of methods have been proposed for data integration (for other reviews see Jorge et al., 2014, Huster et al., 2012, Ritter et al., 2006, Laufs et al., 2012). This review will focus on the challenges faced when recording simultaneous EEG-fMRI, including the nature of the signals that we record and when we can expect them to overlap. We argue that simultaneous EEG-fMRI is not always necessary, and present a flow chart of questions that researchers should ask themselves before recording EEG and fMRI simultaneously rather than in separate experimental sessions. Although this list is not exhaustive, we hope that it will provide a good starting point for any researcher considering combined EEG-fMRI recording.

6.2 When is combined EEG-fMRI necessary?

It is tempting for researchers to turn to combined EEG-fMRI without considering whether two separate experiments would be sufficient to answer their research question. This may be due to the naive theory that 'more data is better', that combined EEG-fMRI can always improve their knowledge of a particular phenomenon, or simply that it is 'more impressive'. However, we argue that combined EEG-fMRI is not necessary in a large number of cases, and that researchers should be thoughtful about their research question and experimental design before jumping to record EEG and fMRI simultaneously.

We constructed a flow chart to help researchers decide whether simultane-

ous EEG-fMRI is necessary, or whether separate EEG and fMRI experiments would be more appropriate (see figure 6.1). We begin the flow chart with the assumption that researchers would like to acquire both fMRI data and EEG data, but are deciding whether these need to be recorded during the same experimental session, or separately (the decision whether to record just EEG OR just fMRI data would require an additional set of statements).

1) Do you expect the brain activity in response to your task to be represented in both EEG and fMRI data?

The first question that we ask the researcher is perhaps the most important, but also the most difficult to answer. A pre-requisite for combined EEG-fMRI is that you expect both modalities to capture and reflect the activation that you hypothesise to find. If you do not expect this, then there is little to gain from recording both data sets, either simultaneously or not. If this is the case, then we recommend that you refine your experimental design and hypotheses before coming back to this question.

For example, if you are looking at a visual ERP, such as the visual P1, then you can have a reasonable expectation that this activity will be represented in both EEG and fMRI data. This is because a large number of studies have previously localised the visual P1 to the extrastriate cortex, and we also know that BOLD activation can be expected in this area during visual tasks. By definition as an ERP, we know that we can find the P1 in the EEG data.

Conversely, if you are interested in amygdala activity then the answer is more complex. Although we can expect to find amygdala activation in fMRI data, the exact contribution of deep brain structures to EEG is unknown. As the signals degrade over space, and are subject to volume conduction (explained in more detail below), the ability to detect activity from regions below the cortex is highly debated for EEG. If you do not expect to record EEG activity that reflects amygdala processing, your region of interest, then you will not gain additional information from combined EEG-fMRI. It is possible

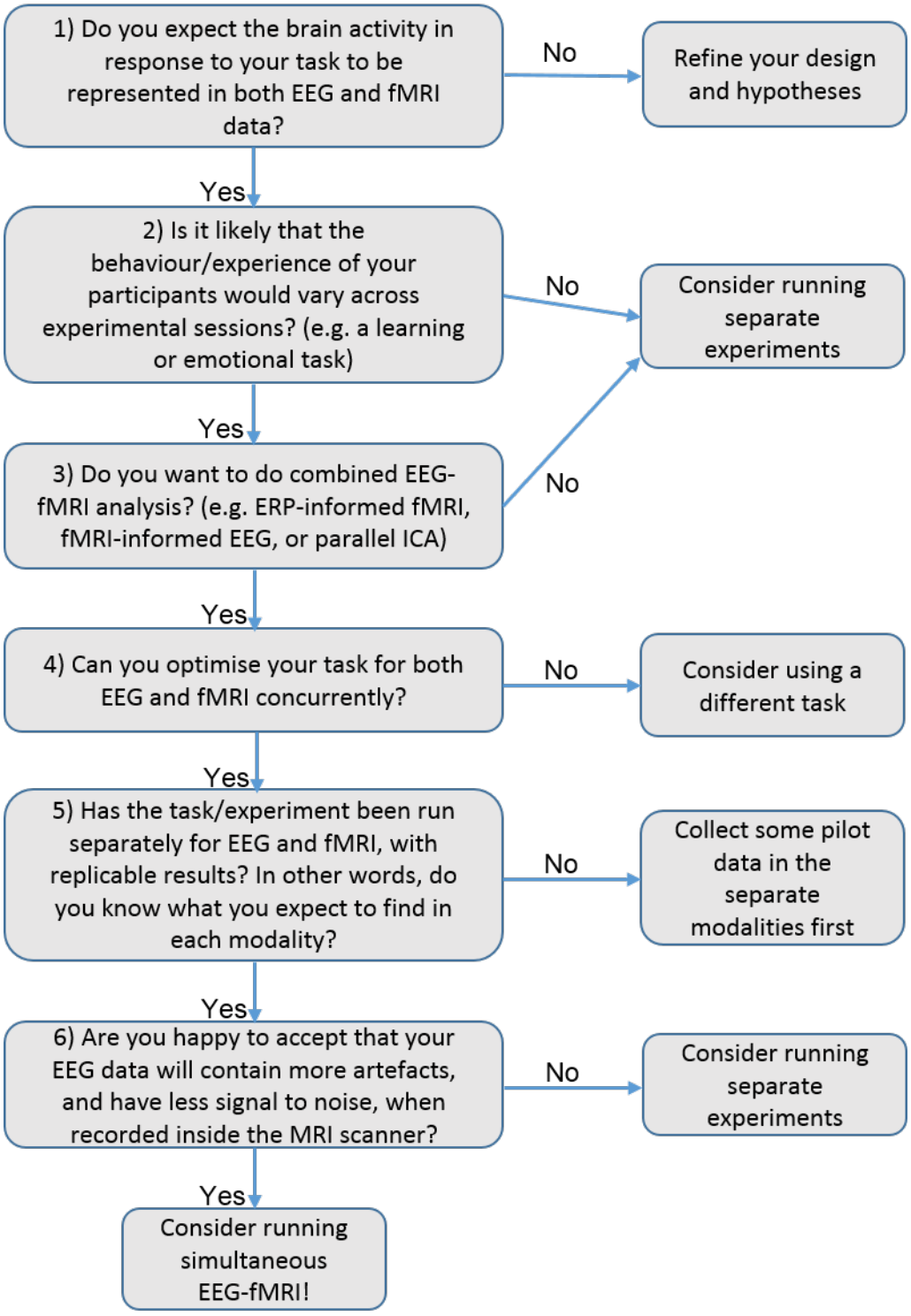

Figure 6.1: Flow chart: is simultaneous EEG-fMRI necessary?

that you will find other EEG signals that correlate with your amygdala activation, but without clear hypotheses you may find yourself wandering along a dangerous path of spurious correlations and multiple comparisons.

Importantly, even if you can reasonably assume that your activation of interest will be captured with both EEG and fMRI, there are a number of reasons why you may fail to detect any common signal. Below we summarise the biological signals that EEG and fMRI are thought to measure, as well as how/when they may not overlap.

Physiological Basis of EEG

An ideal neuroimaging recording would reveal activity at a neuronal level across all areas and structures of the brain with high temporal resolution. Without invasive recording, however, the greatest temporal resolution is achieved using EEG, which measures electrical activity at the level of the scalp. The so called 'inverse problem' represents the difficulty of identifying sources within the brain, given that EEG at the scalp can often be explained by several different source distributions. This is complicated further by our inability to full identify the pathway between a singular action potential and the signal at an electrode.

Although contributing greatly to intracellular recordings, action potentials in the neuron have a small contribution to EEG due to their rapid decrease in strength outside of a 50 μm range (Henze et al., 2000). The summation of many spikes may be detectable when averaging over epochs in relation to external stimuli, but their contribution is still relatively small. The most widely accepted view is that local field potentials (LFPs) are the main source of EEG activity (Logothetis et al., 2001), reflecting the electrical potential in the extracellular space of gray matter. This is the summation of postsynaptic potentials from sections of the cortex, in which pyramidal neurons are aligned perpendicularly with the surface of the brain, therefore creating open fields.

Each section of active cortex contains many possible arrangements of excitatory and inhibitory potentials throughout each layer, but is represented by an average amplitude in EEG. A disadvantage of this, as demonstrated by animal studies, is that similar EEG topographies can be found despite vary-

ing underlying sink source distributions (Mégevand et al., 2008). Other cells creating open fields can also contribute to the signal (Tenke & Kayser, 2012), further complicating the inverse problem.

Closed fields are also possible if cells are organised in such a way as to cancel out each other's signal. In this case, their activity is not reflected in EEG. Invasive recordings suggest that cells must activate in synchrony over an area of at least 4-6 cm^2 to generate 10-20 μV at the surface (Murakami & Okada, 2006), ideally with consistent organisation.

It is still debated whether deep brain structures, such as the hippocampus or amygdala, contribute to the EEG signal with a large enough effect to be detected. Both simulation (Attal et al., 2007) and experimental studies (Lantz et al., 2001; Michel et al., 2004) suggest that signals from these areas can be identified, although criticisms arise around the accuracy of the localisation methods used. Studies aiming to identify activity related to deep brain structures may therefore need to be careful in the conclusions they draw from their results.

Volume Conduction

The pathway of electrical fields from a current generator travelling through biological tissue is known as volume conduction, which is complicated by the varying resistance of each possible medium. Electrical potentials have to pass through several mediums on their way from cells to EEG electrodes, including other brain tissue, cerebrospinal fluid (CSF) and the skull. Forward models of EEG require the modelling of conductivity within each of these regions, which is further complicated by individual differences. For example, thickness of CSF is typically age-related (Nunez & Silberstein, 2000), and skull shapes are variable. However, the use of individual MRI scans to inform skull models may help to improve model success (Fuchs et al., 2007).

The relative influence of surface versus deep structures on EEG signals is thought to be related to skull conductivity, which acts as a low pass filter

(Mulert & Lemieux, 2010). As a result, low frequency signals from deeper sources may be less influenced by the conductivity of the skull and should still be considered as contributors to EEG. It should be remembered, however, that signals from these sources travel further to reach the scalp and are therefore subject to greater tissue inhomogeneity, leading to a reduction in signal.

Physiological Basis of fMRI

It is also important to question the data that can be gained from fMRI, as it is an indirect measure of neuronal activity. Although not the only method available for MRI, we focus here on blood oxygenation level dependent (BOLD) functional imaging, which has a relatively unknown relationship with activity at the neuronal level. Based on the physical properties of oxygenated and deoxygenated blood, the BOLD contrast reveals areas of the brain in which oxygen is increased in concentration. The assumption is that oxygen is supplied to areas of high activity in order to facilitate neural functioning, however, the exact neurovascular coupling is debated (Rosa et al., 2010). Evaluations of the BOLD signal are based on how well haemodynamic changes reflect activity in the neurons and whether all neural processes are equally represented.

Neurovascular coupling refers to the relationship between neural activity and changes in cerebral blood flow (CBF). Neuronal functioning in the brain requires a constant supply of oxygen and glucose, from which energy is synthesised (ATP). This energy is used to maintain and restore the ionic gradients responsible for action potentials and is therefore crucial. Cerebral blood flow increases following oxygen and glucose consumption, in order to replenish cell supplies. This results in a net increase in oxygen levels in active areas, which is the basis of the BOLD signal (Attwell & Iadecola, 2002)

The exact mechanisms involved in neurovascular coupling are still debated. Originally, it was assumed that increases in CBF were in direct response to energy demand (Roy & Sherrington, 1890). In this approach, bi-products of respiration such as carbon dioxide trigger activity in smooth muscle cells,

resulting in vasoconstriction/dilation and therefore changing CBF. This is criticised, however, as providing oxygen after certain levels have been consumed may not guarantee sufficient supplies (Hillman, 2014).

A second approach suggests the involvement of neurotransmitters, in which astrocytes are responsible for glutamate cycling and the production of vasoactive agents such as GABA, acetylcholine, norepinephrine and serotonin (Raichle & Mintun, 2006). It has also been suggested that astrocytes provide temporary energy supplies in the form of lactate through anaerobic glycolysis. This model, known as the astrocyte-neuron lactate shuffle (Pellerin et al., 2007) is still subject to debate (Hillman, 2014). Our understanding of neurovascular coupling is further complicated by its variation across brain regions and cognitive processes (Huettel et al., 2004).

Neural Correlates of BOLD

Comparared to EEG, evidence suggests that BOLD is correlated most strongly with local field potentials (Logothetis, 2008). Estimations by Attwell & Iadecola (2002) suggest that over 70% of total energy consumption can be attributed to post-synaptic potentials, with only 10% devoted to action potentials themselves. This theory supports LFPs as the basis of BOLD, as metabolic changes probably reflect the synaptic processes consuming the most oxygen.

However, the relationship between LFPs and BOLD appears to be region dependent. For example, several animal studies have identified a lack of correlation between LFPs and hippocampal BOLD signal (Ekstrom et al., 2008; Angenstein et al., 2009). In comparison, studies analysing the parietal and frontal cortex often report LFP/BOLD correlations (Scheeringa et al., 2008).

The correlation between spike rate and LFPs under certain conditions indicates that, in some instances, spike rate may also contribute to BOLD (Rees et al., 2000). There is some evidence to suggest that these two signals exhibit high dependence in the sensory cortex (Heeger & Ress, 2002) but large independence in midbrain regions (Kraskov et al., 2007).

Overall, it can be argued that the signals contributing to BOLD, and their relationship with metabolic processes, are dependent on region and behaviour (Huettel et al., 2004). This reflects the varying vascular architecture and functioning across areas of the brain, as well as their specific energy consumption and distance from draining vessels. All of these factors have implications on experimental design and analysis, as well as the conclusions about neuronal activity that can be drawn from BOLD data.

How do EEG and fMRI Overlap?

Rosa et al. (2010) suggest that observed discrepancies between modalities occur due to decoupling between electrophysiological and haemodynamic activity, or to signal failure (false positive/negative results). Another suggestion from Nunez & Silberstein (2000) is that the cell populations responsible for changes in EEG and fMRI do not always overlap, providing null results for co-localisation. There may therefore be instances in which correspondence between modalities fails.

For example, it is possible that the cell population contributing an EEG signal is not co-localised with the vascular branch supplying blood to the neurons, so the change in CBF would occur at a different location. Attempting to localise both EEG and fMRI to the same source would therefore produce uninformative results. However, both of the signals may indirectly reflect the same source of activity, and therefore produce correspondence between the dynamics of the data. In this case, correlation analysis between time series may be more be appropriate.

Due to the fact that EEG and fMRI are sensitive to activity at different spatial and temporal scales, it is likely that their representation will be unequal across conditions (Nunez & Silberstein, 2000). There are several situations in which it is possible for responses in fMRI to be identified in the absence of EEG:

- Areas with high metabolic load may contribute to BOLD but not to EEG (Ritter & Villringer, 2006). Processes independent of cognitive functioning can also contribute, such as maintaining resting potential and neurotransmitter synthesis (Patel et al., 2004).

- Activation detectable in BOLD can be invisible at the scalp level, if the electrophysiological activity is non-synchronised or forms a closed source. For example, Stellate cells have a high metabolic load but are non-pyramidal and do not contribute to EEG (Connors & Gutnick, 1990).

- Activity in deep structures may be unrepresented in EEG, due to volume conduction and signal decay (Henze et al., 2000).

Similarly, EEG detection in the absence of BOLD is plausible:

- Low energy consuming processes represented in EEG may not result in haemodynamic changes detectable in fMRI.

- Dynamics with high temporal resolution may be missed or smoothed by BOLD due to the slow sampling rate.

- Signals may fluctuate in opposite directions, as seen with the alpha rhythm. For example, reduced inhibitory cell activity can decrease metabolic load but increase pyramidal activity (Nunez & Silberstein, 2000).

Relationship with Stimuli

Although reflecting neuronal activity to varying extents, both EEG and fMRI represent individually biased views on brain activity as a whole. For example, it is possible for correlations to occur between behaviour and each modality separately, without any overlap in location between EEG and fMRI themselves (Nunez & Silberstein, 2000; Rosa et al., 2010). A clear distinction can therefore be made between a) neural activity that reflects processes from the same cell population (a localisation approach), and b) neural activity that

reflects a network of populations responding to the same stimuli or event (a stimuli-related approach). The preference of the experimenter for one or other approach will influence decisions made during experimental design, analysis and interpretation.

Researchers should have some idea that their signal of interest is detectable with both EEG and fMRI, from previous research or knowledge of the source location. If this is not the case, then simultaneous EEG-fMRI may not provide any additional information, and relationships between the EEG and fMRI signals may not be detectable.

2) Is it likely that the behaviour/experience of your participants would vary across experimental sessions? (e.g. a learning or emotional task)

Question 2 asks the researcher if it is likely that the behaviour/experience of their participants would vary across experimental sessions, for example in a learning or emotional task. If this is the case, then even if identical paradigms were recorded using the same participants across separate EEG and fMRI experiments, it would be difficult to ensure that the brain activity at each individual trial was the same. We would therefore recommend simultaneous recording, as a strong argument for the robustness of an effect and homogeneity over trials would be needed to justify comparisons across different experimental sessions.

3) Do you want to do combined EEG-fMRI analysis? (e.g. ERP-informed fMRI, fMRI-informed EEG, or parallel ICA)

Question 3 asks the researcher if they wish to run combined EEG-fMRI analysis, for example ERP-informed fMRI, fMRI informed EEG, or parallel ICA. These analyses require both data sets to be entered into the same analysis, and it is therefore necessary for the data to reflect the same brain activity. The previous question is therefore equally relevant here; if you assume that the

activity at each trial will vary across experimental sessions, then simultaneous recording is necessary. For example, if you are using a task where the response of the participant determines the trial type/condition.

In comparison, if the researcher plans to run separate EEG analysis and fMRI analysis, then combined EEG-fMRI may not be necessary. Traditional EEG analysis uses averages, and therefore the variance over single trials is not as important. If an ERP effect is robust (for example the P300), then the average ERP results are likely to be reproducible over a number of separate experimental sessions. Similarly, if you expect a similar average response over sessions, then you can record the fMRI in a separate session with a reasonable assumption that it represents similar brain activation to that recorded with the EEG.

Analysis methods

Assuming that you have acquired all of the relevant hardware/software required for simultaneous EEG-fMRI, and are confident that you are able to remove artifacts in a satisfactory manner, the next challenge to be overcome is the decision regarding, and application of, the analysis method. The existing analysis methods can be broadly grouped into two categories; symmetrical and asymmetrical analysis, of which the latter is most popular given its (relative) simplicity (see Huster et al., 2012 for a review).

In asymmetrical analysis, one modality is used as a predictor for the other. For example, ERP-informed fMRI uses ERP values extracted from single trials as a parametric regressor in a standard fMRI GLM analysis (general linear model). This identifies voxels with activation that co-varies with fluctuation in the EEG signal over trials, in a similar way that researchers could use reaction times or task difficulty. A variation of this method uses single trials values estimated in the frequency domain, for example alpha power. In the opposite direction, fMRI-informed EEG uses BOLD fMRI results to guide and/or constrain EEG source localisation. In theory, this enables a more informed source

localisation, given that areas relevant to the task can be identified using the standard GLM approach.

In symmetrical analysis, researchers avoid giving preference to one modality by modelling relationships between the data or calculating joint independent components (Moosmann et al., 2008; Daunizeau et al., 2007). Current symmetrical analysis can be divided into data-driven and model-based methods. ICA (Calhoun et al., 2006) and information theory (Ostwald & Bagshaw, 2011) are data-driven, as they do not require modelling of hemodynamics or neurovascular coupling. In contrast, dynamic causal modelling (Friston et al., 2003) and other model-based methods attempt to determine the underlying neural components of EEG and fMRI using individual forward models. Complete fusion of the two data sets would require a model that accurately maps neuronal, metabolic, and haemodynamic responses, accounting for neurovascular coupling and electrical propagation through the brain. The complexity of this task, accompanied by the lack of agreement on even small scale models of neural activation, has so far prevented its realisation (Rosa et al., 2010).

Singe-trial analysis

Many EEG-informed studies use single-trial values taken from the EEG signal. This deviates from tradition methods for EEG analysis which typically compares averaged ERPs across conditions. When used as a parametric regressor, these amplitudes can inform researchers of the temporal sequence of BOLD activation (Mulert et al., 2008). The majority of concurrent EEG-fMRI studies with combined analysis use single-trial ERP estimates, rather than averaged amplitudes for each condition. This provides a distinct time-series of EEG activation in response to participant behaviour that can be compared to BOLD activation. It could be argued that taking an average ERP amplitude per condition would remove the requirement for concurrent recording, especially if the ERP in question is robust. Single-trial analysis therefore rationalises the combination of imaging modalities, as fluctuations on this scale could not be

examined accurately with data taken from different recording sessions. For example, studies focusing on participant error and subsequent ERN amplitudes would be incapable of examining the neural implications of making a mistake on following trials, as performance is likely to fluctuate across testing sessions.

Similarly, it is unlikely that emotional ERPs will remain constant across time, since they can be influenced by participant mood, environmental setting, habituation to the task, and other potential confounds. Another validation of single-trial analysis lies in the ability to separate individual ERPs within each trial, and therefore determine individual ERP correlates within BOLD activation. Several studies have utilised this method, which may not be possible using averaged trial information. Debener et al. (2006) argues that variation over single trials in simultaneous EEG-fMRI represents behaviourally relevant activation, and is therefore significant for analysis and understanding. Similarly, Snyder & Raichle (2010) suggest that this type of analysis validates the use of simultaneous EEG-fMRI acquisition, as the focus is the trial-by-trial variation demonstrated by individual subjects.

In an assessment of the extent to which single-trial variation in EEG reflects changes in BOLD activation, Bagshaw & Warbrick (2007) measured participant responses to a visual checkerboard. Comparable results were found using three different approaches to the extraction of single-trial values (independent component analysis, wavelet denoising and GLM), as well as a correlation between ERP and BOLD latencies. They argue that the strong agreement between EEG and MRI data sets supports the assumption that they reflect common neural activity at a single-trial level, therefore validating the use of this method for simultaneously recorded signals. However, this study used separately recorded EEG and fMRI and relied only on correlations between features of both data sets (latency and amplitude) for comparison.

Conversely, De Vos et al. (2012) question the generalisabilty and accuracy of single-trial analysis. In an assessment of frequently used methods for determining single-trial ERP amplitudes, including ICA and regression-based

estimation, the researchers tested their ability to identify the face-sensitive N170. The researchers concluded that ICA was the best method to discriminate between single-trial peak fluctuations, but still questioned its ability to identify smaller ERPs, such as visual P1. Currently the best way to overcome these issues is to ensure that enough trials are recorded to provide a reliable estimate of trial-to-trial fluctuations.

Other arguments against ERP-informed fMRI are based on the possible divergence between measurements, and the likelihood that they represent different views of the same neural activity. Eichele et al. (2008) argue that the spatial and temporal mixing of both signals results in the possibility that an ERP predictor and the associated BOLD response do no co-vary. This can lead to false-negative predictions if the EEG predictor is the summation of two or more sources with differing dynamics (Moosmann et al., 2008). Arguably, this is possible with all EEG-fMRI methods of analysis. However, Moosmann et al. (2008) suggest that ERP based analysis is more susceptible to false-negatives, due to the small proportion of information that is used from the available EEG data.

Overall, single-trial ERP-informed fMRI is a useful method of analysis, provided that the underlying assumptions are not forgotten in the interpretation of results. For example, fMRI voxels that are found to correlate with single-trial P300 results are not necessarily the exact neural source of this ERP. The fact that these voxels are correlated with trial-to-trial fluctuations in EEG can suggest that they have similar dynamics, where EEG and fMRI may respond in a similar way to experimental stimulation. However, it is possible that neither method represents the true source of neuronal activity, and therefore suggesting that the P300 response is formulated in these voxels is misleading. Instead, it can be said that these voxels co-vary with ERP amplitudes and therefore that they are involved in the same pattern/modulation of activation.

Previous literature

As a useful reference, we conducted a literature search for all papers reporting simultaneous EEG-fMRI recording between the years of 2000-2015. In the tables below, we summarise the EEG feature of interest, the experimental task, and analysis method used in each study. Papers with a medical focus, for example epilepsy, are not included. Table 6.1 summarises all papers using time domain ERP analysis, whereas table 6.2 includes all papers using frequency based analysis.

4) Can you optimise your task for both EEG and fMRI concurrently?

Question 4 is important for the design of a combined EEG-fMRI experiment, and asks if the task can be optimised for both EEG and fMRI. Traditional EEG experiments are fast with short trial durations, given that researchers are usually interested in activity within the first 600 ms after stimulus onset. The BOLD response measured using fMRI is much slower, and therefore fast paradigms must be designed with caution. Any task where presentation, responses, and inter-trial intervals cannot be varied across trials may not be suitable for fMRI. Although most paradigms can be modified to suit fMRI, researchers more familiar with EEG should seek advice from fMRI specialists and be cautious in their experimental design.

5) Has the task/experiment been run separately for EEG and fMRI, with replicable results?

Question 5 asks if the task, once agreed on, has been run separately for EEG and fMRI, with replicable results. This question is very similar to question 1, which asks the researcher if their signal of interest is measurable using both EEG and fMRI. However, we wish to emphasise the importance of checking whether you can identify your desired signals in EEG and fMRI data individually. If previous research has found measurable signals in both data sets using

a similar paradigm, then you may not need to run these experiments again. However, if you have designed a new paradigm, or plan to use a new analysis method, we strongly recommend checking that your expected signal can be found in both EEG and fMRI before running combined recording.

A further point to emphasise here is that researchers should have some idea of what they expect to find in both modalities using their analysis technique(s). When confronted with a large data set from simultaneous recording, it is important to have some idea of the analysis pipelines that will be run and what results are reasonable to expect. With more data comes an increased chance of spurious results, and without a clear direction for analysis it is possible to find *something*, even if not meaningful. Researchers should have an understanding of the data from both modalities before attempting to combine it.

6) Are you happy to accept that your EEG data will contain more artifacts, and contain less signal to noise, when recorded inside the MRI scanner?

The final question asks researchers if they are happy to accept that their EEG data will contain more artifacts, and have less signal to noise, when recorded inside the MRI scanner. We appreciate that this statement is vague, and it is not possible for anyone to quantify the loss of signal that will occur for a given paradigm or EEG feature.

When EEG is recorded during MRI acquisition, a number of additional EEG artifacts are incurred. Perhaps the most simple to remove is that caused by the gradient pulse, and therefore known as the gradient artifact. As this is related to the sequence of the MRI scanner, which is known, and is stable over time, the gradient artifact can be successfully removed by subtracting a template of its form (see Allen 2000).

The second artifact is the ballistocardiogram (BCG), which is considerably harder to remove, and a reliable solution to this is yet to be achieved. The BCG artifact is related to the heartbeat of the participant lying down in

the scanner. More specifically, expansions and contractions of arteries in the scalp causes movements in the electrodes and wires in the EEG cap (Goldman et al., 2000). This movement of blood also influences the static magnetic field and can result in artifacts with larger power than the EEG signal of interest (Ritter & Villringer, 2006). With similarities to the removal of the gradient artifact, one method used to remove the BCG is to construct a template of the heartbeat artifact, identify its occurrence across the recording, and subtract it from the EEG signal (Allen et al., 1998). This requires an additional electrode placed on the participant's back to record their heartbeat. This is the method implemented in BrainVision analyzer (Brain Products), and is frequently used.

However, there are a number of factors that reduce the success of this method. Unlike the gradient artifact, the heartbeat of the participant is not stable over time. They may end the experiment in a more relaxed state than they began, causing a difference in the frequency of their heartbeat. Even if relaxed, their heartbeat may vary in speed across the duration of the recording. Another problem is caused by movement of the participant which may dislodge the electrode positioned on their back and disrupt the signal. If you cannot measure the heartbeat then you cannot identify where the peaks occur, and it becomes increasingly difficult to remove the artifact. A number of other methods for removing the BCG artifact have been suggested, for example ICA (Srivastava et al., 2005) and adaptive filtering based on a time varying finite impulse response (Bonmassar et al., 2001). However, no method claims to successfully removed all BGC artifacts for all participants.

The third artifact present in EEG recorded inside the MRI environment is caused by the helium pump, which results in widespread peaks across the frequency spectrum, far above the amplitude range of normal EEG (Mullinger et al., 2008). Given the spread across frequencies, and the difficulty in distinguishing true neural signal from helium pump noise, this can be difficult to remove. A further complication is the large between-site differences in helium artifact, driven by factors such as the scanner manufacturer and physical set

up. One way to avoid this artifact is to switch off the helium pump before running the experiment (Laufs et al., 2008). However, as the helium pump is essential for the continued functioning of the MRI scanner, this cannot be left switched off for long time periods.

Some researchers have also reported spurious correlations between EEG and fMRI signals that are related to motion artifacts, rather than a common neural source. For example, EEG power in the frequency domain was found to be significantly higher during trials with high motion, compared to low motion, especially in low frequency bands (Fellner et al., 2016).

6.3 Conclusions

In summary, we advise that simultaneous EEG-fMRI is unnecessary if; you plan to run traditional EEG and fMRI analysis separately, using averages over trials, rather than combined analysis and single-trial data; you assume that participant behaviour and neurological responses would be relatively stable across experimental sessions; if you do not expect your signal of interest to be detected in both EEG and fMRI signals; and if you cannot find a suitable paradigm that can be optimised for EEG and fMRI concurrently. If you do not know what to expect in the individual modalities or how you would analyse your data, we suggest that you run pilot studies in each modality first before coming back the question of simultaneous EEG-fMRI.

In comparison, simultaneous EEG-fMRI is necessary if; you plan to run combined EEG-fMRI analysis with the assumption that the same behaviour and neurological activity is represented in both modalities at each trial; if you expect that this behaviour and neurological activity would vary across experimental sessions; if you can reasonably expect that your signal of interest to be detected in both EEG and fMRI signals; if you know what to expect from each modality individually and are therefore interested in running combined analysis to extract more information.

Lead author	Date	ERP(s)	Task	Analysis method
Bonmassar	2001	N75 N100	Checkerboard	fMRI informed EEG
Kruggel	2001	P1 N1	Oddball	Separate
Liebenthal	2003	MMN	Oddball	Separate
Mulert	2004	P300	Oddball	Localisation comparison
Scarff	2004	AEPs	None	Localisation comparison
Mulert	2005	N1	None	Localisation comparison
Debener	2005	ERN	Flanker	ERP informed fMRI
Henning	2005	P1 N1	Dot patterns	Separate
Otzenberger	2005	P300	Oddball	Separate
Benar	2007	ST P300	Oddball	ERP informed fMRI
Sabri	2007	MMN	Oddball	Separate
Strobel	2008	P300	Oddball	fMRI informed EEG
Mulert	2008	N1 P300	Forced choice	ERP informed fMRI
Sadeh	2010	N170	Attention	Correlation: face selectivity
Mayhew	2010	AEPs	None	ERP informed fMRI
Novitski	2011	P1 N1	Checkerboard	ERP informed fMRI
Hesselmann	2011	P3	Attention	ERP informed fMRI
Donamayor	2012	ERN	Flanker	fMRI informed EEG
Diukova	2012	P300	Oddball	Separate
Jaspers-Fayer	2012	EPN	Emotional	ERP informed fMRI
Regenbogen	2012	N1 P2	WM/Oddball	ERP informed fMRI
Plichta	2013	CNV	Reward anticipation	ERP informed fMRI/DCM
Walz	2013	P3 N1 P2	Oddball	ERP informed fMRI
Geukes	2013	N400	Semantic priming	Localisation comparison
Wirsich	2014	N170	Face recognition	ERP informed fMRI
Eichele	2014	ST	Oddball	ERP informed fMRI
Walz	2014	P3	Oddball	ERP informed fMRI
Karch	2014	N2	Go/no-go	ERP informed fMRI
Nguyen	2014 a	N170	Face recognition	ERP informed fMRI
	2014 b	N170	Face recognition	DCM
Baumeister	2014	N2 P3	Flanker/no-go	ERP informed fMRI
Iannacconne	2015	ERN	Flanker	ERP informed fMRI

Table 6.1: Simultaneous EEG-fMRI experiments using ERP components for analysis. MMN = mismatch negativity, ERN = error related negativity, AEPs = auditory evoked potentials, EPN = early posterior negativity, CNV= contingent negative variation, DCM = dynamic causal modelling.

Lead author	Date	Frequency(s)	Task	Analysis method
Goldman	2002	Alpha	Resting state	EEG informed fMRI
Laufs	2003	Alpha	Resting state	EEG informed fMRI
Goncalves	2006	Alpha	Resting state	EEG informed fMRI
Horovitz	2007	Theta Alpha	Sleep	EEG informed fMRI
Mantini	2007	A B G	Resting state	EEG informed fMRI
Scheeringa	2008	Theta	Resting state	EEG informed fMRI
Ritter	2009	Alpha Beta	Action	EEG informed fMRI
Jann	2009	Alpha	Resting state	EEG informed fMRI
Wu	2010	Alpha	Resting state	EEG informed fMRI
Michels	2010	A B G	Working memory	EEG informed fMRI
Sadaghiani	2010	Alpha	Resting state	EEG informed fMRI
Britz	2010	Microstates	Resting state	EEG informed fMRI
Hanslmayr	2011	T A B	Memory	EEG informed fMRI
Picchioni	2011	Infraslow	Sleep	EEG informed fMRI
Bayram	2011	A B G	SSVEP	EEG informed fMRI
Scheeringa	2011	Alpha	Attention	EEG informed fMRI
Yuan	2012	Microstates	Resting state	EEG informed fMRI
Lehongre	2013	D T G	Passive viewing	EEG informed fMRI
Bridwell	2013	T A B G	Resting state	EEG informed fMRI
Omata	2013	Alpha	Resting state	EEG informed fMRI
Meyer	2013	Microstates	Resting state	EEG informed fMRI
Razavi	2013	D T A B	Resting state	EEG informed fMRI
Liu	2014	Alpha	Visual attention	EEG informed fMRI
Scharinger	2014	Alpha	Auditory categorisation	EEG informed fMRI

Table 6.2: Simultaneous EEG-fMRI experiments using frequency components for analysis. DMN = default mode network, RSN = resting state network, SSVEP = steady state visual evoked potential, D = delta, T = theta, A = alpha, B = beta, G = gamma.

The final chapter offers an overall discussion of our findings, explains how these fit into the wider literature, and highlights the main contributions to the field.

Chapter 7

General discussion and conclusions

7.1 Aims

The overall aim of this thesis was to explore the existence and nature of the *sense* condition in the change blindness paradigm, using a combination of EEG, fMRI, and behavioural measures. More specifically, we had the following theoretical and methodological aims:

1. To build upon existing EEG results comparing the *sense* condition with other levels of awareness, such as when participants are completely *blind* to the change

2. To ascertain whether the brain activity related to the *sense* condition is different from that related to other levels of awareness, using fMRI

3. To identify brain areas with activity that co-varies with fluctuations in the EEG signal, using ERP-informed fMRI analysis

4. To investigate the influence of the MRI environment on EEG and behavioural results

To achieve these aims we conducted two experiments, one using EEG alone (described in Chapter 3) and the other using simultaneous EEG-fMRI (study pre-registration described in Chapter 4, and results in Chapter 5). We then collated the data from both experiments to run a combined analysis and examine the influence of the MRI environment on EEG and behavioural results

(described in Chapter 6). In addition, this thesis contains a literature review of previous papers investigating the *sense* condition (described in Chapter 2), as well as a literature review detailing the use of combined EEG-fMRI (described in Chapter 7). This discussion summarises the main results found across all chapters, comments on the implications of these results, and suggests avenues for future research.

7.2 Behavioural results

7.2.1 Task difficulty and individual differences

As outlined in the literature review in Chapter 2, change blindness paradigms often result in uneven trial distributions and a low number of *sense* trials. This is due to individual differences in response strategy, variable working memory performance, and the nature of the *sense* condition as an uncertain level of awareness. We therefore attempted to control for individual differences in task performance by varying the task difficulty over time, based on the assumption that increasing the set size would reduce accuracy. By adjusting the difficulty in real time in relation to participants' responses, we attempted to prevent floor or ceiling performance and increase the number of *sense* trials available for analysis. This was achieved as all of our participants were able to *sense* changes during the task, unlike in other experiments (Ball & Busch, 2015; Thornton & Fernandez-Duque, 2001).

In the EEG only experiment, we found no correlation between the mean task difficulty and performance on the task (detection accuracy, location accuracy, or d'prime), suggesting that increasing the difficulty did not inhibit success on the task. However, in the EEG-fMRI experiment we found that mean task difficulty correlated with both location accuracy and d'prime (but not detection accuracy). In the combined analysis with all 40 data sets, we also found a correlation between mean task difficulty and location accuracy.

Intuitively, one would expect a negative relationship between difficulty and

localisation accuracy, as increasing the number of items on the display should make it harder to identify the object that changed. Instead, a positive correlation was found, as participants who were better at the task received a more difficult version of the paradigm. The task difficulty was also adjusted based on the localisation response, so a relationship between these two variables can be expected.

The fact that several participants were performing at 80% localisation accuracy despite receiving the most difficult version of the paradigm highlights the need to control for individual performance. Had these participants been completing the task at a lower difficulty level, their accuracy would have been at ceiling, and therefore few *sense* or *false alarm* trials would have been available for analysis. These results also corroborate other literature suggesting large individual differences in performance on working memory tasks (Vogel et al., 2005; Luck & Vogel, 2013).

The number of *sense* trials increase as a task becomes more difficult, and can be reduced to chance level if a task is too simple (Ball & Busch, 2015). However, this mirrors the decrease in rates of full awareness. When given more items to store and compare across image presentations, observers become worse at describing details of a change, and are more likely to report *sensing*. This suggests that *sensing* changes may be a useful mechanism when we are overloaded with information, such that our explicit knowledge is limited. Attention can only be directed towards a subset of items in our visual world, but other items may be equally important for our survival.

Although important to establish a range of trial types within participants, it was a concern that adjusting the task difficulty could produce confounds in the EEG. It is known that the properties of the visual stimulus can influence ERPs within the recording, especially visual P1/N1, and it is therefore possible that the range of visual stimuli presented across the experiment could introduce a confound (Luck, 2014). This is particularly relevant given that the task difficulty influences performance, and subsequently trial categorisa-

tion; changes during easy trials are more likely to be detected and localised correctly, whereas those in difficult trials are more likely to be missed. To investigate this we correlated single-trial ERP amplitudes with task difficulty over trials. After correcting for the large number of comparisons we found no significant relationships between task difficulty and ERP amplitudes. However, as considered in the EEG-fMRI chapter discussion, it is possible that our method for extracting single-trial values was not optimal, and future studies could employ different methods (for example, using single-trial values derived using ICA).

As far as we are aware, we are the first to control individual difficulty in real time when investigating the *sense* condition, and therefore these findings are a valuable contribution to the literature.

7.2.2 Comparison of *sense* and *false alarm* trials

Across the EEG only and EEG-fMRI experiment and the combined analysis, participants had more *sense* trials than *false alarms*. This suggests a distinction between these two trial types, as if *sense* trials were simply incorrect responses then you would expect a balanced number of incorrect responses over change and no-change trials. Additional evidence for the distinction comes from the increased reaction times for *sense certain* trials compared to *false alarms* across all data sets.

However, in the EEG and combined analysis the percentages of *sense* and *false alarm* trials were correlated, such that participants with more *sense* trials also had more *false alarms*. This could provide evidence that *sensing* occurs due to a liberal response criterion, as suggested by (Mitroff et al., 2002). Indeed, participants who were more willing to report that they detected a change in the face of uncertainty (as they did not know the location of the change) were also more likely to make mistakes, detecting a change during a no-change trial. It is unclear whether this causes a problem for the investigation of the *sense* condition, or if this is an inevitable occurrence when adopting a liberal

response criterion.

7.2.3 Comparison of *blind* trials and no-change trials

In the combined analysis, 75% of participants were slower to respond when they were *blind* during a change trial, compared to no-change trials, and the difference in reaction times was significant across all participants. The difference was greater still when comparing *blind uncertain* trials to no-change trials, as these trials had even slower reaction times. This suggests that the presence of a change slowed down observer responses, even when they failed to detect it, and corroborates other research using the registration without detection definition of *sensing* where *blind* and no-change trials are compared. If participants had no information about the change when they failed to detect it, then *blind* trials should have a similar reaction time to no-change trials. The repeated finding that reaction times are increased provides evidence that behaviour is influenced in some way by the presence of the change, even if undetected.

7.2.4 Comparison of *sense* and *blind* trials

Across all data sets, reaction times for *sense certain* trials were significantly faster than *blind* trials. Participants were therefore slower when they failed to detect the change, suggesting that these trial types are separable in the behavioural data. Perhaps the strongest evidence for the distinction between these trials comes from the fMRI results, where a network of visual and parietal activation was found for the contrast *sense* > *blind* (to be discussed in more detail below).

7.3 EEG results

7.3.1 Summary

One aim of this thesis was to build upon existing EEG results comparing the *sense* condition to other levels of awareness, such as when participants are completely *blind* to the change. To achieve this, we conducted two experiments, one using EEG alone, and the other using simultaneous EEG-fMRI. The EEG results reported across these experiments were variable. In the EEG only experiment, significant differences between all awareness conditions were found in the late positivity ERP (LP). Additionally, both *localise* and *sense* conditions had a significantly larger N2pc than *blind* trials. No effects were found in the VAN, P1, or N1.

In the EEG-fMRI experiment, the LP for *localise* trials was significantly greater than *blind*, as before. However, the *sense* condition was no longer separable from *blind*. No effect of hemisphere was found in the N2pc window, and only a general increased in negativity was found between *localise* and *blind* conditions. Again, no effects were found in the VAN, P1, or N1.

When all data was included in a combined analysis, the results more closely matched the EEG only experiment, perhaps due to the increased power with a larger sample size. Both *localise* and *sense* conditions had significantly larger LP and N2pc amplitudes than *blind*. There were no significant differences between *localise* and *sense* conditions. No effects were found for the VAN. However, additional effects were found within the visual ERPs. There was a main effect of awareness on the N1 amplitude, with a larger amplitude for both *localise* and *sense* conditions compared to *blind*. There was also an interaction between awareness and display found for the P1, as *blind* trials had a larger amplitude in the first display whereas *localise* and *sense* amplitudes were larger in the second display.

7.3.2 LP

Across all analyses, our results indicate a larger LP amplitude for both *sense* and *localise* conditions when compared to *blind* trials. Although we also found a significant difference between *localise* and *sense* in the EEG only experiment, this finding was not replicated. As it was also close to the threshold of significance, $p = .048$, we will not draw any further inference from this finding.

An increased late positivity for change detected trials versus *blind* trials is the most commonly reported finding within the EEG literature, and all of the papers considered in the review by Koivisto et al. (2010) report this finding. This may be due to the relatively large size of this ERP, peaking anywhere between 300 and 700ms after a change stimulus and across large time windows. This could explain why the effect remained in the EEG-fMRI data set, despite the additional noise and artifacts caused by the MRI environment.

Previous literature has suggested that the LP reflects conscious aspects of task processing and is generally associated with complete or 'access' awareness (Lamme, 2004). However, we found an LP for both *localise* and *sense* trials, even though participants were not able to report the location of the change in the *sense* condition. One explanation for this is that full awareness is not necessary for the LP, and that any knowledge of the change is sufficient to provide an increase in the LP time window. This finding was also reported by Busch et al. (2009), although they refer to this ERP as a P3.

A second explanation is that our *sense* trials are simply *localise* trials with a response error. Given that we were unable to provide any distinction between these trial types in behavioural, EEG, or fMRI data, this explanation remains plausible. The phenomenological experience of participants suggests that a version of the *sense* condition does exist, but whether we have successfully measured this with our paradigm is unclear.

7.3.3 N2pc

The majority of change blindness papers listed by Koivisto (2010) reported enhanced negativity in the N1-N2 range (with the exception of Fernandez-Duque et al. 2003; Niedeggen et al. 2001). Busch et al. (2010) found that an N2pc was evoked only when the change was fully identified, and not in the *sense* or *blind* conditions. Based on this, they draw the conclusion that for *sense* trials, the change did not induce a shift in attention towards the location of the change, and therefore the features of the change were not available for further recognition. This is based on the assumption that the N2pc represents the allocation of attention towards the object of interest, which is supported by a number of previous studies (Luck & Ford, 1998).

In our EEG experiment, we found that both awareness conditions were significantly different to *blind* trials, indicating a shift in the allocation of attention for all identified changes, regardless of subsequent success/failure to localise. It may therefore be that *sense* trials elicited a shift in attention to the correct hemifield of change (and subsequently an N2pc), but that it was not specific enough to determine whether the change occurred in the upper or lower field within that hemifield. Woodman and Luck (2003) also identified an N2pc for 'unaware' stimuli which were obscured by object substitution masking, suggesting that the N2pc does not necessarily represent conscious awareness of changes (Woodman & Luck, 2003). It is suggested, however, that the amplitude is increased for 'aware' stimuli (Schankin & Wascher, 2007), which our findings support.

In the EEG-fMRI experiment, we failed to find a significant contralateral negativity or N2pc. Instead, we found general negativity increase for *localise* trials compared to *blind*. One explanation for this is the reduced signal to noise ratio. However, the results from the combined analysis matched the EEG results, and both *localise* and *sense* trials had a significantly greater N2pc than *blind*. Hence, we can conclude that, overall, both awareness conditions

produced a shift in attention towards the location of the change, as indexed by the increased contralateral negativity (N2pc). Therefore, complete awareness was not necessary for the increase in N2pc, although the amplitude was greater when changes were both detected and localised.

7.3.4 VAN

Although there was a main effect of awareness within the VAN in the EEG only experiment, the corrected post-hoc tests were not significant, and only *localise* was significantly different to *blind* using an uncorrected threshold ($p = .044$). We also failed to find any other effects within the VAN time window across experiments.

In comparison, (Busch et al., 2010) identified a VAN for their *sense* condition, compared to *blind*. The VAN is thought to be dependent on spatial attention, and requires both the location and identity of an object to be stored such that it is available for conscious report (Koivisto et al., 2008). As participants were not able to identify the location of change in our *sense* condition, this may explain the lack of significant VAN ERP. In another study (Koivisto et al., 2008), VAN was found to be reduced when participants were asked to keep their eyes fixated at the centre of the screen. This was the case in this experiment, which may also have contributed to the lack of significant finding within the VAN window.

Although our choice of the time window and electrode sites were based on previous literature, it is possible that the VAN was present at a different location. As with all ERPs, a range of electrodes are used across the literature, making informed decisions difficult. The VAN in particular had large variation across reported studies. All of our analyses were decided before data collection began, based on recommendations from Luck et al. (2017), in an attempt to control the rate of false positives. However, post-hoc multivariate analysis could be beneficial to detect any additional effects at other sites or time windows.

7.3.5 P1

Significant effects for the visual P1 and N1 were only found in the combined analysis, possibly due to the increase in participants and statistical power. The lack of consensus in the literature may therefore be due to low participant numbers that fail to detect small effects in these visual ERPs (Koivisto & Revonsuo, 2010).

Our P1 results contradict findings from Pourtois et al. (2006) who found a larger P1 at the first display for change detected trials. They suggest that the larger P1 during the first display reflected increased attentional resources allocated towards the display. This results in increased sensory processing of the first image, and therefore successful comparison to the changed image. As the P1 increase was bilateral, they suggest that this attentional increase was not specific to a particular area of the display, but instead indicated an improved encoding of the whole visual display. Other studies have also linked the attentional state of the observer prior to a change with the increased likelihood of detection. For example, the phase and amplitude of EEG alpha power has been found to predict task success (Hanslmayr et al., 2007; Dijk et al., 2008; Romei et al., 2010). However, a recent pre-registered study only found evidence for the effect of pre-stimulus alpha amplitudes on visual perception; alpha phase did not predict perception (Ruzzoli et al., 2019).

In our combined results, using the EEG data from both experiments, we found a larger P1 at the first display for change *blind* trials, rather than change detected trials. The P1 for detected changes (*sense* and *localise*) were smaller in amplitude. This is the opposite result to Pourtois et al. (2006). Additionally, there was an interaction between awareness and display found for the P1, as *blind* trials had a larger amplitude in the first display (compared to the second) whereas *localise* and *sense* had a larger P1 in the second display. During the second display, *sense* trials had the largest P1, followed by *blind*, and then *localise*.

These results do not support the hypothesis that an increased P1 reflects enhanced encoding of the pre-change display, as we found a larger P1 for trials where the change was missed. This is the opposite of what would be expected, as an enhanced representation of the pre-change display should facilitate change detection. We did, however, find a decrease in the P1 for *blind* trials during the second display. If we accept P1 as an reflection of attentional state, then this could indicate a decrease in attention during the second display that could be the cause of change blindness. It should be noted that this hypothesis is based on the explanations of change blindness related to display encoding, rather than display comparison (Simons, 2000).

7.3.6 N1

There was a main effect of awareness on the N1, with larger amplitudes for both *localise* and *sense* conditions compared to *blind*. This suggests an early increased negativity over the visual cortex for detected changes compared to those that were missed. This negativity was present even when participants failed to localise the change correctly.

The visual N1 ERP is greater in amplitude for stimuli that occur at an attended location, compared to an unattended location (Luck et al., 1990; Vogel & Luck, 2000). This ERP is therefore linked to the discrimination of visual information in the visual cortex, based on the location of attention. In short, attention leads to an increase in sensory input for items in the attended location, reflected by an increased N1 component (Martınez et al., 2006; Di Russo et al., 2003). This effect only occurs when participants are asked to discriminate between stimuli at the attended location, disappearing when only a button press is required (Luck et al., 1990; Vogel & Luck, 2000), suggesting that the N1 may also reflect discrimination between stimuli in an attended location.

We did not explicitly compare the N1 amplitudes for attended versus unattended locations, as we divided our trials based on awareness condition. However, we know that attention facilitates change detection (Rensink et al., 1997;

Simons, 2000), and therefore these two processes are related. It is possible that the increased N1 for *localised* trials reflects the allocation of attention to the change location, which resulted in successful detection and localisation of the change. The decreased amplitude for *blind* trials may reflect the opposite. If we accept this hypothesis, then the increased N1 for *sense* also indicates an allocation of attention towards the correct location, suggesting that some additional knowledge about the change was available to the participant. However, the participant did not report the correct change location. It should be noted that this hypothesis make large assumptions about the relationship between attentional allocation and change detection. Previous research indicates that changes occurring within the focus of attention can also be missed, and therefore this explanation of the N1 results is likely to be an oversimplification. It does, however, provide more evidence that *sense* trials are more similar to *localise* trials than *blind* trials, in terms their associated neurological signals.

7.3.7 Certainty

Unfortunately, the trial numbers collected prevented us from dividing our main awareness conditions into certain and uncertain trials for EEG analysis, as suggested by other researchers (Pourtois et al., 2006; Galpin et al., 2008). This was mostly due to the low number of *sense* trials for some participants, which could not have been divided into two further conditions for reliable EEG analyses. However, we ran a post-hoc EEG analysis that divided all trials based on the participant's response of certain and uncertain. Although all ERPs were examined (P1/N1/VAN/N2pc/LP), the only significant finding was a significantly increased LP for certain versus uncertain trials. This supports previous theories that this late ERP reflects subjective experience and certainty, rather than, specifically, awareness (Koivisto & Revonsuo, 2010; Eimer & Mazza, 2005; Railo et al., 2011; Pourtois et al., 2006). The certainty measure appears to be a useful way to separate trials in addition to the *blind*/*localise*/*sense* categorisation, as shown in our reaction time results, and future studies could

focus on obtaining higher trial numbers to facilitate full EEG analysis across both awareness and certainty.

7.3.8 Influence of the MRI environment

Another aim of this thesis was to investigate the influence of the MRI environment on EEG and behavioural results. When comparing ERP amplitudes across the EEG only EEG-fMRI experiments, we found an overall increase in amplitudes inside the MRI scanner. In particular, the first and second N1 had significantly larger amplitudes in the EEG-fMRI experiment. Almost all latencies were earlier in the EEG-fMRI data set, except for the N1 which was later. Although the mean variance in the single-trial ERPs was similar across experiments, the standard deviations of the variance across participants was greater in the EEG-fMRI data set, suggesting larger distributions of the single-trial ERP values. This could be caused by additional noise, differences in pre-processing steps, between-group differences, or a combination of the above.

7.4 fMRI results

A further aim of this thesis was to improve our knowledge of the neurological basis of the *sense* condition through the addition of fMRI results. We found largely overlapping activation for both *localise* and *sense* conditions when contrasted with trials where participants were *blind* to the change in coloured square. Both awareness conditions had significantly greater activation in the early visual cortex (B18, V2), the left supramarginal gyrus in the inferior parietal lobe (BA40), and the left pre-motor cortex (BA6). These results are similar to previously reported *see* versus *blind* contrasts in fMRI (Pessoa, 2004; Beck et al., 2001). In the conjunction analysis, the visual cortex and supramarginal gyrus were significantly activated for both contrasts.

The posterior parietal cortex and early visual cortex are commonly impli-

cated as storage sites for the contents of visual working memory (Todd and Marois, 2004; Edin et al., 2009; D'Esposito et al., 2015, Christophel et al., 2012), and previous fMRI studies of change detection also found activations in these areas (Beck et al., 2001; Pessoa, 2004). As discussed in Chapter 5, these results suggest that an inability to localise the change during *sense* trials may not be caused by a lack of visual representation, as activity in the dorsal stream (BA18 and BA40) was present for both *sense* and *localise* contrasts. If it is not the representation of the stimuli that failed, then another explanation of change blindness is required, for example the failure of stimuli comparison (Simons, 2000).

Activations found only in the *localise* contrast (but not for *sense*) were located in the primary sensory cortex (BA1), putamen (BA49), insula (BA13), and angular gyrus in the inferior parietal lobe (BA39). This forms a wider network of activation than the *sense* versus *blind* contrast, including mid-brain structures. The insula and putamen are both hypothesised to act as hubs in key brain networks relating to cognitive control, and their activation specific to *localise* trials may indicate their role in facilitating full awareness of the change (Menon & Uddin, 2010; Eckert et al., 2009; Uddin et al., 2017; van Belle et al., 2014; Karnath & Rorden, 2012).

In comparison, activation in the anterior cingulate cortex (ACC) was found exclusively in the *sense* versus *blind* contrast. However, the direct comparison of *localise* versus *sense* revealed no significant activations. Although ACC activation has been found to boost attention towards task-relevant stimuli (Orr & Weissman, 2009; Kim et al., 2016), this does not explain why ACC activation was not present in the *localise* trials were participants had full awareness of the change. We therefore suggest that the ACC activation for *sense* trials may reflect error monitoring, as participants incorrectly localised the change (Iannaccone et al., 2015; Debener, 2005).

It is possible that during *sense* trials, the ACC activation reflected a mis-match between the intended response and the actual response (Dehaene, 2018).

Although participants had represented the stimuli in visual working memory (indexed by the increased visual and parietal activation that was similar to *localise* trials), and planned the correct response, their actual response did not match their intended one leading to ACC activation. In *blind* trials, participants had significantly reduced visual and parietal activation, and may not have known which response was correct. This could explain the increased ACC activation when comparing *sense* versus *blind trials*, however, we acknowledge that is very difficult to assess this hypothesis without knowing the participant's intended response.

7.4.1 Certainty and difficulty

Using participant certainty at each trial as a parametric regressor, we found significant activations in the right visual cortex (BA18, V2), and bilateral supramarginal gyrus (BA40). These regions largely overlap with the cortical activations also found to increase with awareness of the change (*localise* and *sense* trials), possibly due to the relationship between awareness and certainty. Specifically, when participants were aware of the change and could localise it correctly, they were likely to report higher certainty in their responses. The parametric regressor of task difficulty (the number of squares presented per trial) revealed significant activation in the visual cortex (BA18, V2). This finding probably reflects the greater visual stimulation associated with a more complex visual array.

7.5 ERP-informed fMRI

Our pre-registered analysis method of ERP-informed fMRI revealed no significant results. We therefore failed to identify voxels with activation that significantly co-varied with fluctuations in the EEG. It is acknowledged that EEG-BOLD couplings are weak, as they measure the effects remaining after the mean evoked BOLD responses are explained (Liu et al., 2014). However, pre-

vious combined EEG-fMRI experiments have managed to identify correlates of EEG using ERP-informed fMRI (Debener, 2005; Eimer & Mazza, 2005), even if at liberal correction thresholds. It is our intention to explore further analysis techniques on this data set, and consider the implications of all processing methods in single-trial ERP extraction for combined EEG-fMRI.

One pre-processing step that was chosen during pre-registration was baseline correction of the single-trial ERP values. When running traditional ERP analyses, each epoch is baseline corrected before a grand average is formed, usually by subtracting the mean value of the 200 ms preceding stimulus onset. The rationale for this is to bring all conditions to the same baseline value, therefore facilitating meaningful comparisons between them. It also removes the influence of signal drift over time, as each epoch is corrected using a section of the time-series immediately preceding it.

One problem with using the preceding 200 ms is that any noise contained in this section of the time series will be added to the epoch of interest. This is likely to occur if participants move or blink deliberately between trials. An alternative method is to subtract the mean across the epoch from the time series, which therefore avoids adding outside noise to the trial. This method was chosen for baseline correcting the single-trial ERP values used in ERP-informed fMRI analysis.

However, when running the ERP-informed analysis, it became clear that this baseline correction influenced the results. For many participants, using single-trial values that had not been baseline corrected revealed significant activations in visual/parietal voxels that we had expected to see. In comparison, baseline corrected values removed most of these effects. This raises several questions - is the baseline correction removing the meaningful content of the single-trial values? Or, are the non-baseline corrected values spuriously correlating due to an artifact occurring over time that matches the MRI signal in some voxels? On reflection, most EEG-fMRI papers do not explicitly mention baseline correcting their single-trial values, despite the fact that is

an essential step in traditional ERP analyses. Our decision to baseline correct may have been misinformed, but influence of a simple baseline correction on ERP-informed fMRI analysis is an interesting finding, which could benefit from further study.

7.6 Contributions to the field

As far as we are aware, we ran the first fMRI (and EEG-fMRI) study investigating the *sense* condition in a change blindness paradigm. We are therefore able to provide initial evidence for the distinction between *sense* and *blind* conditions in BOLD fMRI data. The change blindness paradigm that we used was also novel, as the difficulty was modulated in real time. Using this paradigm we were able to overcome issues faced by other researchers, such as low trial numbers in the *sense* condition.

As we collected EEG data from the same paradigm across two experimental sessions (EEG only and EEG-fMRI), we were able to compare the EEG data recorded inside and outside of the scanner. Only a handful of other studies have focused on this comparison, and none were using the change blindness paradigm.

We further provide two novel literature reviews. One focused on the definition and investigation of the *sense* condition in behavioural and neuroimaging experiments (Chapter 2), which will provide a strong foundation for further work in the field. The other focused on simultaneous EEG-fMRI and considered when combined recording is really necessary (Chapter 7). This review will benefit future researchers planning to use this multi-modal approach, ensuring that their paradigm and research aims align with the method.

7.7 Limitations

fMRI paradigm design

One limitation of the EEG-fMRI experiment was the fast timing of the paradigm, which may have prevented the separation of BOLD signal relating to different task processes (such as encoding, maintenance, and response). For the fMRI experiment we increased the ITI from 2-3 seconds to to 5-7 seconds, in an attempt to maximise the time between trials. However, there are other modifications to the paradigm that would have improved our ability to model the haemodynamic responses across each trial.

Adding jitter to the 100 ms fixation following the change display would have improved the experimental design. The fixed duration prevented us from separating change related activity from response activity, as the timing between the change display and the first response was always the same. After the first response onset, the timing of the following responses did vary with participant reaction times. However, the disadvantage of this was that faster participants were able to reduce the overall time between the end of one trial and start of the next. A pertinent modification would be to fix each trial to a certain length and present a fixation during the remaining time after responses have been completed.

MRI response device

In question 2 of the experimental paradigm, participants were asked to localise the change based on a 2x2 grid displayed on the screen. The ideal response device for this paradigm would have been a pad with four buttons arranged in a similar manner, making it simple for the participants to translate the location on the screen to the location on the response pad. This is particularly relevant for the fMRI experiment, as participants could only respond using one hand. However, we only had access to a typical MRI button box, which has a

1x4 layout.

We acknowledge that the responses to button 4, made with the little finger, are likely to have been slower than button 1, made with the index finger. However, all of our reaction time results were taken from question 1, which had only two responses (button 1: yes, button 2: no), and therefore were not influenced by this. We also found that participants who incorrectly localised the change were just as likely to get the upper changes incorrect (button 1: top left, button 2: top right) as lower changes (button 3: lower left, button 4: lower right). Therefore, having to respond with the third and fourth finger did not increase the likelihood of incorrect detection.

Sensing definition

In our paradigm, if participants responded that they did not see a change during question 1 ('Did you see a change?'), then question 2 was skipped ('Where did the change occur?'). This was because we were not primarily interested in these trials, and removing this question allowed for additional trials to be added to the experiment. However, on reflection, it would have been interesting to retain these trials, as they would capture 'implicit' detection based on the definition from Fernandez-Duque et al. (2000) and others. Given that we do no know how the different definitions relate to each other, it would have been interesting to compare these to our *sense* condition. (Recap; 'implicit' awareness is here defined as incorrect detection but correct identification; *sensing* is here defined as correct detection but incorrect identification.)

Between-groups design

For our comparison of EEG recorded inside and outside of the scanner, we used a between-group design with two different participant groups. This results in reduced statistical power for the EEG comparison, and individual differences will cause an unknown percentage of the variance between the EEG data sets. The most optimal design would be to use the same set of participants for both

the EEG and EEG-fMRI experiments. However, these experiments would then need to be counterbalanced across participants to remove practice effects for the change blindness task. This would have introduced time constraints, given that the fMRI scanner at Reading was upgraded around the time of the EEG only experiment. We therefore opted for two separate experiments, allowing time for the EEG experiment to be completed and written up while the scanner was not in use.

7.8 Further analysis

MVPA

As discussed throughout this thesis, a prevailing limitation of the *sense* condition is the lack of consistent definition. Although an implicit awareness by definition, we rely on explicit responses from participants to categorise trials and infer whether they are *sensing* or *seeing* a change.

The traditional GLM approach to BOLD fMRI analysis informs us how well the activity at each voxel matches a model of activity that we assume based on our paradigm and our trial labels. Multivariate pattern analysis (MVPA) is a different approach that attempts to distinguish between patterns of activity across subsets of voxels. For example, the pattern of activity across voxels in the early visual cortex differs significantly when observers are shown an image of a face compared to that of a house. Using these patterns of activity we are able to identify which image observers are looking at with a high level of accuracy (Haxby, 2001; Haxby et al., 2014; Norman et al., 2006).

Given that our BOLD results suggest a number of brain regions where the activity differs for *sense* and *localise* trials, it would be interesting to compare the patterns of activity within these regions using MVPA. If we are able to successfully predict the trial type based on these patterns, then we have additional evidence that these brain areas are responding differently depending on the level of awareness. Further, it would be interesting to compare the

classification accuracy of trials depending on different definitions of awareness and/or *sensing*. We assume that the brain categorises trials in the same way that we do, using our definitions of *sense*, *localise* and *blind*, but this may not be the case. For example, it may be that classification is most successful when patterns are divided into certain versus uncertain trials.

ERP-informed fMRI

As discussed previously, we did not find any voxels with activity that co-varied with fluctuations in the LP ERP. In this pre-registered analysis we used the most simple method for extracting single-trial values. However, other methods are also used in the literature, such as ICA, and we plan to compare the single-trial values extracted using this method. Typically, temporal ICA is run on each data set, and the ICA component with the highest correlation with the ERP of interest is identified. Single-trial values are then taken from the time course of this component, rather than the raw EEG. This method relies on the success of ICA to isolate the ERP of interest, and the ability of the researcher to identify the correct component.

Multivariate EEG analysis

All ERP analysis was defined before data collection, and based on a selection of previous EEG papers investigating *sensing*. However, it would be useful to run a post-hoc multivariate EEG analysis including all electrodes and time points. Although it is more difficult to control for false-positives using this approach, it would identify any effects that lie outside of the ERPs already analysed.

Error-related negativity

One possible explanation of *sense* trials is that they are simply *localise* trials with a response error. One way to establish this would be to compare the error-related negativity (ERN) ERP amplitudes for each awareness condition.

This ERP peaks around 100ms after a response, occurring maximally at front-central electrode sites, with an increased amplitude for incorrect responses (Gehring et al., 2018). If ERN amplitudes are greatest for the *sense* condition, then they would provide evidence in favour of this hypothesis.

The role of distractors

The number, location, and colour of all the distractor squares presented at each trial was stored during the EEG-fMRI experiment. This was based on the hypothesis that particular arrangements of distractors may have facilitated or hindered change detection. However, this data has not yet been analysed, as it did not relate to any of our key research questions.

7.9 Future directions

Despite the allure of complicated EEG-fMRI experiments and high-level analysis techniques, it became increasingly obvious throughout the course of this thesis that what the field of *sensing* lacks is a coherent definition and concrete experimental paradigm. While we opted for the only *sensing* definition that had been used with EEG previously, there were a number of paradigms that we could have chosen. It is rare that researchers attempt to compare or combine definitions, and only mention in passing that they may reflect different types of processing. A behavioural paradigm that could capture several definitions of *sensing* would be beneficial to the field, particularly if adopting a psychometric approach of many trials across fewer participants (although the more the better). Including enough trials for each awareness level across definitions of *sensing*, and possibly across participant certainty levels and abilities, would greatly improve our knowledge of this condition in change detection.

7.10 Overall conclusions

Does *sensing* really exist?

Regardless of any EEG or fMRI results, it remains true that participants commonly report a 'sense' that something has changed during change detection tasks, but without full awareness of what or where this occurred. This phenomenological experience of observers is reported widely in the literature, and I am certain that this sensation does exist (having experienced it many times myself). What remains unclear, however, is how this relates to behavioural and neurological data, and whether we have created experimental designs that are able to distinguish these occurrences from others.

To summarise the discussion from previous chapters, there are a few possible explanations for the phenomenological experience of 'sensing' a change:

- *a)* when we *sense* a change, we have not explicitly detected anything. Given the nature of the task, we know that we are looking for changes, and therefore we sometimes overestimate our detection of differences. We may also be able to guess the correct responses using strategies facilitated by the experimental design, but this is not based on our own awareness (Mitroff et al., 2002). Therefore, only our phenomenological experience is different to that when we are *blind* to the change.

- *b)* when we *sense* a change, we do have some representation of the change (at a neural level), but we do not have full access to that information. Our report of any specific details is therefore lacking. The brain activity for *sensing* should therefore be greater than *blind* trials, but may also differ from *localise* trials. This links to hierarchical theories of visual processing and awareness (Dehaene et al., 2006).

- *c)* when we *sense* a change, we have full knowledge of that change, but we simply make a mistake in our report of the details. For example, we

press the wrong response button. The brain activity should therefore be similar to *localise* trials.

Based on the range of results reported in this thesis, our conclusion is that the *sense* condition (as we have defined it), may be best explained by both item *b* and *c* above. In behavioural, EEG, and EEG-fMRI data, we found a range of evidence to suggest that the *sense* condition is distinguishable from the *blind* condition. For example, reaction times for *sense* certain trials are faster than *blind* trials; late positivity, N2pc, and N1 amplitudes are larger for *sense* trials compared to *blind*; and a range of visual and parietal areas showed increased BOLD activation for the *sense* versus *blind* condition. All of this evidence suggests that differential processing occurs during *sense* trials that cannot be found when participants completely miss the change. This is strong evidence against item *a* above.

The support for item *b* is not as clear. As mentioned above, we have evidence to suggest a distinction between *blind* and *sense* trials. However, we have found very little evidence to dissociate *localise* and *sense* conditions. In our combined EEG results, no significant differences were found between these two conditions. Similarly, no significant voxels were found in the contrast between *sense* and *localise* in the fMRI analysis, and they revealed largely overlapping networks when contrasted with *blind* trials. Although we did find midbrain activation that was specific to the *localise* > *blind* contrast (putamen and insula), and ACC activation specific to the *sense* > *blind* contrast, we are cautious about drawing conclusions based on these separate analyses. Additionally, no activations were found to be significantly higher in the *localise* > *blind* contrast than *sense* > *blind*, and visa versa.

The lack of a clear distinction between *localise* and *sense* conditions therefore leads us towards statement *c* above. It is possible, based on our definition of *sensing*, that participants knew the correct location but pressed the wrong response button. In this case, trials were categorised as a *sense* trial when

they should not have been (also suggested in Thornton et al., 2001).

Evidence to the contrary, as also stated by Ball and Busch (2015), is that the incorrect location responses were random. If participants responded incorrectly, you may expect them to press the neighbouring location response by mistake. However, there was no significant relationship between the correct location and the reported location. Roughly half of our participants were more accurate in their left/right report (for example, if the correct answer is bottom left, they might respond top left instead), whereas the other half were more accurate in their top/bottom report (if the correct answer is top left, they might respond top right).

As previously suggested, one way to help rule out this confound is to provide an additional response screen for participants to report if they made a mistake in their answer. These trials could then be excluded from the *sense* condition. However, this is only useful for trials where the participant is aware that their response was incorrect, which may not always be the case. It may also be beneficial to remove the time constraint for the localisation question (Q2), to reduce the chance of response errors.

What does *sensing* tell us about change blindness?

As discussed in Chapter 1, there are a number of explanations for the phenomenon of change blindness. One theory is that blindness occurs due to a failure to encode either the pre- or post-change display. If no information is stored, then differences cannot be detected. However, previous researchers have argued that the presence of the *sense* condition provides evidence against this hypothesis, as observers can identify object from both displays above chance level (Simons et al., 2005a; Hollingworth et al., 2001). Even when observers are not explicitly aware of a change, their ability to correctly identify items from both displays suggests that they have stored some information about them.

The fMRI data reported in this thesis provides evidence against the hypoth-

esis that change blindness occurs when representations are not encoded. Even when participants could only *sense* that a change had occurred, significantly increased activation was found in early visual and parietal areas. These areas are commonly implicated as storage sites for the representation of items in visual working memory, and therefore their activation during *sense* trials may suggest that some relevant information was stored. An alternative explanation is therefore required.

One such explanation is that change blindness occurs due to a failure to compare the two visual displays. Despite storing information of the pre- and post-change scene, the observer fails to compare the contents of these representations, and therefore cannot provide details about the change. It is possible that the increased BOLD activation we found for *localised* trials in a wider network of visual, parietal, and mid-brain structures, is what facilitated successful report of the change location.

However, it is also possible there are multiple explanations for change blindness and *sensing*. In an experiment intending to dissociate between different explanations, Varakin et al. (2007) found that multiple explanations could be applied to distinct subsets of participants. Observers who missed changes and had low confidence in their ability also had poor memory for the pre- and post-change items, suggesting that they failed to represent the information. By contrast, observers who missed changes but had high confidence demonstrated good memory for these objects, indicating a comparison failure as the cause of their change blindness. Therefore, even within one experiment, several explanations were plausible.

It is also possible that, within one experiment, multiple types of *sensing* occur, such that different aspects of the change are available for explicit report during different trials. Based on this, an important question to ask is what determines the availability of this information for conscious report, and how can we improve our awareness of changes in the world around us? Further, is explicit and detailed knowledge of these changes necessary for our successful

interaction with the world, or is *sensing* enough to enable our survival?

Summary

Overall, using a combination of behavioural, EEG, and fMRI measures, we found evidence supporting the distinction between *sensing* a change and being *blind* to it. In EEG, the late positivity potential, N2pc, and N1 amplitudes were larger for *sense* trials compared to *blind*. Additionally, a range of visual (BA18), parietal (BA40), and midbrain (anterior cingulate) areas showed increased fMRI BOLD activation when a change was *sensed*. These visual and parietal areas are commonly implicated as the storage sites of visual working memory, and we therefore argue that *sensing* may not be explained by a lack of representation of the visual display.

However, it is less clear whether *sensing* a change is quantifiably different to complete awareness, given the lack of evidence supporting the distinction between these two conditions. We argue that *sensing* is a phenomenologically different experience to complete awareness, but suggest that further research is needed to determine how our choice of definition for *sensing* may influence our ability to measure it accurately.

Given the relevance of change detection in our everyday lives, the notion that our brain is processing more information than we are explicitly aware of is a pleasing one. If we can identify the mechanisms that 'push' our knowledge of changes from merely *sensing* them to being fully aware of their characteristics, then we can apply this knowledge to situations where detection of changes in the outside world is paramount.

Appendix A

Additional analysis

A.1 EEG pilot

Before beginning the EEG and EEG-fMRI experiments that form the main body of this PhD, a small EEG pilot with three participants was conducted. The main aim of this was to ensure that the extended change blindness paradigm worked in the way that it was intended, that participants were able to understand the task, and could respond within the restricted response time. We also wanted to ensure that the difficulty modulation worked equally well across the three participants at stabilising their accuracy over time. Several changes were made to the paradigm as a result of this pilot, such as the removal of one confidence question, and the adjustment of square positioning on the screen. All of the changes made as a result of this pilot will be justified and discussed.

Materials and Methods

Participants

Four subjects (male, mean age 28) with no history of psychiatric or neurological disorders participated in this EEG recording. All had normal-to-corrected vision and were not colour blind. Unfortunately, one participant was removed due to excessive noise and a subsequent lack of usable trials.

Stimuli and Presentation

The paradigm used throughout this thesis is a modified change blindness paradigm. In a standard paradigm, participants are shown two displays with an interrupting fixation and are asked to report if they detected a change between them. This is equivalent to question 1 in our paradigm, which is displayed in figure A.1. The modification in our version comes from the addition of three questions, asking the participant to localise the change, and indicate the level of confidence they have in their responses.

Although many types of visual display have been used to investigate change blindness, such as household objects, naturalistic scenes, and faces, we opted for the simplest display in order to avoid increasing the number of potential confounds. Our display therefore consisted of a number of small squares, and the change was manipulated by changing the colour of one square from the first to the second display.

Difficulty was modulated in real time by adding and removing squares from the display, with the assumption that more squares on the screen, or more distractors, makes the task more difficult. Performance over the previous two trials was used to update the current trial in a two-up-two-down method; two correct answers increased the number of squares, two incorrect reduced the number, and one of each resulted in no change. Answers were taken from the localisation question, as it was predicted that this would have lower accuracy than the identification question (Yes/No), and therefore using this would prevent the task from becoming too difficult.

Overall, participants completed a total of 5 blocks each containing 50 trials, meaning a total of 250 trials. One third of these were no change trials, leaving 165 change trials. No change trials were of less interest, and therefore were of a smaller percentage. Participants were asked in a debrief questionnaire, after completing the experiment, to indicate what percentage of the trials they thought were change trials. If a participant was able to answer this question

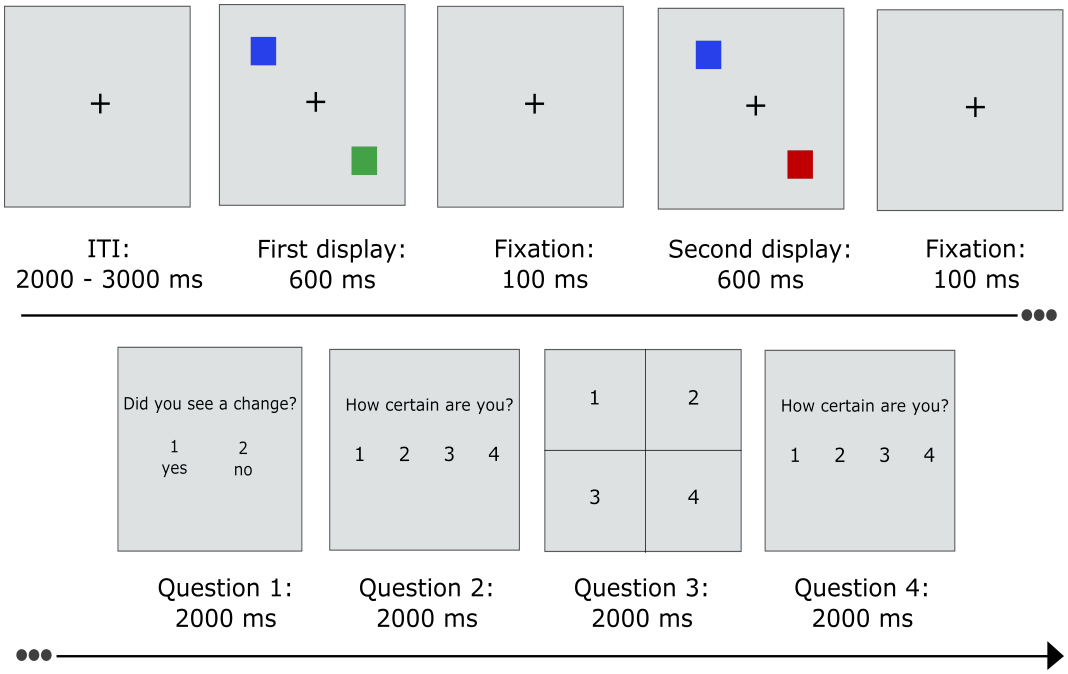

Figure A.1: Illustration of the original experimental paradigm, used in the EEG pilot. The number of squares presented varied from 2 to a maximum of 36. Question 1 asked 'Did you see a change?' to which participants could respond 'Yes' or 'No'. Question 2 asked how certain participants were of their responses, ranging from '1: Very Certain' to '4: Very Uncertain'. Question 3 asked participants to localise the change, based on a grid from top left to bottom right. Question 4 was another certainty question. The correct answers in this instance would be '1: yes' there was a change, in quadrant '4', bottom right. Participants were asked all questions on all trials.

correctly, then it would indicate that their responses may be biased to the correct ratio of change versus no change, and they may have an advantage over other participants. However, no one in any of the following studies reported the correct change percentage.

Data Acquisition

EEG data were recorded with BrainVision EasyCap, with 64 electrodes including an IO channel. The reference electrode was placed at FCz and the ground at AFz. Impedance was kept below 10k. EEG recordings were performed with Brain Vision Recorder Software (Brain Products) at a sampling rate of 5000 Hz. This was later down-sampled to 500 Hz for analysis.

Pre-processing

Raw EEG data was pre-processed using Brain Vision Analyzer version 2.1 (Brain Products). The data was first filtered with a high-pass of 0.1 Hz, low-pass of 50 Hz and a notch of 50 Hz. The data was then corrected for eye-blinks using ICA to identify artifactual components, which were manually inspected in semi-automatic mode. Two components were removed, corresponding to eye-blinks and lateralised eye movements. Raw data inspection identified artifacts based on a maximum voltage step of 50 V/ms and lowest allowed activity in intervals 0.5 V. The majority of these were identified outside of trial timings. Segments were then taken from -200 to 6000 ms to include the whole trial, and baseline corrected from -200ms to 0. These were time locked to the onset of the first display. Trials containing marked artifacts were excluded from further analysis, as well as those where a response was not made within the response time.

The trials in which a change occurred could be separated into: 'change blind' (no change reported), 'identified' (change reported and correctly localised), and 'sensing' (change reported but incorrectly localised). Trials in which no change occurred can be divided into 'correct rejection' (no change reported) and 'false alarm' (change incorrectly reported). In the pilot, another condition was possible, as participants were always asked to localise the change, even if they reported seeing no change. This condition contains trials where participants report no change, but then accurately located the change (possibly by chance).

Behavioural Results

Trials per condition

Table A.1 shows the number of trials in each condition, for all three participants. Table A.2 shows the same data in percentage of total trials. Identified and blind are represented as a percentage of all change trials. Localised and

sensing are represented as a percentage of all identified trials. Correct rejection and false alarm are represented as the percentage of all no change trials.

One aim of the difficulty modulation was to ensure that participants did not perform at floor or ceiling, and had trials in all awareness conditions. The overall performance of the participants was around 70% accuracy for the yes/no detection question, with a larger percentage of trials falling into the identified condition. This ratio was also observed between the localise and sensing conditions, with roughly twice as many trials in the localised condition. In an ideal case, these trials would be split 50/50, meaning equal trials in each condition for subsequent analysis. However, it is hard to manipulate this. For example, consider a situation where the yes/no detection accuracy was higher, for which the task would have to be easier. In this case, there would be more trials to split between localised and sensing conditions, but it is likely that the ratio would also be more biased towards the localising condition due to the task being easier. Although the difficulty measure successfully managed to stabilise the accuracy across participants, it is important to remember for analysis that the number of trials contributing to each ERP is very likely to vary across conditions.

There has been some suggestion that a response pattern with a higher number of false alarms would be associated with a higher number of sensing trials, if sensing is due to a less conservative response bias (Howe & Webb, 2014). The results here are contradictory, as participant 3 had a higher number of sensing trials than the other participants, but a lower number of false alarms.

Reaction Times and Accuracy

Table A.5 contains both reaction time and accuracy data for participants 1, 2, and 3. An average was taken within each block, as well as an overall average. Difficulty represents the number of squares presented to the participant, with a higher number indicating a greater difficulty. Participant 1 was able to perform at 70% accuracy in the identification of a change (Yes/No) with an

Number of Trials per Condition						
	Change Trials				No Change Trials	
Participant	Identified	Blind	Localised	Sensing	Corr. Reject.	False Alarm
1	114	51	75	36	44	41
2	122	43	84	39	46	39
3	112	53	58	54	68	17
Average	116	49	72	43	53	32

Table A.1: The number of trials in each awareness condition. Identified: trials where the participant correctly identified a change in the Yes/No question. Blind: trials where they failed to identify a change in the Yes/No question. Localised: trials where they correctly localised the change. Sensing: trials where they incorrectly localised the change. Correct Rejection (Corr. Rect.): correctly identified no change. False alarm: incorrectly suggested a change in a no change trial.

Percentage of Trials per Condition (%)						
	Change Trials				No Change Trials	
Participant	Identified	Blind	Localised	Sensing	Corr. Reject.	False Alarm
1	73	27	70	30	54	46
2	74	26	68	32	55	45
3	68	32	52	48	80	20
Average	72	28	63	37	63	37

Table A.2: Percentage of the total number of change or no change trials, respectively. Identified and blind sum to 100% of change trials. Localised and sensing sum to 100% of identified trials. Correct rejection and false alarm sum to 100% of no change trials.

average of 14-18 squares, whereas participant 2 performed at this level with 9-10 squares. This demonstrates the ability of the difficulty modulation to stabilise performance across participants, by adjusting the difficulty of the task to suit their individual ability.

From visual inspection, it appears that reaction times generally decreased over the experiment, as you might expect when the participants became more familiar with the task. For participants 2 and 3, difficulty increased over time, possibly reflecting learning or increased familiarity. For participant 1, however, there is greater fluctuation, with the highest difficulty occurring in block 3. The subsequent decrease in difficulty could reflect a reduction in motivation during later blocks. For more detailed information, the frequency of each difficulty level performed by each participant is displayed in table A.6.

D Prime

Table A.3 shows the d prime scores for each participant. D prime is a measure of response bias, with the equation $d = z(\text{hit rate}) - z(\text{false alarm rate})$, and is defined as the difference between the means of signal and noise distributions, normalised by the variance. Hit rate is defined as the proportion of hits in the change trials $P(yes|change)$, and false alarm rate as the proportion of false alarms in no change trials $P(falsealarm|nochange)$. It therefore takes into account performance accuracy as well as tendency to over estimate the number of changes during the experiment. Someone always responding 'yes' they saw a change will have a high accuracy, but also a high false alarm rate, and therefore a low d prime score.

D Prime Scores		
Participant	d'	c
1	0.72	-0.26
2	0.79	-0.26
3	1.31	0.19

Table A.3: D prime (d') and criterion (c) calculated for each subject, using their responses to the yes/no question. Higher d' means greater discrimination between change and no change conditions. $c > 0$ indicates a bias towards 'no' responses. $c < 0$ indicates bias towards 'yes'. $c = 0$ indicates no bias.

Response bias, or criterion, can also be calculated, where $c = -0.5 * (z(\text{hit rate}) + z(\text{false alarm rate}))$. $c = 0$ indicates no response bias to either 'yes' or 'no' responses. $c > 0$ indicates a bias towards 'no' responses, with fewer hits and fewer false alarms. $c < 0$ indicates bias towards 'yes', with more hits but also more false alarms.

The d prime scores provide more evidence for the varying response patterns demonstrated by the participants. While participants 1 and 2 had a very similar number of false alarm trials, and therefore d prime scores, participant 3 had much fewer false alarms. This is reflected in their higher d prime score of 1.3, as well as their positive criterion. This suggests that participant 3 had a more conservative response pattern, with a slight tendency towards misses

rather than false alarms.

Certainty

Table A.4 show the responses to both certainty questions (Q2 and Q4 in the paradigm shown in figure A.1). Within each trial, the responses were highly correlated: participant 1, $r = .83, p < .000$, participant 2, $r = .75, p =< .000$, and participant 3, $r = .73, p < .000$. From a visual comparison, it appears that the three participants had different response patterns. For example, participants 1 and 2 were less likely to respond very uncertain than participant 3. Overall, there seems to be a bias towards the responses very certain and uncertain.

Yes/No Question Certainty								
	Correct				**Incorrect**			
Participant	VC	C	U	VU	VC	C	U	VU
1	56	11	30	1	4	21	10	1
2	82	14	27	0	19	2	22	0
3	67	12	19	14	20	16	7	10
Location Question Certainty								
	Correct				**Incorrect**			
Participant	VC	C	U	VU	VC	C	U	VU
1	45	8	19	3	4	9	17	0
2	56	6	35	2	11	10	17	1
3	49	7	12	16	17	1	19	17

Table A.4: The number of times participants responded with different levels of certainty to the yes/no question (top) and the location question (bottom). This includes both change and no change trials, comparing correct versus incorrect answers. VC: very certain, C: certain, U: uncertain, VU: very uncertain.

EEG Analysis

Ghost Markers

When using a parallel port to send EEG event markers, from the paradigm presentation computer to the EEG recording computer, it is important to reset the port to zero after each initiation. This was not correctly implemented in the pilot study, and therefore certain markers in certain combinations sum-

mated to create 'ghost' markers. These did not correspond to intended marker codes, and therefore created problems for subsequent analysis. Although corrections were attempted for this after recording, not all trials could be correctly identified, and those that were may be subject to human error. There were therefore fewer usable trials for each person and each condition than intended, and the EEG results are subsequently more noisy. This mistake was corrected in the code for all further experiments.

Despite this error, and the small sample size of 3 participants, event related potentials (ERPs) were calculated from the pilot data in order to facilitate a visual inspection on the quality of the usable EEG.

P1 and N1

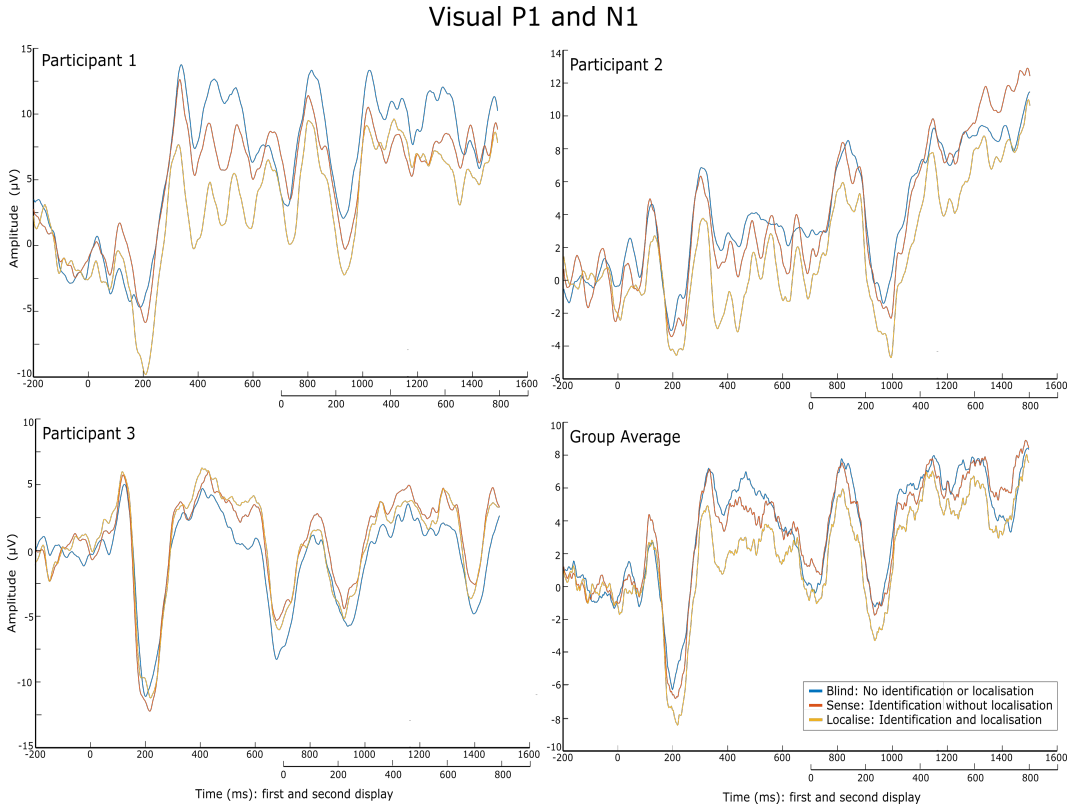

Figure A.2: P1 and N1 amplitudes taken from electrodes PO7 and PO8, displayed individually for the three pilot participants as well as a group average.

Checking for visual ERPs in response to a visual task is a good way to assess the quality of your data. If we had failed to find visual ERPs, it may

have indicated a design flaw or confound within the experimental paradigm, or within the analysis pipeline. As shown in figure A.2, visual ERPs were present in the data for all three participants.

Participant 1 has the noisiest data, in terms of the visibility of visual ERPs, and the P1 for the first display is small. It is, however, followed by the negative N1 and then a subsequent P2, as expected. It is also clear where the onset of the second display falls, as you see another clear P1/N1/P2 time course (within 100 to 400 ms after second onset). Participant 2 also has clear ERPs, but the data appears to contain some drift towards the end of the trial, indicating a deterioration in data quality over time. If this was not a pilot data set, this slow component drift would need to be examined in further detail, and perhaps more pre-processing steps applied to correct for it. Participant 3 has the cleanest ERPs, demonstrating the typical time course of visually evoked components over the visual cortex. Despite the noise in Participant 1's data, the timing of the visual ERPs was fairly consistent across the three participants, and therefore these peaks are still clearly visible in the group average.

N2pc

The N2pc was defined as the mean difference between contralateral and ipsilateral recording electrodes, in relation to the position of the change stimuli, within 200-400 ms after the second display. Occipital electrodes PO7 and PO8 were chosen based on previous literature. This component should have a greater amplitude when changes are seen, and correctly identified in the Yes/No question, compared to changes that are not seen.

Figure A.3 shows the results from the three participants. The ERPs for participants 2 and 3 are very noisy, and difficult to draw conclusions from. This is likely to be due to the low trial numbers as a results of incorrect trial markers. Although it appears that participant 1 has a cleaner and better signal, the differences in N2pc are apparent over almost the whole trial for the localised condition, which you would not necessarily expect. As the amplitude

of the N2pc is much greater than those for participants 2 and 3, the group average mostly reflects the data from this participant. This highlights the potential difficulty in using group data to represent single subject activity, when the overall amplitude range varies between individuals.

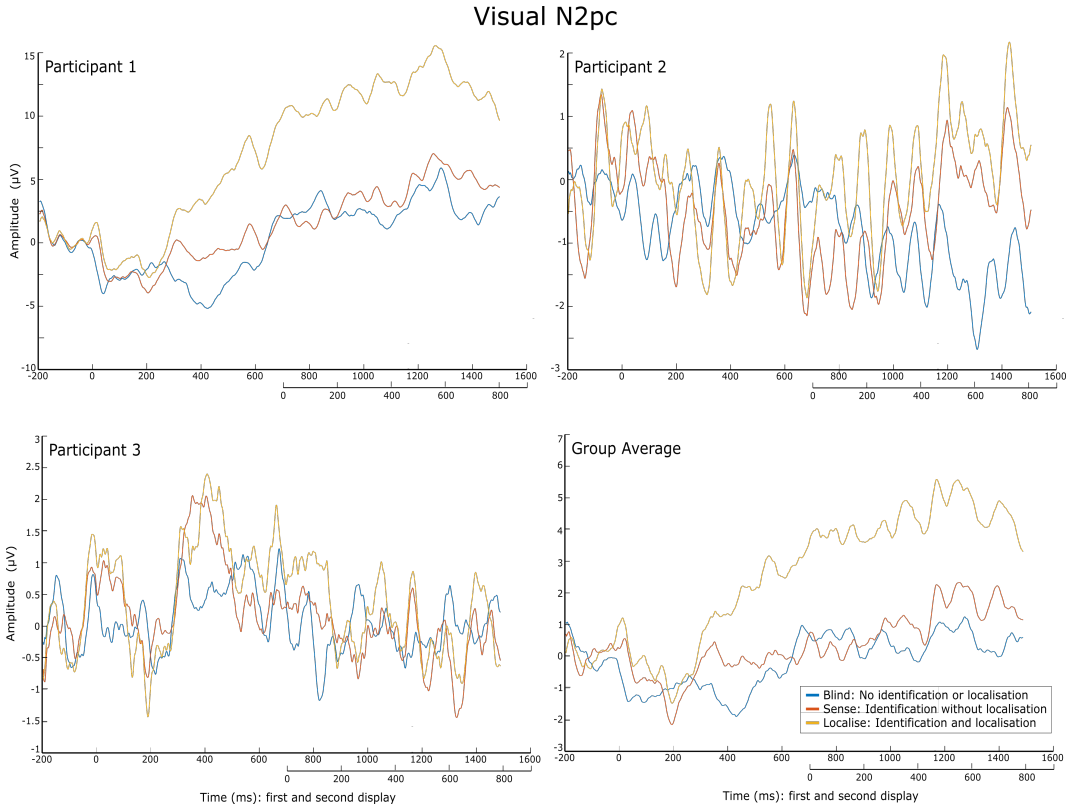

Figure A.3: N2pc amplitudes taken from electrodes PO7 and PO8, displayed individually for the three pilot participants as well as a group average.

Conclusions and Changes

Based on the results of the EEG pilot described above, the following changes were made to the paradigm:

- The answers to the two certainty questions were very correlated within each trial. This matched participant reports that their responses within one trial were very similar. We therefore removed the first certainty question, leaving only one at the end of each trial. A secondary positive outcome of this change was the reduction in total running time of the experiment.

- Originally, the certainty question response ranged from '1: very certain' to '4: very uncertain'. However, participants all reported that this order was not intuitive, and so the order was changed to '1: very uncertain' to '4: very certain'. This change meant that a higher number corresponded to a higher degree of certainty.

- The necessity of four possible responses for the certainty question, rather than two or three, was also mentioned. Due to the variable response pattern across the small subset of participants, it was decided that the four responses should remain, but that they would probably be collapsed into certain and uncertain for future analysis.

- Participants reported that it was sometimes difficult to distinguish between the top and bottom locations of the squares. The distance between the possible square locations in the top and bottom parts of the screen was therefore increased in the experiment code.

- In the original paradigm used in the pilot, participants were always asked all four of the questions. However, the participants reported that the localisation question was unnecessary and confusing when they responded that they did not see a change. There were also very few trials in which the location was subsequently identified correctly (ranging from 2 to 7), from which no meaningful analysis could be conducted. The pradigm was therefore changed, so that if participants responded 'no change' in question 1, they were not asked to localise the change, and were only asked their confidence before moving to the next trial.

- The difficulty modulation appeared to work as expected, and therefore no changes were made.

- All participants responded well within the limited response time for each question, so no changes were made to the allocated time.

- Ghost markers were appearing in the EEG data due to a failure to reset

the parallel port codes to zero after every initiation. This was therefore implemented for the future.

Behavioural Results: Participant 1							
	Reaction Times (s)				**Accuracy (%)**		
Block	Q1	Q2	Q3	Q4	Yes/No	Location	Difficulty
1	0.67	0.73	0.62	0.44	75	55	12.80
2	0.73	0.55	0.64	0.33	65	27	20.95
3	0.57	0.53	0.49	0.32	58	33	25.10
4	0.53	0.45	0.50	0.36	70	43	18.20
5	0.56	0.34	0.53	0.26	70	44	13.85
Average	0.61	0.52	0.56	0.34	68	40	18.18
SD	0.34	0.37	0.31	0.30	6.43	10.81	2.87
Behavioural Results: Participant 2							
	Reaction Times (s)				**Accuracy (%)**		
Block	Q1	Q2	Q3	Q4	Yes/No	Location	Difficulty
1	0.86	0.70	1.00	0.79	72	28	9.28
2	0.71	0.53	0.67	0.42	70	37	9.20
3	0.60	0.41	0.68	0.37	70	32	9.64
4	0.62	0.42	0.44	0.30	80	38	15.56
5	0.67	0.47	0.73	0.29	72	26	15.20
Average	0.69	0.50	0.70	0.43	72	32	11.78
SD	0.27	0.31	0.40	0.37	4.15	5.31	2.59
Behavioural Results: Participant 3							
	Reaction Times (s)				**Accuracy (%)**		
Block	Q1	Q2	Q3	Q4	Yes/No	Location	Difficulty
1	0.76	0.89	0.66	0.53	70	43	8.08
2	0.76	0.94	0.92	0.75	72	51	8.96
3	0.76	0.79	0.71	0.59	72	34	14.28
4	0.82	0.82	0.74	0.59	66	43	14.60
5	0.67	0.58	0.65	0.51	68	29	16.96
Average	0.75	0.80	0.74	0.59	70	40	12.58
SD	0.05	0.14	0.11	0.09	2.61	8.60	3.86

Table A.5: Average reaction times in seconds (s) for each block are presented separately for each question, in the order which they were asked. Accuracy is recorded separately for identification of a change (Yes/No) and the localisation of the change, in percentage (%). Difficulty represents the average number of squares presented to the participant during the block.

	Participant		
No. of Squares	*1*	*2*	*3*
2	1	8	7
4	1	21	17
6	1	21	27
8	9	25	25
10	9	30	30
12	21	47	31
14	25	35	27
16	17	27	23
18	16	14	22
20	32	8	19
22	24	10	13
24	18	2	7
26	16	2	2
28	8	0	0
30	2	0	0
32	0	0	0
34	0	0	0
36	0	0	0

Table A.6: Difficulty of the task, manipulated by the number of squares presented during each trial. Here the frequencies of each difficulty level for each participant are presented. The maximum number of squares was 36, with a 6x6 grid. An equal number of squares was always presented on the left and right side of the display, and therefore the difficulty levels are multiples of 2.

A.2 EEG experiment

Additional results from the EEG only experiment.

Reaction times: average reaction times were compared using a repeated measures ANOVA calculated for question (1/2/3) and accuracy (correct/incorrect). The main effect of question was significant, $F(2, 38) = 9.543, p < .001, eta = .334$. (Q1: $M = .680$, Q2: $M = .842$, Q3: $M = .784$). The detection question (Q1) was significantly faster than the localisation question (Q3), $p = .002$, and the certainty question (Q3), $p = .004$. However, the localisation and certainty questions were not significantly different, $p = .106$.

There was a significant main effect of accuracy, $F(1, 18) = 43, 899, p < .001$, as reaction times for correct answers were faster ($M = .706$) than incorrect

answers ($M = .831$). The interaction between question and accuracy was not significant, $F(2, 38) = 0.054, p = .948$, as the difference in reaction times between the three questions was the same within correct and incorrect answers.

Task difficulty: this was defined as the number of squares presented to the participant during each trial. We hypothesised that the mean and maximum difficulty level achieved by the participants, performing at a similar accuracy, would vary based on individual capabilities in the task.

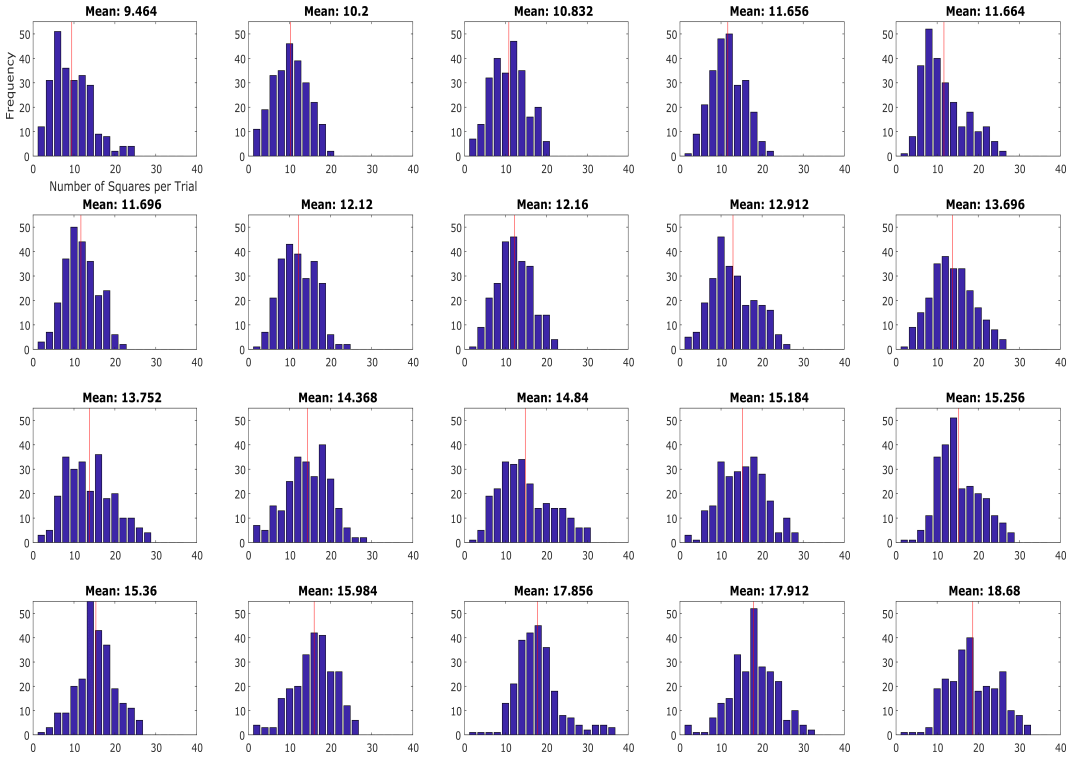

Figure A.4: Each histogram corresponds to a single participant, and the red line represents their mean difficulty.

As shown in figure A.4, a range of difficulty distributions were demonstrated by the participants, supporting our hypothesis. The averaged difficulty distribution across all participants can be found in figure A.5.

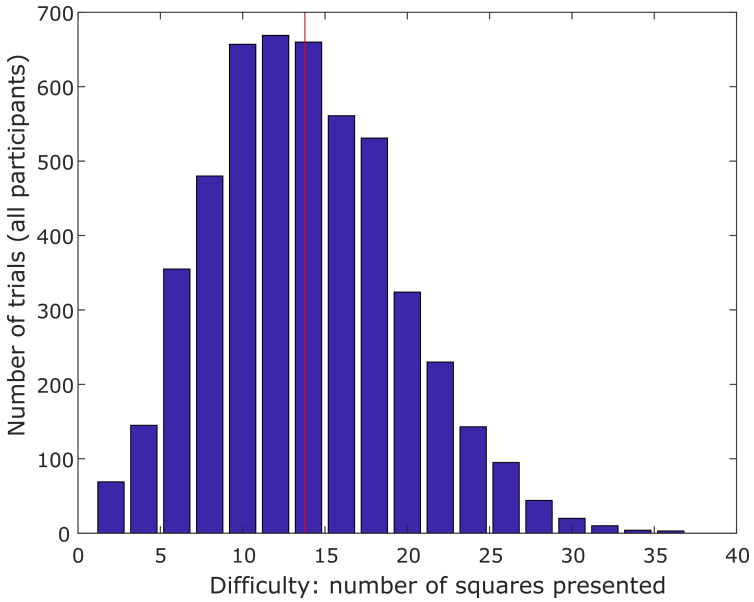

Figure A.5: Histogram of task difficulty across the group. The red line represents their mean difficulty.

A.3 EEG-fMRI pre-registration

Introduction

Change blindness is a phenomenon in which changes to a visual scene are missed if their presentation is interrupted, and was originally identified due to the fact that changes during a visual saccade are more likely to be missed (Irwin, 1991; Simons & Ambinder, 2005). In order to manipulate this in an experimental setting, the change blindness paradigm typically consists of two images displayed in quick succession that are interrupted by a blank screen or distractor image. In some instances, the second image will be identical to the first, and in others, some aspect will have changed. The complexity of these images varies across paradigms, ranging from coloured rectangles (Koivisto & Revonsuo, 2003) and coloured dots (Schankin & Wascher, 2007), to facial expressions (Eimer & Mazza, 2005), detailed visual scenes (Fernandez-Duque et al., 2003) and household objects (Busch et al., 2010). In all cases, although complete visual information is available, participants often fail to identify changes.

Most versions of the change blindness paradigm only ask participants to detect the presence of a change across two image presentations, meaning that trials can only be categorised as one of two options: seen or unseen changes. In a previous EEG only experiment, we extended the possible trial categorisation by including an additional question, where participants were asked to *localise* the change in space. Specifically, we divided the visual display into quadrants, and asked participants to select the quadrant in which the change occurred. We were therefore able to distinguish between trials where the participants could correctly identify the change, as well as its location, versus those where the identified that there was a change, but could not provide additional information about it. In line with the terminology used by Rensink (2004), Busch (2010), and others (Galpin et al., 2008; Fernandez-Duque & Thornton, 2000; Laloyaux et al., 2006; Mitroff et al., 2002), we categorised full awareness trials as *localise* trials, and partial awareness as *sense* trials.

Based on the results from this previous EEG experiment, we have evidence to suggest that trials in which the participants are able to provide complete information about the change (*localise*), are represented differently in the electrophysiological signals to trials where the participants can only detect the presence of the change (*sense*). This was particularly apparent in an early visual negativity waveform (N2pc), and a late positivity (LP). Although several papers have examined the differences between detected and undetected changes using fMRI (Beck et al., 2001; Pessoa, 2004), to our knowledge, the additional trial type of *sensing* has not been explored using fMRI. We therefore aim to identify brain areas with activations associated with the *sensing* condition, in contrast to both *blind* and *localise* trials. Given the limited spatial resolution of EEG, we hope to gain additional information about the sources of differences that we identified in the EEG waveform, using combined EEG-fMRI.

Aims and hypotheses

The first aim of this combined EEG-fMRI experiment will be to replicate previous EEG and behavioural findings in a different set of participants. The second aim will be to build on these findings with the addition of fMRI data. The third aim will be to compare the ERPs recorded from inside and outside of the scanner, as an assessment of the influence of the fMRI environment on the ERP results.

Aim 1: To identify behavioural and neural differences between full and partial awareness of colour changes in a change blindness paradigm. These hypotheses are based on the results from a previous EEG experiment using the same paradigm.

1. Reaction times will be significantly increased for incorrect responses, and will be slowest for the localisation question (Q2)
2. Participants will report higher certainty for trials where they can both correctly identify and localise the change (*localise* trials)
3. The mean and max difficulty level achieved by the participants, performing at a similar accuracy, will vary based on individual capabilities at the task
4. Certain colour changes will be easier to detect, and will therefore be associated with correct responses more frequently
5. The location of the change (left or right hemisphere) should not influence the likelihood of correct detection
6. P1 and N1 peaks will not be significantly different across awareness conditions, or across the first and second visual display
7. N2pc amplitudes will be significantly different across awareness conditions

8. Visual awareness negativity (VAN) amplitudes will be significantly different between *blind* and *localise* conditions only

9. Late positivity (LP) amplitudes will be significantly different across awareness conditions

10. Single-trial ERP values will not correlate with difficulty level or confidence

Aim 2: To build on the results gained from a previous EEG only experiment, using the same paradigm, with the addition of fMRI data

11. BOLD contrasts for *blind* versus *see* conditions collapsed will reveal parietal and frontal activations

12. BOLD contrasts for *blind* versus separate *sense* and *localise* conditions may reveal different activation patterns

13. BOLD contrasts using single-trial ERPs as parametric modulators will reveal a smaller subset of voxels from those identified above, where activation is correlated with fluctuations in single-trial EEG.

14. BOLD contrasts using confidence scores at each trial as a parametric modulator will reveal a subset of voxels, in which activation is correlated with fluctuations in confidence levels of the participant

15. BOLD contrasts using the the number of squares presented per trial as a parametric modulator will reveal a subset of voxels, in which activation is correlated with fluctuations in the difficulty level of the task

Aim 3: To compare the ERPs recorded from inside and outside of the scanner, as an assessment of the influence of the fMRI environment on the ERP results.

16. ERPs may vary in amplitude and latency when recorded inside the scanner, due to the smaller signal to noise ratio. This hypothesis is non-directed as we do not have a clear hypothesis on the effect of the MR environment on each ERP.

Materials and Methods

Participants

A sample size of between 20 and 30 participants will be collected, with the aim of acquiring at least 20 usable data sets. More than the desired quantity will be collected due to the high probability that some will be excluded on the basis of movement artifacts in the EEG, or fMRI.

Subjects will be recruited from within the School of Psychology at the University of Reading, primarily through their research participation scheme (SONA), where students take part in studies in return for credits towards their degree. This will also be advertised on school mailing lists. All participants will be right-handed and will need to pass an fMRI safety screening that excludes individuals on several accounts, for example having metal artifacts anywhere in their body, being pregnant, or having epilepsy. Due to the nature of the paradigm, participants who are colour blind will also be excluded. It is expected that several months will be dedicated to data collection, due to the limited access to the fMRI scanning equipment. This experiment has received ethical approval from the University of Reading ethics committee.

Change Blindness paradigm

This experiment will use an extended change blindness paradigm, with three response questions. In each trial, participants will be presented with two visual displays of coloured squares, interrupted by a short fixation screen. In a standard change blindness paradigm, participants are typically asked to report only if they detected a change. Our paradigm has the addition two questions, asking the participant to localise the change on the screen, and indicate the level of confidence they have in their responses during a particular trial.

Participants will be asked to fixate on a central fixation cross and identify changes between consecutive displays of coloured squares. These will be interrupted by a short fixation display, in order to facilitate the change blindness

phenomenon (see figure A.6). On change trials, one of the squares will change in colour from the first to the second display. On no change trials, the displays will be identical. This will be followed by two or three questions, depending on participant response to the first question.

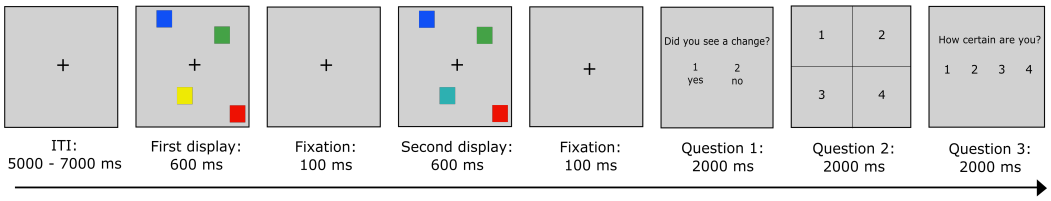

Figure A.6: Illustration of the experimental paradigm. The number of squares presented will vary from 2 to a maximum of 36, depending on participant performance. Question 1 asks 'Did you see a change?' to which participants can respond 'Yes' or 'No'. Question 2 asks participants to localise the change, based on a grid from top left to bottom right. Question 3 asks how certain participants are of their responses, ranging from '1: Very Uncertain' to '4: Very Certain'. If participants respond 'no change' to question 1, they will move instead to a filler response question that asks them to 'Press any button'.

In order to prevent floor or ceiling performance, difficulty will be modulated in real time by adding and removing squares from the display, with the assumption that more squares on the screen, or more distractors, makes the task more difficult. The maximum difficulty will be 36 squares, increasing and decreasing the difficulty in multiples of 2, to balance the number of squares presented on the left and right of the display. Performance over the previous two trials will be used to update the current trial in a two-up-two-down method; two correct answers increases the number of squares, two incorrect answers reduces the number, and one of each results in no change. In addition, if the difficulty remains constant for more than 4 consecutive trials, and therefore performance is around chance level, the difficulty will be reduced. Answers will be taken from the localisation question (question 2), as it is predicted that this will have lower accuracy than the identification question (question 1), and therefore using this will prevent the task from becoming too difficult.

Trial and block design

This study has a within-subjects repeated measures design, where each participant will complete 5 blocks of 50 trials, meaning a total of 250 trials. Of these 250 trials, 165 will contain a change in coloured square, and the remaining trials will contain no change. The ratio will not be kept at 50/50, as the trials containing the change are of most interest for analysis. However, after the experiment participants will be asked to report the percentage of trials that they believed contained a change.

Scan session

After successful screening, participants will be invited to a scanning session lasting a maximum of 2 hours and 30 minutes. The EEG set up is expected to last between 30 minutes and 1hr, depending on the length of time required to reduce impedance to the desired level of less than 10kΩ. Once completed, the participants will be placed head first supine into the MRI scanner, enclosed in a 64 channel head and neck coil. The paradigm will be displayed on a screen displayed approximately 47cm away from the centre of the scanner bore. This will be viewed by the participant through a mirror mounted onto the coil, at approximately 12cm from the participant's eyes. In their left hand, the participant will hold an alarm ball, and in their right they will hold a 4 key button box. They will use all of the 4 keys to respond to the task. All programming and task presentation will be carried out using Psychtoolbox in MATLAB, and displayed on a monitor with a refresh rate of 60Hz and resolution of 1920 x 1080.

The scanning session consists of a total of 5 blocks, during which one ERI sequence will run continuously. After each block of 50 trials, the participants will be presented with a break screen, and advised to take as long of a break as they need. The participant will be able to continue the experiment at their discretion by pressing any button on the button box. Before beginning the

main task, participants will be given a short block of 10 trials in which to practice responding to the paradigm with the button box. The data from this practice block will not be analysed. The maximum amount of time spent in the scanner depends on the reaction times of the participant, as well as the length of the breaks that are taken. Based on the maximum length of the paradigm, we estimate that the maximum length of scan time will be 50 minutes.

EEG recording

EEG data will be recoded with an MRI-compatible cap equipped with carbon-wired Ag/AgCL electrodes (Braincap MR) from 64 scalp positions according to the international 10-10 system. The reference electrode will be placed at FCz and the ground at AFz. An additional ECG electrode will be positioned on the back to measure heart rate. An MRI-compatible EEG amplifier will be used (Brain-Amp MR, Brain Products) with a sampling rate of 5000Hz. Impedance will be kept below 10kΩ for EEG channels and 5kΩ for the ECG. EEG recordings will be performed with Brain Vision Recorder Software (Brain Products) and timings kept constant using a BrainProducts SyncBox to synchronise EEG with the MRI system clock.

fMRI recording

MRI data will be acquired using a 3.0-T whole-body MRI scanner (Prisma, Siemens) and a 64 channel coil for functional imaging. Interleaved slices will be recorded using a 2D echo planar imaging (EPI) sequence [repetition time (TR) 1630ms; echo time (TE) 30ms; flip angle 90°; voxel size 3mm x 3mm; thickness 3mm; encoding direction A to P, distance factor 20%, FOV read 192mm]. A total of 30 slices will be taken per image, with transversal orientation and anterior to posterior phase encoding. Three dummy scans will be acquired at the beginning of each block. As well as the functional scans, two anatomical scans of the entire brain will be acquired [3D MPRAGE; saggital; TE 2.37ms; TR 1800ms; flip angle 8°; voxel size 0.98mm x 0.98mm; FOV read 250mm; slice

thickness 0.85mm; slices per slab 208; ascending acquisition; phase encoding direction A to P].

EEG pre-processing

Raw EEG data will be pre-processed using Brain Vision Analyzer version 2.1 (Brain Products). Correction for the MR gradient artifact will be performed using a baseline corrected sliding average of MR volumes (Allen, 2000). Removal of cardioballistic artefacts will involve the subtraction of heartbeat artifacts on a second by second basis, using a sliding average of 21 (Allen, 1998). The delay will be detected using the CBC detection solution in Analyzer. Peaks will then be detected semi-automatically, with a manual check of the algorithm's estimations. ICA will used to remove any BCG residual artifact, as well as eye movement related artifacts, inspected in semi-automatic mode. The data will be additionally filtered with a high-pass of 0.1 Hz, low-pass of 50 Hz and a notch of 50 Hz. Raw data inspection will be used to identify remaining artifacts based on a maximum voltage step of 50 μV/ms and lowest allowed activity in intervals 0.5 μV.

Segments will be taken from -200 to 7000 ms to include the whole trial, and baseline corrected from -200ms to 0. These will be time locked to the onset of the first display. Trials containing marked artifacts will be excluded from further analysis, as well as those where a response to any question is not made within the response time. The trials in which a change occurred can be divided into several conditions: '*blind*' (no identification or localisation), '*localise*' (change identification and localisation), and '*sense*' (change identification without localisation). Trials in which no change occur can be divided into 'correct rejection' (no change reported) and 'false alarm' (change incorrectly reported).

fMRI pre-processing

Images will be pre-processed using the procedure recommended in SPM12 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK). This includes re-alignment of functional images using a mean image, co-registration of the structural image to aligned functional images, segmentation of white and gray matter, normalisation of functional images using the deformation field created during segmentation, and normalisation of the functional to structural. The registration of images will be checked at each stage. Images not registered to the MNI template will be used for any single subject co-registration of EEG and fMRI.

Analysis

Variables

Table A.7 contains a description of the independent variables that are manipulated in the experimental paradigm. For example, the type of trial that is presented (change or no change), and the colour of the changed square. Table A.8 contains a description of the variables that are saved during the experiment, and exported into a text file for each participant and each block. For example, the key responses to each question, and the number of squares that were presented to the participant during each trial. Table A.9 contains a description of other variables that are calculated using the variables recorded (as shown in table A.8). For example, the mean accuracy level achieved by each participant, and the mean difficulty or number of squares that they are presented during the experiment.

Behavioural

All results will be reported using the Greenhouse-Geisser correction if Mauchley's test of sphericity is to be significant. Post-hoc p-values will be corrected using the false discovery rate correction.

Hypothesis 1: Reaction times will be significantly increased for incorrect responses, and will be slowest for the localisation question (Q2)

- Reaction times will be compared over question and awareness using a repeated measures ANOVA calculated for question (1/2/3) and accuracy (correct/incorrect)

Hypothesis 2: Participants will report higher certainty for trials where they can both correctly identify and *localise* the change

- The percentage of trials across conditions will be compared with a repeated measures ANOVA for accuracy (correct/incorrect localisation) and certainty (certain/uncertain).

Hypothesis 3: The mean and maximum difficulty level achieved by the participants, performing at a similar accuracy, will vary based on individual capabilities in the task

- The standard deviation of the mean and maximum difficulty will be calculated across participants
- Difficulty histograms will be plotted for each participant
- Difficulty over trials will be correlated with detection accuracy, localisation accuracy, and D'prime scores

Hypothesis 4: Certain colour changes will be easier to detect, and will therefore be associated with correct responses more frequently

- A chi-square test will be used to compare the frequency of each colour change occurring across awareness conditions (*see*/*blind*). Due to the large number of possible random colour changes, *localise* and *sense* trials will be combined into *see* trials. Each possible colour change pairing will be coded with a number between 1 - 21

Hypothesis 5: The location of the change should not influence the likelihood of correct detection

- A chi-square test will be used to compare the frequency of each location occurring across awareness conditions (*see*/*blind*). Due to the large number of possible change locations, *localise* and *sense* trials will be combined into *see* trials

- Within the incorrectly *localise* trials, we will run a paired t-test to determine if participants had more trials where they were still able to identify the correct hemisphere (left or right). For example, if the change occurred in the top left, were they more likely to identify the bottom left hemisphere than the top right

ERP

Hypothesis 6: P1 and N1 peaks will not be significantly different across awareness conditions, or across the first and second visual display

- The latency of P1 and N1 peaks will be identified based on a grand average of all conditions (Luck & Gaspelin, 2017), at electrodes P07 and P08

- P1 amplitudes will be compared across display and awareness condition using 2x3 repeated measures ANOVA: display (first/second) and awareness (*blind*/*localise*/*sense*) Post-hoc comparisons will compare significant differences between all pairs of awareness conditions across displays

Hypothesis 7: N2pc amplitudes will be significantly different across awareness conditions

- The N2pc is pre-defined as the mean within 200-400ms after the second display, at contralateral and ipsilateral occipital electrodes PO7 and PO8

- N2pc differences across hemispheres will be analysed with a 2x3 repeated measures ANOVA: hemisphere (contralateral/ipsilateral) and awareness (*blind*/*localise*/*sense*). Post-hoc comparisons will compare significant differences between all pairs of awareness conditions.

Hypothesis 8: Visual awareness negativity (VAN) amplitudes will be significantly different between *blind* and *localise* conditions only

- The VAN is pre-defined as the mean within 130-330ms (Busch et al., 2010) after the second display, at central parietal electrodes Cz, Pz, CPz
- VAN Differences between awareness conditions will be calculated with a repeated measures ANOVA for awareness condition (*blind*/*localise*/*sense*). Post-hoc comparisons will compare significant differences between all pairs of awareness conditions

Hypothesis 9: Late positivity (LP) amplitudes will be significantly different across awareness conditions

- The LP is pre-defined as the mean within 400-600ms (Busch et al., 2010) after the second display, at central parietal electrodes Cz, Pz, CPz
- LP differences between awareness conditions will be calculated with a repeated measures ANOVA for awareness condition (*blind*/*localise*/*sense*). Post-hoc comparisons will compare significant differences between all pairs of awareness conditions

Single-trial EEG

For each participant, time courses will be constructed from the single-trial values of each ERP, at each channel, as an average of the data within the pre-defined ERP windows. Each single-trial value will be baseline corrected by subtracting the mean of the trial from which they were extracted.

Hypothesis 10: Single-trial ERP values will not correlate with difficulty level or confidence

- Single-trial ERP values for each participant, for each ERP, at each channel, will be correlated with the difficulty level across trials. P-values will be corrected for false positives using false discovery rate correction.

- Single-trial ERP values for each participant, for each ERP, at each channel, will be correlated with the confidence level across trials. P-values will be corrected for false positives using false discovery rate correction.

fMRI

Following the standard method of fMRI data analysis, general linear models (GLM) with event-related designs will be conducted in SPM12, to identify voxels activated in response to each awareness condition. Regressors will be created by convolving the stimulus onset times of each awareness level individually with the hemodynamic response function, across all blocks. The onset will relate to the second display of colour squares, meaning the display in which the change occurs. Scanning will be continuous, so blocks can be considered in the same regressor, unless for unforeseen reasons the scan has to be stopped during the experiment. Additional regressors for subject motion will be included. Aly et al. (2013 cite) found that hippocampal activation during a change detection task of complex visual scenes increased linearly with participant confidence in suspected changes. We will therefore use participant confidence to idetify voxels that demonstrate this trend.

In a change blindness fMRI experiment, Beck et al. (2001) identified bilateral superior parietal lobule (BA 7; talairach coordinates -24, -60, 60) and right dorsolateral prefrontal (BA 46; 51, 30, 24) activations in a contrast of change detected versus change missed. Heuttel et al., (2001) identified activation in a range of occipital-parietal (fusiform gyri, 28, -66, -16; intraparietal sulcus, 22,-72,34; parahipocampal guris, -22,-34,-8) and frontal areas (precentral gyrus, -28,-8,54; medial frontal gyrus, 8,6,54; superior frontal gyrus, 16,66,-10; precentral sulculs, 48,2,30; inferior frontal gyrus, -42,16,26). Our hypotheses about the expected activation for detected versus undetected changes are based on these results.

Hypothesis 11: BOLD contrasts for *blind* versus *see* conditions collapsed

will reveal parietal and frontal activations, based on results from Beck et al. (2001) and Pessoa (2004)

- A GLM will be constructed with separate regressors for the onset of the expected hemodynamic responses to *see* and *blind* trials. These will be compared with a standard contrast (1,-1). We will also look at the main effects (1,0).

Hypothesis 12: BOLD contrasts for *blind* versus separate *sense* and *localise* conditions may reveal different activation patterns

- A GLM will be constructed with separate regressors for the onset of the expected hemodynamic responses to *blind*, *localise*, and *sense* trials. All pairs will be compared with contrasts (1,-1). We will also look at the main effects (1,0).

Hypothesis 13: BOLD contrasts using confidence scores at each trial as a parametric modulator will reveal a subset of voxels, in which activation is correlated with fluctuations in confidence levels of the participant

- A a GLM regressor will be constructed using the confidence scores over time as a parametric modulator

Hypothesis 14: BOLD contrasts using the the number of squares presented per trial as a parametric modulator will reveal a subset of voxels, in which activation is correlated with fluctuations in the difficulty level of the task.

- A a GLM regressor will be constructed using the number of squares presented per trial as a parametric modulator

EEG-informed fMRI

Hypothesis 15: BOLD contrasts using single-trial ERPs as parametric modulators will reveal a smaller subset of voxels from those identified above, where activation is correlated with fluctuations in single-trial EEG

- For each ERP, a GLM regressor will be constructed for all stimuli onset times, with normalised single-trial ERP values as parametric modulators

Influence of the fMRI environment

Hypothesis 16: ERPs may vary in amplitude and latency when recorded inside the scanner, due to the smaller signal to noise ratio

- ERP amplitudes across recording condition will be analysed using a mixed repeated measures ANOVA: within-subject factors of awareness level (*blind*/*localise*/*sense*), and ERP type (P1/N1 /N2pc/LP/VAN). Between-subject factor: recording condition (inside/outside scanner)
- ERP latencies will be analysed in the same way as ERP amplitudes, but with latency as the dependent variable.

Inclusion and exclusion criteria

Participants who accurately report 60% change trials may be considered to have had an advantage over other participants, and may be excluded from analysis if their accuracy is more than three standard deviations from the mean accuracy across participants. If a participant has less than 20 usable trials in any of the awareness conditions, across behavioural, EEG, and fMRI data, they will be considered for exclusion. Factors potentially causing less than 20 usable trials include:

- Participant response behaviour (floor/ceiling accuracy, response bias, or lack of attention)
- Large eye movement, muscle, or fMRI based artifacts in the EEG data, that cannot be successfully removed during pre-processing
- Large movement artifacts in the fMRI data, over 2 voxels in any direction
- Any unforeseeable technical faults that may occur

Statement of study progress

At the date of submission to the Open Science Framework (27/06/18), the following has been completed:

- 17 data sets have been collected. However, no decision on the usability of these data sets has been made
- EEG pre-processing has been completed for 9 data sets
- fMRI pre-processing has been completed for 6 data sets
- First level fMRI analysis has been completed for 1 data set, to establish that all required timing and trial information was correctly saved
- No EEG analysis has been completed
- Behavioural analysis scripts have been run to determine the timings of each trial type. However, other behavioural results have not been examined

Table A.7: Manipulated Variables: the independent variables that are manipulated in the experimental paradigm

Variable Name	Type	Description	Details	Variable Contents
Trial type (IV)	Binary	Each trial can be classified as a change trial, or a no change trial. The order of presentation is randomly selected	165 trials contain a change, out of a total of 250 trials for each participant	1 or 0
Colour of change (IV)	Nominal	The colour of the change is randomly selected from 7 possible colours. The colour before and after the change is stored	Possible colours: red, blue, cyan, magenta, white, green, yellow	r/b/c/m/w/g/y
Location of change (IV)	Nominal	The location of the change is randomly selected, but with equal numbers on the left and right of the fixation	36 possible locations represented as indices, and by the quadrants of the screen: 'TL', 'TR', 'BL', 'BR'	Indices: 1 to 36 Quadrants: TL/TR/BL/BR
Difficulty (IV)	Discrete	The number of squares presented during each trial	The maximum difficulty possible is 36. The difficulty is always a multiple of 2	A multiple of 2 from 2 to 36
Distractors (IV)	Nominal	The colour and location of the distractor squares. The number is dependent on the difficulty	All locations are stored, so the presence of an rbg value indicates the presence of a distractor at that location	000 or rgb
Colour change code (IV)	Nominal	Each possible colour change, in each direction, is coded with a number		1 to 41

Table A.8: Recorded Variables: the variables that are saved during the experiment, and exported into a text file for each participant and each block

Variable Name	Type	Description	Details	Contents
Identification (DV)	Nominal	Response to question 1: 'Did you see a change?'	1 = yes 2 = no Nan = did not respond in time	1, 2, or NaN
Localisation (DV)	Nominal	Response to question 2: 'Where did the change occur?'	1 = top left 2 = top right 3 = bottom left 4 = bottom right Nan = did not respond in time	1, 2, 3, 4, or NaN
Response Certainty (DV)	Ordinal	Response to question 3: 'How certain are you are your response?'	1 = very uncertain 2 = uncertain 3 = certain 4 = very certain Nan = did not respond in time	1, 2, 3, 4, or NaN
Running Accuracy (DV)	Discrete	The accuracy used to update the difficulty, based on responses to question 2 in the previous 2 trials	0 = previous two trials incorrect, next trial decreases in difficulty 1 = previous two trials correct, next trial increases in difficulty 0.5 = one correct and one incorrect, so no change	0, 0.5, or 1
Difficulty (DV)	Discrete	The number of squares presented during each trial. This varies as a function of a participant's running accuracy	The maximum difficulty possible is 36. The difficulty is always a multiple of 2	2 to 36

Table A.9: Calculated Variables: other variables that are calculated for each participant using the recorded variables

Variable Name	**Type**	**Description**	**Details**	**Contents**
Awareness (IV)	Nominal	Trial classification based on participant responses and trial types	·*blind*: failure to identify a change ·*sense*: successful identification, but incorrect localisation ·*localise*: correct identification and localisation ·*see*: *sense* and *localise* combined ·False alarm: false identification in a no change trial ·Correct rejection: correct detection of a no change trial	1 or 0
Awareness (DV)	Continuous	The number and percentage of trials in each condition	·*blind*: failure to identify a change ·*sense*: successful identification, but incorrect localisation *localise*: correct identification and localisation ·*see*: *sense* and *localise* combined ·False alarm: false identification in a no change trial ·Correct rejection: correct detection of a no change trial	
Accuracy (DV)	Continuous	The mean accuracy achieved during the experiment	·Yes no accuracy: for the identification question ·Localisation accuracy: for the localisation question	
Difficulty (DV)	Discrete Ordinal (frequency)	The number of squares presented to the participant on each trial The frequency of trials at each difficulty level	Mean and maximum difficulty	2 to 36 Frequency table (response, frequency, percentage)
No response (DV)	Discrete	The total number of questions where the participant did not respond	For each of the three questions: identification, location, and certainty	
DPrime (DV)	Continuous	The response pattern of the participant	$d = z(\text{hit rate}) - z(\text{false alarm rate})$	
Criterion (DV)	Continuous	The response bias of the participant	$c = -0.5 * (z(\text{hit rate}) + z(\text{false alarm rate}))$	
Reaction time (DV)	Continuous	The response time of the participant (seconds)	·For each of the three questions: identification, location, and certainty ·For each of the trial types: change and no change ·For each of the awareness conditions: *see*, *blind*, *localise*, and *sense*	
Certainty (DV)	Ordinal (frequency) Discrete (mean)	The frequency of responses for each certainty level The mean certainty response, where 1 is very uncertain, 2 is uncertain, 3 is certain, and 4 is very certain	·For each certainty response: very uncertain, uncertain, certain, very certain ·For each of the trial types: change and no change ·For each of the awareness conditions: *see*, *blind*, *localise*, and *sense*	Frequency table (response, frequency, percentage) Rounded to the nearest whole number

A.4 EEG-fMRI experiment

Additional behavioural analysis from the combined EEG-fMRI experiment. Results will be summarised with reference to the related hypotheses from the pre-registration.

Hypothesis 1: Reaction times will be significantly increased for incorrect responses, and will be slowest for the localisation question (Q2). Reaction times were compared using a repeated measures ANOVA calculated for question (1/2/3) and accuracy (correct/incorrect).

The main effect of question was not significant, $F(2, 38) = 3.077, p = .058$. (Q1: $M = .585$, Q2: $M = .578$, Q3: $M = .652$). There was a significant main effect of accuracy, $F(1, 18) = 43,899, p < .001$, as reaction times for correct answers were faster ($M = .568$) than incorrect answers ($M = .642$). The interaction between question and accuracy was also significant, $F(2, 38) = 3.857, p = .030$. This was driven by the fact that accuracy did not significantly modulate responses for Q2 (correct $M = .567$, incorrect $M = .588$). However, for Q1 and Q3, incorrect answers were slower than correct answers (Q1 correct $M = .535$, incorrect $M = .636$, Q3 correct $M = .602$, incorrect $M = .702$).

These findings support the hypothesis that incorrect answers were slower than correct answers, but only for Q1 and Q3. We failed to support the hypothesis that the localisation question (Q2) would have the slowest responses, as the certainty question (Q3) had the largest reaction times overall.

Hypothesis 2: Participants will report higher certainty for trials where they can both correctly identify and localise the change (*localise* trials). The percentage of trials across conditions was compared with a repeated measures ANOVA for accuracy (correct/incorrect localisation) and certainty (certain/uncertain).

There was a significant main effect of accuracy, $F(1, 19) = 6.788, p = .017$, with a higher percentage of correct ($M = 49.924$) than incorrect tri-

als ($M = 48.779$). The main effect of certainty was also significant, $F(1, 19) = 12.720, p = .002$, with a higher percentage of certain trials ($M = 62.873$) than uncertain ($M = 35.830$). The interaction between accuracy and certainty was also significant, $F(1, 19) = 52.836, p < .001$. This was driven by the similar percentage of certain ($M = 48.599$) and uncertain trials ($M = 48.960$) when participants incorrectly localised the change. Conversely, when participants correctly localise the change, they had more certain ($M = 77.148$) than uncertain trials ($M = 22.700$).

Our hypothesis was supported, as participants were more certain of their responses when they were able to localise the change (*localise* trials). When they were unable to localise the change (*sense* trials), they were equal in their responses of certain and uncertain.

Hypothesis 3: The mean and maximum difficulty level achieved by the participants, performing at a similar accuracy, will vary based on individual capabilities in the task. The standard deviation of the mean and maximum difficulty will be calculated across participants, difficulty histograms will be plotted for each participant, and difficulty over trials will be correlated with detection accuracy, localisation accuracy, and D'prime scores.

As shown in figure A.7, a range of difficulty distributions were demonstrated by the participants, supporting our hypothesis. The averaged difficulty distribution across all participants can be found in figure A.8. The mean difficulty level given to each participant ranged from 6 to 23 ($M = 16, SD = 4$), with the maximum difficulty experienced by each participant ranging from 18 to 36 ($M = 27, SD = 5$). Mean difficulty correlated with mean location accuracy ($r = .590, p = .008$) and d'prime ($r = -.601p = .005$), but not with mean detection accuracy ($r = -.371, p = .107$). Maximum difficulty also correlated with mean location accuracy ($r = .537, p = .015$) and d'prime ($r = -.482p = .031$), but not with mean detection accuracy ($r = -.349, p = .131$).

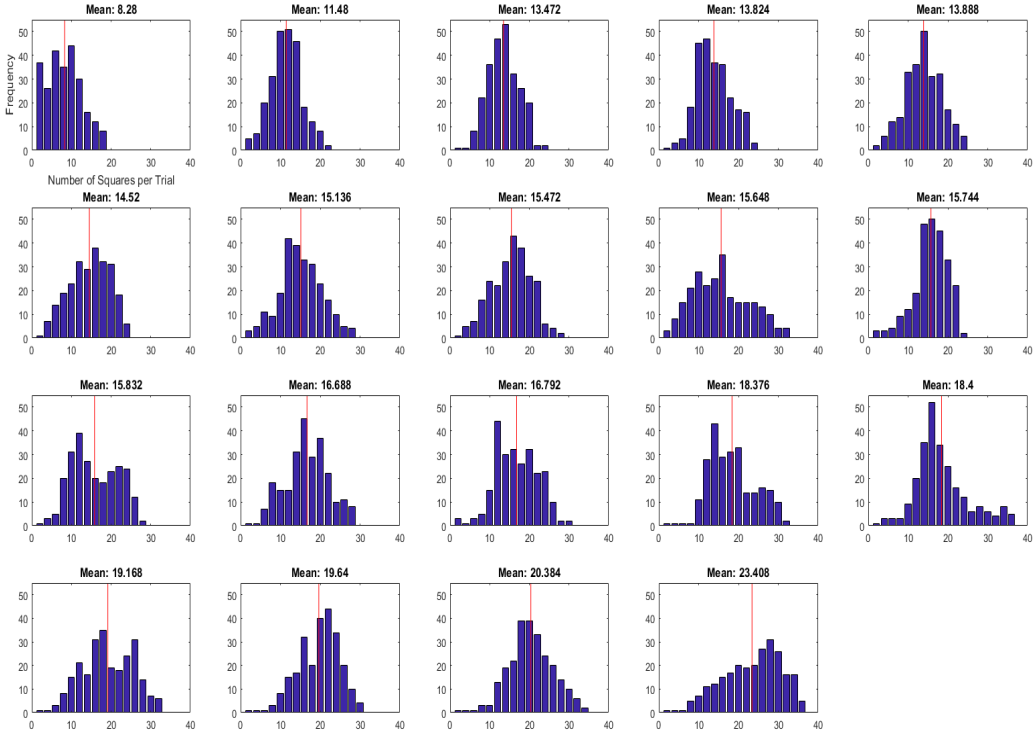

Figure A.7: Each histogram corresponds to a single participant, and the red line represents their mean difficulty.

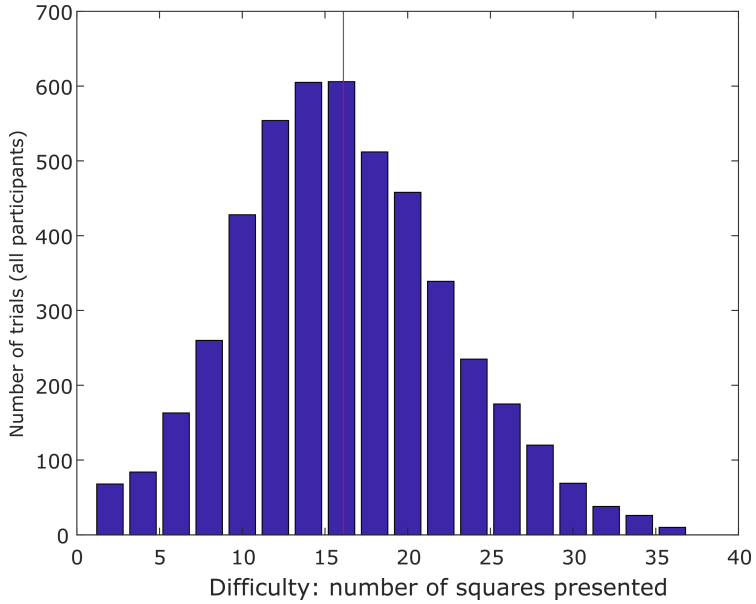

Figure A.8: Histogram of task difficulty across the group. The red line represents the mean difficulty.

Hypothesis 4: Certain colour changes will be easier to detect, and will therefore be associated with correct responses more frequently. A chi-square test was used to compare the frequency of each colour change occurring across

awareness conditions (*see*/*blind*). Due to the large number of possible random colour changes, *localise* and *sense* trials were combined into *see* trials. Each possible colour change was coded with a number between 1 - 21 (such that a change in either direction was codes with the same number). Colour changes with an adjusted standardised residual greater than 1.96 were considered to have a significant difference between conditions.

There was a significant effect of colour change on awareness, $\chi^2(20) = 120.342, p < .001$. Colour change pairs with a significantly higher number of *blind* trials were: yellow/green, yellow/cyan, green/cyan, blue/cyan, white/cyan. Colour change pairs with a significantly higher number of *see* trials were: magenta/red, cyan/red, blue/red, white/red.

The main finding from these results is that colour changes including red were much more likely to be detected, given the prevalence of colour pairings that included red in the *see* condition. The other findings is that colour changes including cyan were more likely to be missed, as there were more *blind* trials with cyan in the colour pairing.

Hypothesis 5: The location of the change should not influence the likelihood of correct detection. Also, within the incorrectly localised trials (*sense* trials), we will run a paired t-test to determine if participants had more trials where they were still able to identify the correct side of the display (left/right).

During the localisation question, participants were asked to identify the location of the change square using a 2 x 2 response grid. The display was therefore divided into four squares: upper left, lower left, upper right, and lower right. During *sense* trials, participants reported an incorrect location for the change. The purpose of this analysis was to determine if, for incorrect answers, participants were more likely to identify the correct side of the display (left/right) or the correct field (upper/lower).

Nine out of twenty participants had more trials where they correctly identified the correct side of the display (left/right), but the incorrect hemisphere

(upper/lower). The remaining eleven had more trials where they were more likely to report the correct hemisphere (upper/lower) but incorrect side (left/right). In an paired t-test, we found no significant difference between the number of trials in each condition, $t(19) = -0.697, p = 0.494$. There is therefore no clear distinction between the likelihood of identifying the correct side or hemisphere for incorrectly localise trials.

Hypothesis 10: Single-trial ERP values will not correlate with difficulty level or confidence. Single-trial ERP values for each participant, for each ERP, at each channel, will be correlated with the difficulty and certainty levels across trials. P-values will be corrected for false positives using false discovery rate correction.

After correcting for multiple comparisons using FDR correction ($q = .05$), no significant correlations were found for either task difficulty or participant certainty.

A.5 EEG comparison

Accuracy and Difficulty

Accuracy for question 1, in which participants had to identify a change, had a mean of 51.31% (range $= 32 - 73\%, SD = 11$). Accuracy for question 2, in which participants had to *localise* the change, had a mean of 70.93% (range $= 55 - 87\%, SD = 8$). The mean difficulty level given to each participant ranged from 6 to 24 squares ($M = 15, SD = 4$) had a mean of 14.6 squares displayed, with the maximum difficulty experienced by each participant ranging from 18 to 36 ($M = 27, SD = 5$).

D'prime scores ranged from .744 to 2.31 ($M = 1.28, SD = .37$). Three participants had a negative criterion, meaning that they had a response bias towards false alarms. All other participants had positive criterion, indicating a conservative response strategy ($M = .61, SD = .33$). D'prime scores were sig-

nificantly different to zero in a one-sampled t-test, indicating that participants could identify between change and no change trials, $t(38) = 23.192, p < .001$. The number of false alarms ranged from 1 to 28 ($M = .60, SD = .37$), which correlated with the number of incorrectly localised trials ($r = .633, p < .001$).

Mean difficulty correlated with mean location accuracy ($r = .391, p = .013$) and d'prime ($r = -. - 0.361, p = .022$), but not with mean detection accuracy ($r = -.128, p = .431$). Maximum difficulty did not correlate with mean location accuracy ($r = 0.270, p = .092$), d'prime ($r = -.292p = .068$), or mean detection accuracy ($r = -.091, p = .577$).

Comparison of *sense* and *false alarm* trials

The percentage of *false alarm* trials ($13.44\% \pm 10.03$) was lower than the percentage of *sense* trials ($29.19\% \pm 7.86$) $t(39) = -9.96, p < .001, g_{rm} = 1.70$, suggesting that *sense* trials occurred more often than participants made false alarms. However, the percentage of false alarms was significantly correlated with the percentage of *sense* trials ($r = .396, p = .011$).

Reaction times for *sense* and *false alarm* trials were compared, to determine if *sense* trials were different to trials where the participant incorrectly reported a change during a no change trial. Reaction times for all *sense* trials (0.690 ± 0.166 s), regardless of certainty, were not significantly different to *false alarm* trials (0.721 ± 0.173 s), $t(39) = -1.579, p = .123, g_{rm} = 0.17$. However, *sense certain* trials (0.580 ± 0.140 s) were significantly faster than *false alarm* trials, $t(38) = -6.515, p < .001, g_{rm} = 0.87$. Therefore, when participants were certain that a change occurred, they responded more quickly than when they were simply making a false alarm.

Reaction times for *sense certain* trials (0.580 ± 0.140 s) were also significantly faster than *false alarm uncertain* trials (0.778 ± 0.214 s), $t(38) = -6.480, p < 001, g_{rm} = 0.29$. However, this may be explained by the general finding that, across all conditions, certain trials (0.587 ± 0.143) were faster than uncertain trials (0.778 ± 0.171 s), ($t(39) = -10.701, p < .001, g_{rm} = 1.16$)

Comparison of *sense* and *blind* trials

Reaction times for *sense* trials (0.690 ± 0.166 s) were not significantly different to *blind* trials (0.698 ± 0.181 s), $t(39) = -0.276, p = .784, g_{rm} = 0.00$. However, reaction times for *sense certain* trials (0.580 ± 0.140 s) were significantly faster than *blind* trials, $t(38) = -4.432, p < .001, g_{rm} = 0.66$. Therefore, on trials where the participant did not see the change (*blind*), they responded more slowly than when they suspected a change but could not provide additional information about it (*sense*).

Comparison of *blind* trials and no-change trials

Out of the 40 participants included in the analysis, 30 were slower to respond when they were *blind* to the change, compared to no-change trials (75%). Reaction times for *blind* trials were significantly slower than no-change trials (0.661 ± 0.176 s), $t(39) = 4.030, p < .001, g_{rm} = 0.16$. Similarly, *blind uncertain* trials ($.804 \pm 0.223$ s) were significantly slower than no-change trials, $t(39) = 5.838, p < .001, g_{rm} = 0.83$. Therefore, despite being *blind* to the change, the presence of a change in the display increased reaction times, particularly for trials where the participant was uncertain.

Post-hoc EEG analysis

We compared ERP amplitudes across trial type (change/no change) and certainty response (certain/uncertain). In our previous analysis, trials were categorised based on a combination of detection and localisation responses. However, there is a chance that participants responded incorrectly and therefore that trials were wrongly categorised. We hypothesised that participant certainty may also be an indication of participant awareness. This is supported by the finding that participants were certain of their response in 70% of correctly localised trials, compared to 50% of incorrectly localised trials.

The only significant result was a post-hoc comparison between LP ampli-

tude for change trials where participants were certain ($M = 2.05, SD = 3.33$) versus uncertain ($M = 1.28, SD = 2.72$), $t(39) = 2.019, p = .049$. Therefore, when participants were certain about the change, the amplitude of the LP was greater. For no change trials, LP amplitudes did not vary with participant certainty, $t(39) = 0.394, p = .696$.

Bibliography

Allen, P. J., Josephs, O., & Turner, R. (2000). A Method for Removing Imaging Artifact from Continuous EEG Recorded during Functional MRI. *NeuroImage*, *12*(2), 230–239.

Allen, P. J., Polizzi, G., Krakow, K., Fish, D. R., & Lemieux, L. (1998). Identification of EEG events in the MR scanner: the problem of pulse artifact and a method for its subtraction. *Neuroimage*, *8*(3), 229–239.

Angelone, B. L., Levin, D. T., & Simons, D. J. (2003). The Relationship between Change Detection and Recognition of Centrally Attended Objects in Motion Pictures. *Perception*, *32*(8), 947–962.

Angenstein, F., Kammerer, E., & Scheich, H. (2009). The BOLD Response in the Rat Hippocampus Depends Rather on Local Processing of Signals than on the Input or Output Activity. A Combined Functional MRI and Electrophysiological Study. *Journal of Neuroscience*, *29*(8), 2428–2439.

Attal, Y., Bhattacharjee, M., Yelnik, J., Cottereau, B., Lefevre, J., Okada, Y., Bardinet, E., Chupin, M., & Baillet, S. (2007). Modeling and detecting deep brain activity with MEG & EEG. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, (pp. 4937–4940). IEEE.

Attwell, D. & Iadecola, C. (2002). The neural basis of functional brain imaging signals. *Trends in neurosciences*, *25*(12), 621–625.

Bagshaw, A. P. & Warbrick, T. (2007). Single trial variability of EEG and fMRI responses to visual stimuli. *NeuroImage*, *38*(2), 280–292.

Ball, F. & Busch, N. A. (2015). Change detection on a hunch: Pre-attentive vision allows "sensing" of unique feature changes. *Attention, Perception, & Psychophysics*, *77*(8), 2570–2588.

Baumann, O., Endestad, T., Magnussen, S., & Greenlee, M. W. (2008). Delayed discrimination of spatial frequency for gratings of different orientation: behavioral and fMRI evidence for low-level perceptual memory stores in early visual cortex. *Experimental Brain Research*, *188*(3), 363–369.

Beck, D. M., Rees, G., Frith, C. D., & Lavie, N. (2001). Neural correlates of change detection and change blindness. *Nature neuroscience*, *4*(6), 645–650.

Becker, R., Ritter, P., Moosmann, M., & Villringer, A. (2005). Visual evoked potentials recovered from fMRI scan periods. *Human Brain Mapping*, *26*(3), 221–230.

Benjamini, Y. & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300.

Bonmassar, G., Hadjikhani, N., Ives, J. R., Hinton, D., & Belliveau, J. W. (2001). Influence of EEG electrodes on the BOLD fMRI signal. *Human Brain Mapping*, *14*(2), 108–115.

Bregadze, N. & Lavric, A. (2006). ERP differences with vs. without concurrent fMRI. *International Journal of Psychophysiology*, *62*(1), 54–59.

Breitmeyer, B. G., Kropfl, W., & Julesz, B. (1982). The existence and role of retinotopic and spatiotopic forms of visual persistence. *Acta Psychologica*, *52*(3), 175–196.

Bridgeman, B., Hendry, D., & Stark, L. (1975). Failure to detect displacement of the visual world during saccadic eye movements. *Vision Research*, *15*(6), 719–722.

Busch, N. A. (2009). Objective and Subjective Measures of Change Blindness: Where Are the Real Pitfalls? *Journal of Cognitive Neuroscience*, *22*(9), 1903–1905.

Busch, N. A. (2010). Objective and Subjective Measures of Change Blindness: Where Are the Real Pitfalls? *Journal of Cognitive Neuroscience*, *22*(9), 1903–1905.

Busch, N. A. (2013). The fate of object memory traces under change detection and change blindness. *Brain Research*, *1520*, 107–115.

Busch, N. A., Dürschmid, S., & Herrmann, C. S. (2010). ERP effects of change localization, change identification, and change blindness:. *NeuroReport*, *21*(5), 371–375.

Busch, N. A., Fründ, I., & Herrmann, C. S. (2009). Electrophysiological Evidence for Different Types of Change Detection and Change Blindness. *Journal of Cognitive Neuroscience*, *22*(8), 1852–1869.

Busch, N. A., Fründ, I., & Herrmann, C. S. (2010). Electrophysiological Evidence for Different Types of Change Detection and Change Blindness. *Journal of Cognitive Neuroscience*, *22*(8), 1852–1869.

Bénar, C.-G., Schön, D., Grimault, S., Nazarian, B., Burle, B., Roth, M., Badier, J.-M., Marquis, P., Liegeois-Chauvel, C., & Anton, J.-L. (2007). Single-trial analysis of oddball event-related potentials in simultaneous EEG-fMRI. *Human Brain Mapping*, *28*(7), 602–613.

Calhoun, V., Adali, T., Pearlson, G., & Kiehl, K. (2006). Neuronal chronometry of target detection: Fusion of hemodynamic and event-related potential data. *NeuroImage*, *30*(2), 544–553.

Caplovitz, G. P., Fendrich, R., & Hughes, H. C. (2008). Failures to see: Attentive blank stares revealed by change blindness. *Consciousness and Cognition*, *17*(3), 877–886.

Chetverikov, A., Kuvaldina, M., MacInnes, W. J., Jóhannesson, I., & Kristjánsson, (2018). Implicit processing during change blindness revealed with mouse-contingent and gaze-contingent displays. *Attention, Perception, & Psychophysics*, *80*(4), 844–859.

Cohen, M. X. (2014). *Analyzing neural time series data: theory and practice*. Issues in clinical and cognitive neuropsychology. Cambridge, Massachusetts: The MIT Press.

Comi, E., Annovazzi, P., Silva, A. M., Cursi, M., Blasi, V., Cadioli, M., Inuggi, A., Falini, A., Comi, G., & Leocani, L. (2005). Visual evoked potentials may be recorded simultaneously with fMRI scanning: A validation study. *Human Brain Mapping*, *24*(4), 291–298.

Cowan, N., Naveh-Benjamin, M., Kilb, A., & Saults, J. S. (2006). Life-Span Development of Visual Working Memory: When is Feature Binding Difficult? *Developmental psychology*, *42*(6), 1089–1102.

Daunizeau, J., Grova, C., Marrelec, G., Mattout, J., Jbabdi, S., Pélégrini-Issac, M., Lina, J.-M., & Benali, H. (2007). Symmetrical event-related EEG/fMRI information fusion in a variational Bayesian framework. *NeuroImage*, *36*(1), 69–87.

De Vos, M., Thorne, J. D., Yovel, G., & Debener, S. (2012). Let's face it, from trial to trial: Comparing procedures for N170 single-trial estimation. *NeuroImage*, *63*(3), 1196–1202.

Debener, S. (2005). Trial-by-Trial Coupling of Concurrent Electroencephalogram and Functional Magnetic Resonance Imaging Identifies the Dynamics of Performance Monitoring. *Journal of Neuroscience*, *25*(50), 11730–11737.

Debener, S., Mullinger, K. J., Niazy, R. K., & Bowtell, R. W. (2008). Properties of the ballistocardiogram artefact as revealed by EEG recordings at 1.5, 3 and 7 T static magnetic field strength. *International Journal of Psychophysiology*, *67*(3), 189–199.

Debener, S., Ullsperger, M., Siegel, M., & Engel, A. K. (2006). Single-trial EEG–fMRI reveals the dynamics of cognitive function. *Trends in Cognitive Sciences*, *10*(12), 558–563.

Dehaene, S. (2018). The Error-Related Negativity, Self-Monitoring, and Consciousness. *Perspectives on Psychological Science*, *13*(2), 161–165.

Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in Cognitive Sciences*, *10*(5), 204–211.

Delorme, A. & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21.

Di Russo, F., Martínez, A., & Hillyard, S. A. (2003). Source Analysis of Event-related Cortical Activity during Visuo-spatial Attention. *Cerebral Cortex*, *13*(5), 486–499.

Digirolamo, G. J. & Hintzman, D. L. (1997). First impressions are lasting impressions: A primacy effect in memory for repetitions. *Psychonomic Bulletin & Review*, *4*(1), 121–124.

Dijk, H. v., Schoffelen, J.-M., Oostenveld, R., & Jensen, O. (2008). Prestimulus Oscillatory Activity in the Alpha Band Predicts Visual Discrimination Ability. *Journal of Neuroscience*, *28*(8), 1816–1823.

Driver, J., Baylis, G. C., & Rafal, R. D. (1992). Preserved figure-ground segregation and symmetry perception in visual neglect. *Nature*, *360*(6399), 73–75.

Eapen, M., Zald, D., Gatenby, J., Ding, Z., & Gore, J. (2011). Using High-Resolution MR Imaging at 7t to Evaluate the Anatomy of the Midbrain Dopaminergic System. *American Journal of Neuroradiology*, *32*(4), 688–694.

Eckert, M. A., Menon, V., Walczak, A., Ahlstrom, J., Denslow, S., Horwitz, A., & Dubno, J. R. (2009). At the heart of the ventral attention system: The right anterior insula. *Human Brain Mapping*, *30*(8), 2530–2541.

Eichele, T., Calhoun, V. D., Moosmann, M., Specht, K., Jongsma, M. L., Quiroga, R. Q., Nordby, H., & Hugdahl, K. (2008). Unmixing concurrent EEG-fMRI with parallel independent component analysis. *International Journal of Psychophysiology*, *67*(3), 222–234.

Eimer, M. (2000). Effects of face inversion on the structural encoding and recognition of faces: Evidence from event-related brain potentials. *Cognitive Brain Research*, *10*(1), 145–158.

Eimer, M. & Mazza, V. (2005). Electrophysiological correlates of change detection. *Psychophysiology*, *42*(3), 328–342.

Ekstrom, A., Suthana, N., Millett, D., Fried, I., & Bookheimer, S. (2008). Correlation Between BOLD fMRI and Theta-Band Local Field Potentials in the Human Hippocampal Area. *Journal of Neurophysiology*, *101*(5), 2668–2678.

Fellner, M.-C., Volberg, G., Mullinger, K. J., Goldhacker, M., Wimber, M., Greenlee, M. W., & Hanslmayr, S. (2016). Spurious correlations in simultaneous EEG-fMRI driven by in-scanner movement. *NeuroImage*, *133*, 354–366.

Fernandez-Duque, D., Grossi, G., Thornton, I. M., & Neville, H. J. (2003). Representation of Change: Separate Electrophysiological Markers of Attention, Awareness, and Implicit Processing. *Journal of Cognitive Neuroscience*, *15*(4), 491–507.

Fernandez-Duque, D. & Thornton, I. M. (2000). Change Detection Without Awareness: Do Explicit Reports Underestimate the Representation of Change in the Visual System? *Visual Cognition*, *7*(1-3), 323–344.

Fernandez-Duque, D. & Thornton, I. M. (2003). Explicit mechanisms do not account for implicit localization and identification of change: An empirical reply to Mitroff et al. (2002). *Journal of Experimental Psychology: Human Perception and Performance*, *29*(5), 846–858.

Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics*. SAGE Publications.

Friedman, A. (1979). Framing pictures: The role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General*, *108*(3), 316–355.

Friston, K., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *NeuroImage*, *19*(4), 1273–1302.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-P., Frith, C. D., & Frackowiak, R. S. (1994). Statistical parametric maps in functional imaging: a general linear approach. *Human brain mapping*, *2*(4), 189–210.

Fuchs, M., Wagner, M., & Kastner, J. (2007). Development of volume conductor and source models to localize epileptic foci. *Journal of Clinical Neurophysiology*, *24*(2), 101–119.

Galpin, A., Underwood, G., & Chapman, P. (2008). Sensing without seeing in comparative visual search. *Consciousness and Cognition*, *17*(3), 672–687.

Garreffa, G., Bianciardi, M., Hagberg, G. E., Macaluso, E., Marciani, M. G., Maraviglia, B., Abbafati, M., Carnì, M., Bruni, I., & Bianchi, L. (2004). Simultaneous EEG–fMRI acquisition: how far is it from being a standardized technique? *Magnetic Resonance Imaging*, *22*(10), 1445–1455.

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (2018). The Error-Related Negativity. *Perspectives on Psychological Science*, *13*(2), 200–204.

Gibson, J. J., Kaplan, G. A., Reynolds, H. N., & Wheeler, K. (1969). The change from visible to invisible:. *Perception & Psychophysics*, *5*(2), 113–116.

Goldman, Stern, J. M, Engel, J., & Cohen, M. S. (2000). Acquiring simultaneous EEG and functional MRI. *Clinical Neurophysiology*, *111(11)*, 1974–1980.

Goldman, R. I., Wei, C.-Y., Philiastides, M. G., Gerson, A. D., Friedman, D., Brown, T. R., & Sajda, P. (2009). Single-trial discrimination for integrating simultaneous EEG and fMRI: Identifying cortical areas contributing to trial-to-trial variability in the auditory oddball task. *NeuroImage*, *47*(1), 136–147.

Haberman, J. & Whitney, D. (2011). Efficient summary statistical representation when change localization fails. *Psychonomic Bulletin & Review*, *18*(5), 855.

Hanslmayr, S., Aslan, A., Staudigl, T., Klimesch, W., Herrmann, C. S., & Bäuml, K.-H. (2007). Prestimulus oscillations predict visual perception performance between and within subjects. *NeuroImage*, *37*(4), 1465–1473.

Haxby, J. V. (2001). Distributed and Overlapping Representations of Faces and Objects in Ventral Temporal Cortex. *Science*, *293*(5539), 2425–2430.

Haxby, J. V., Connolly, A. C., & Guntupalli, J. S. (2014). Decoding Neural Representational Spaces Using Multivariate Pattern Analysis. *Annual Review of Neuroscience*, *37*(1), 435–456.

Hillman, E. M. (2014). Coupling Mechanism and Significance of the BOLD Signal: A Status Report. *Annual Review of Neuroscience*, *37*(1), 161–181.

Hollingworth, A., Williams, C. C., & Henderson, J. M. (2001). To see and remember: Visually specific information is retained in memory from previously attended objects in natural scenes. *Psychonomic Bulletin & Review*, *8*(4), 761–768.

Howe, P. D. L. & Webb, M. E. (2014). Detecting Unidentified Changes. *PLoS ONE*, *9*(1), e84490.

Huettel, S. A., Song, A. W., & McCarthy, G. (2004). *Functional magnetic resonance imaging*, volume 1. Sinauer Associates Sunderland.

Huster, R. J., Debener, S., Eichele, T., & Herrmann, C. S. (2012). Methods for Simultaneous EEG-fMRI: An Introductory Review. *Journal of Neuroscience*, *32*(18), 6053–6060.

Iannaccone, R., Hauser, T. U., Staempfli, P., Walitza, S., Brandeis, D., & Brem, S. (2015). Conflict monitoring and error processing: New insights from simultaneous EEG–fMRI. *NeuroImage*, *105*, 395–407.

Irwin, D. E. (1991). Information integration across saccadic eye movements. *Cognitive Psychology*, *23*(3), 420–456.

Jarbo, K. & Verstynen, T. D. (2015). Converging Structural and Functional Connectivity of Orbitofrontal, Dorsolateral Prefrontal, and Posterior Parietal Cortex in the Human Striatum. *Journal of Neuroscience*, *35*(9), 3865–3878.

Jorge, J., van der Zwaag, W., & Figueiredo, P. (2014). EEG–fMRI integration for the study of human brain function. *NeuroImage*, *102*, 24–34.

Kahneman, D. (1984). Changing views of attention and automacy. *Varieties of attention*, 29–61.

Kanai, R. & Verstraten, F. A. (2004). Visual transients without feature changes are sufficient for the percept of a change. *Vision Research*, *44*(19), 2233–2240.

Kanwisher, N. (2001). Neural events and perceptual awareness. *Cognition*, *79*(1-2), 89–113.

Karnath, H.-O. & Rorden, C. (2012). The anatomy of spatial neglect. *Neuropsychologia*, *50*(6), 1010–1017.

Kastner, S. (1998). Mechanisms of Directed Attention in the Human Extrastriate Cortex as Revealed by Functional MRI. *Science*, *282*(5386), 108–111.

Kim, J., Wasserman, E. A., Castro, L., & Freeman, J. H. (2016). Anterior cingulate cortex inactivation impairs rodent visual selective attention and prospective memory. *Behavioral Neuroscience*, *130*(1), 75–90.

Kimura, M., Katayama, J., & Ohira, H. (2008). Event-related brain potential evidence for implicit change detection: A replication of Fernandez-Duque et al. (2003). *Neuroscience Letters*, *448*(3), 236–239.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, *36(14), 1.*

Koivisto, M., Kainulainen, P., & Revonsuo, A. (2009). The relationship between awareness and attention: Evidence from ERP responses. *Neuropsychologia*, *47*(13), 2891–2899.

Koivisto, M., Lähteenmäki, M., Sørensen, T. A., Vangkilde, S., Overgaard, M., & Revonsuo, A. (2008). The earliest electrophysiological correlate of visual awareness? *Brain and Cognition*, *66*(1), 91–103.

Koivisto, M. & Revonsuo, A. (2003). An ERP study of change detection, change blindness, and visual awareness. *Psychophysiology*, *40*(3), 423–429.

Koivisto, M. & Revonsuo, A. (2010). Event-related brain potential correlates of visual awareness. *Neuroscience & Biobehavioral Reviews*, *34*(6), 922–934.

Kolb, F. C. & Braun, J. (1995). Blindsight in normal observers. *Nature*, *377*(6547), 336–338.

Kraskov, A., Quiroga, R. Q., Reddy, L., Fried, I., & Koch, C. (2007). Local field potentials and spikes in the human medial temporal lobe are selective to image category. *Journal of cognitive neuroscience*, *19*(3), 479–492.

Kruggel, F., Wiggins, C. J., Herrmann, C. S., & von Cramon, D. Y. (2000). Recording of the event-related potentials during functional MRI at 3.0 Tesla field strength. *Magnetic resonance in medicine*, *44*(2), 277–282.

Kuleshov, L. V. & Kuleshov, L. (1974). *Kuleshov on Film: Writings*. University of California Press. Google-Books-ID: 0kfHxRVupxIC.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*.

Laloyaux, C., Destrebecqz, A., & Cleeremans, A. (2006). Implicit change identification: a replication of Fernandez-Duque and Thornton (2003). *Journal of Experimental Psychology: Human Perception and Performance*, *32*(6), 1366.

Laloyaux, C., Devue, C., Doyen, S., David, E., & Cleeremans, A. (2008). Undetected changes in visible stimuli influence subsequent decisions. *Consciousness and Cognition*, *17*(3), 646–656.

Lamme, V. A. (2003). Why visual attention and awareness are different. *Trends in cognitive sciences*, *7*(1), 12–18.

Lamme, V. A. F. (2004). Separate neural definitions of visual consciousness and visual attention; a case for phenomenal awareness. *Neural Networks*, *17*(5), 861–872.

Lantz, G., de Peralta Menendez, R. G., Andino, S. G., & Michel, C. M. (2001). Noninvasive localization of electromagnetic epileptic activity. II. Demonstration of sublobar accuracy in patients with simultaneous surface and depth recordings. *Brain topography*, *14*(2), 139–147.

Laufs, H. (2012). A personalized history of EEG–fMRI integration. *NeuroImage*, *62*(2), 1056–1067.

Laufs, H., Daunizeau, J., Carmichael, D., & Kleinschmidt, A. (2008). Recent advances in recording electrophysiological data simultaneously with magnetic resonance imaging. *NeuroImage*, *40*(2), 515–528.

Levin, D. T., Momen, N., Drivdahl, S. B., & Simons, D. J. (2000). Change Blindness Blindness: The Metacognitive Error of Overestimating Change-detection Ability. *Visual Cognition*, *7*(1-3), 397–412.

Levin, D. T. & Simons, D. J. (1997). Failure to detect changes to attended objects in motion pictures. *Psychonomic Bulletin & Review*, *4*(4), 501–506.

Liu, Y., Bengson, J., Huang, H., Mangun, G. R., & Ding, M. (2014). Top-down Modulation of Neural Activity in Anticipatory Visual Attention: Control Mechanisms Revealed by Simultaneous EEG-fMRI. *Cerebral Cortex*.

Liu, Y., Bengson, J., Huang, H., Mangun, G. R., & Ding, M. (2016). Top-down Modulation of Neural Activity in Anticipatory Visual Attention: Control Mechanisms Revealed by Simultaneous EEG-fMRI. *Cerebral Cortex*, *26*(2), 517–529.

Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature*, *453*(7197), 869–878.

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, *412*(6843), 150–157.

Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique*. MIT Press.

Luck, S. J. & Ford, M. A. (1998). On the role of selective attention in visual perception. *Proceedings of the National Academy of Sciences*, *95*(3), 825–830.

Luck, S. J. & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't): How to get significant effects. *Psychophysiology*, *54*(1), 146–157.

Luck, S. J., Heinze, H. J., Mangun, G. R., & Hillyard, S. A. (1990). Visual event-related potentials index focused attention within bilateral stimulus

arrays. II. Functional dissociation of P1 and N1 components. *Electroencephalography and Clinical Neurophysiology*, *75*(6), 528–542.

Luck, S. J. & Hillyard, S. A. (1994). Spatial filtering during visual search: Evidence from human electrophysiology. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(5), 1000–1014.

Luck, S. J. & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–281.

Luck, S. J. & Vogel, E. K. (2013). Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences*, *17*(8), 391–400.

Lyyra, P., Wikgren, J., Ruusuvirta, T., & Astikainen, P. (2012). Explicit behavioral detection of visual changes develops without their implicit neurophysiological detectability. *Frontiers in Human Neuroscience*, *6*.

Martınez, A., Teder-Salejarvi, W., Vazquez, M., Molholm, S., Foxe, J. J., Javitt, D. C., Russo, F. D., Worden, M. S., & Hillyard, S. A. (2006). Objects Are Highlighted by Spatial Attention. *18*(2), 13.

Mathewson, K. E., Gratton, G., Fabiani, M., Beck, D. M., & Ro, T. (2009). To See or Not to See: Prestimulus Phase Predicts Visual Awareness. *Journal of Neuroscience*, *29*(9), 2725–2732.

Mayhew, S. D., Dirckx, S. G., Niazy, R. K., Iannetti, G. D., & Wise, R. G. (2010). EEG signatures of auditory activity correlate with simultaneously recorded fMRI responses in humans. *NeuroImage*, *49*(1), 849–864.

Menon, V. & Uddin, L. Q. (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Structure and Function*, *214*(5-6), 655–667.

Michel, C. M., Lantz, G., Spinelli, L., de Peralta, R. G., Landis, T., & Seeck,

M. (2004). 128-Channel EEG Source Imaging in Epilepsy: Clinical Yield and Localization Precision:. *Journal of Clinical Neurophysiology*, *21*(2), 71–83.

Michel, C. M., Murray, M. M., Lantz, G., Gonzalez, S., Spinelli, L., & Grave de Peralta, R. (2004). EEG source imaging. *Clinical Neurophysiology*, *115*(10), 2195–2222.

Mitchell, D. J. & Cusack, R. (2008). Flexible, Capacity-Limited Activity of Posterior Parietal Cortex in Perceptual as well as Visual Short-Term Memory Tasks. *Cerebral Cortex*, *18*(8), 1788–1798.

Mitroff, S. R. & Simons, D. J. (2002). Changes are not localized before they are explicitly detected. *Visual Cognition*, *9*(8), 937–968.

Mitroff, S. R., Simons, D. J., & Franconeri, S. L. (2002). The siren song of implicit change detection. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(4), 798–815.

Mitroff, S. R., Simons, D. J., & Levin, D. T. (2004). Nothing compares 2 views: Change blindness can occur despite preserved access to the changed information. *Perception & Psychophysics*, *66*(8), 1268–1281.

Moosmann, M., Eichele, T., Nordby, H., Hugdahl, K., & Calhoun, V. D. (2008). Joint independent component analysis for simultaneous EEG–fMRI: Principle and simulation. *International Journal of Psychophysiology*, *67*(3), 212–221.

Mulert, C., Jäger, L., Schmitt, R., Bussfeld, P., Pogarell, O., Möller, H.-J., Juckel, G., & Hegerl, U. (2004). Integration of fMRI and simultaneous EEG: towards a comprehensive understanding of localization and time-course of brain activity in target detection. *NeuroImage*, *22*(1), 83–94.

Mulert, C. & Lemieux, L. (Eds.). (2010). *EEG - fMRI*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Mulert, C., Seifert, C., Leicht, G., Kirsch, V., Ertl, M., Karch, S., Moosmann, M., Lutz, J., Möller, H.-J., Hegerl, U., Pogarell, O., & Jäger, L. (2008). Single-trial coupling of EEG and fMRI reveals the involvement of early anterior cingulate cortex activation in effortful decision making. *NeuroImage*, *42*(1), 158–168.

Mullinger, K., Brookes, M., Stevenson, C., Morgan, P., & Bowtell, R. (2008). Exploring the feasibility of simultaneous electroencephalography/functional magnetic resonance imaging at 7 T. *Magnetic Resonance Imaging*, *26*(7), 968–977.

Murakami, S. & Okada, Y. (2006). Contributions of principal neocortical neurons to magnetoencephalography and electroencephalography signals: MEG/EEG signals of neocortical neurons. *The Journal of Physiology*, *575*(3), 925–936.

Mégevand, P., Quairiaux, C., Lascano, A. M., Kiss, J. Z., & Michel, C. M. (2008). A mouse model for studying large-scale neuronal networks using EEG mapping techniques. *NeuroImage*, *42*(2), 591–602.

Nguyen, V. T., Breakspear, M., & Cunnington, R. (2014). Fusing concurrent EEG–fMRI with dynamic causal modeling: Application to effective connectivity during face perception. *NeuroImage*, *102*, 60–70.

Niedeggen, M., Wichmann, P., & Stoerig, P. (2001). Change blindness and time to consciousness: Change blindness and time to consciousness. *European Journal of Neuroscience*, *14*(10), 1719–1726.

Niedermeyer, E. & Silva, F. H. L. d. (2005). *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Lippincott Williams & Wilkins. Google-Books-ID: tndqYGPHQdEC.

Noe, A., Pessoa, L., & Thompson, E. (2000). Beyond the Grand Illusion: What Change Blindness Really Teaches Us About Vision. *Visual Cognition*, *7*(1-3), 93–106.

Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, *10*(9), 424–430.

Novitski, N., Alho, K., Korzyukov, O., Carlson, S., Martinkauppi, S., Escera, C., Rinne, T., Aronen, H., & Näätänen, R. (2001). Effects of Acoustic Gradient Noise from Functional Magnetic Resonance Imaging on Auditory Processing as Reflected by Event-Related Brain Potentials. *NeuroImage*, *14*(1), 244–251.

Noë, A. (2005). What does change blindness teach us about consciousness? *Trends in Cognitive Sciences*, *9*(5), 218.

Nunez, P. L. & Silberstein, R. B. (2000). On the relationship of synaptic activity to macroscopic measurements: does co-registration of EEG with fMRI make sense? *Brain topography*, *13*(2), 79–96.

Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences*, *87*(24), 9868–9872.

O'Regan, J. K. (1992). Solving the "real" mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *46*(3), 461–488.

O'Regan, J. K., Deubel, H., Clark, J. J., & Rensink, R. A. (2000). Picture Changes During Blinks: Looking Without Seeing and Seeing Without Looking. *Visual Cognition*, *7*(1-3), 191–211.

O'Regan, J. K. & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and brain sciences*, *24*(05), 939–973.

Orr, J. M. & Weissman, D. H. (2009). Anterior Cingulate Cortex Makes 2 Contributions to Minimizing Distraction. *Cerebral Cortex*, *19*(3), 703–711.

Ostwald, D. & Bagshaw, A. P. (2011). Information theoretic approaches to functional neuroimaging. *Magnetic Resonance Imaging*, *29*(10), 1417–1428.

Overgaard, M., Jensen, M., & Sandberg, K. (2010). Methodological Pitfalls in the "Objective" Approach to Consciousness: Comments on Busch et al. (2009). *Journal of Cognitive Neuroscience*, *22*(9), 1901–1902.

Patel, A. B., de Graaf, R. A., Mason, G. F., Kanamatsu, T., Rothman, D. L., Shulman, R. G., & Behar, K. L. (2004). Glutamatergic Neurotransmission and Neuronal Glucose Oxidation Are Coupled During Intense Neuronal Activation:. *Journal of Cerebral Blood Flow & Metabolism*, 972–985.

Pertzov, Y., Heider, M., Liang, Y., & Husain, M. (2015). Effects of healthy ageing on precision and binding of object location in visual short term memory. *Psychology and Aging*, *30*(1), 26–35.

Pessoa, L. (2004). Neural Correlates of Change Detection and Change Blindness in a Working Memory Task. *Cerebral Cortex*, *14*(5), 511–520.

Pins, D. (2003). The Neural Correlates of Conscious Vision. *Cerebral Cortex*, *13*(5), 461–474.

Pourtois, G., De Pretto, M., Hauert, C.-A., & Vuilleumier, P. (2006). Time Course of Brain Activity during Change Blindness and Change Awareness: Performance is Predicted by Neural Events before Change Onset. *Journal of Cognitive Neuroscience*, *18*(12), 2108–2129.

Prince, S. & Hensley, W. E. (1992). The Kuleshov Effect: Recreating the Classic Experiment. *Cinema Journal*, *31*(2), 59–75.

Raichle, M. E. & Mintun, M. A. (2006). Brain work and brain imaging. *Annu. Rev. Neurosci.*, *29*, 449–476.

Railo, H., Koivisto, M., & Revonsuo, A. (2011). Tracking the processes behind conscious perception: A review of event-related potential correlates of visual consciousness. *Consciousness and Cognition*, *20*(3), 972–983.

Railo, H., Piccin, R., & Lukasik, K. M. (2020). Unconscious perception can be predicted from prestimulus activity. *bioRxiv*, 2020.01.06.896803.

Rensink, R. A. (2004). Visual sensing without seeing. *Psychological Science*, *15*(1), 27–32.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To See or not to See: The Need for Attention to Perceive Changes in Scenes. *Psychological Science*, *8*(5), 368–373.

Ress, D., Backus, B. T., & Heeger, D. J. (2000). Activity in primary visual cortex predicts performance in a visual detection task. *Nature Neuroscience*, *3*(9), 940–945.

Reynolds, M. G. & Withers, K. (2015). Understanding the relationship between implicit and explicit change detection: Evidence from scan path data. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(5), 1184–1189.

Ritter, P. & Villringer, A. (2006). Simultaneous EEG–fMRI. *Neuroscience & Biobehavioral Reviews*, *30*(6), 823–838.

Romei, V., Gross, J., & Thut, G. (2010). On the Role of Prestimulus Alpha Rhythms over Occipito-Parietal Areas in Visual Input Regulation: Correlation or Causation? *Journal of Neuroscience*, *30*(25), 8692–8697.

Rosa, M. J., Daunizeau, J., & Friston, K. J. (2010). EEG-fMRI integration: a critical review of biophysical modelling and data analysis approaches. *Journal of Integrative Neuroscience*, *09*(04), 453–476.

Rowlands, A. (2000). *The Continuity Supervisor*. Taylor & Francis. Google-Books-ID: JbLkr2d9Ei0C.

Roy, C. S. & Sherrington, C. S. (1890). On the regulation of the blood-supply of the brain. *The Journal of physiology*, *11*(1-2), 85–158.

Sammer, G., Blecker, C., Gebhardt, H., Kirsch, P., Stark, R., & Vaitl, D. (2005). Acquisition of typical EEG waveforms during fMRI: SSVEP, LRP, and frontal theta. *NeuroImage*, *24*(4), 1012–1024.

Schankin, A. & Wascher, E. (2007). Electrophysiological correlates of stimulus processing in change blindness. *Experimental Brain Research*, *183*(1), 95–105.

Scheeringa, R., Bastiaansen, M. C., Petersson, K. M., Oostenveld, R., Norris, D. G., & Hagoort, P. (2008). Frontal theta EEG activity correlates negatively with the default mode network in resting state. *International Journal of Psychophysiology*, *67*(3), 242–251.

Scrivener, C. L., Malik, A., Marsh, J., Lindner, M., & Roesch, E. (2019). An EEG study of detection without localisation in change blindness. *bioRxiv*, 513697.

Seeber, M., Cantonas, L.-M., Hoevels, M., Sesia, T., Visser-Vandewalle, V., & Michel, C. M. (2019). Subcortical electrophysiological activity is detectable with high-density EEG source imaging. *Nature Communications*, *10*(1), 753.

Shallice, T. (1964). The Detection of Change and the Perceptual Moment Hypothesis. *British Journal of Statistical Psychology*, *17*(2), 113–135.

Simons, D. & Levin, D. T. (1997). Change Blindness. *Trends in Cognitive Sciences*, 261–267.

Simons, D. J. (2000). Current Approaches to Change Blindness. *Visual Cognition*, *7*(1-3), 1–15.

Simons, D. J. & Ambinder, M. S. (2005). Change Blindness: Theory and Consequences. *Current Directions in Psychological Science*, *14*(1), 44–48.

Simons, D. J. & Chabris, C. F. (1999). Gorillas in Our Midst: Sustained Inattentional Blindness for Dynamic Events. *Perception*, *28*(9), 1059–1074.

Simons, D. J. & Levin, D. T. (2003). What Makes Change Blindness Interesting? In *Psychology of Learning and Motivation*, volume 42 of *Cognitive Vision* (pp. 295–322). Academic Press.

Simons, D. J., Nevarez, G., & Boot, W. R. (2005a). Visual Sensing Is Seeing: Why "Mindsight," in Hindsight, Is Blind. *Psychological Science*, *16*(7), 520–524.

Simons, D. J., Nevarez, G., & Boot, W. R. (2005b). Visual Sensing Is Seeing Why "Mindsight," in Hindsight, Is Blind. *Psychological Science*, *16*(7), 520–524.

Smilek, D., Eastwood, J. D., & Merikle, P. M. (2000). Does unattended information facilitate change detection? *Journal of Experimental Psychology: Human Perception and Performance*, *26*(2), 480–487.

Snyder, A. Z. & Raichle, M. E. (2010). Studies of the human brain combining functional neuroimaging and electrophysiological methods. *Simultaneous EEG and fMRI: Recording, Analysis, and Application*, 47–66.

Sommer, M., Meinhardt, J., & Volz, H. (2002). Combined measurement of event-related potentials (ERPs) and fMRI. *Acta neurobiologiae experimentalis*, *63*(1), 49–54.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, *74*(11), 1–29.

Srivastava, G., Crottaz-Herbette, S., Lau, K., Glover, G., & Menon, V. (2005). ICA-based procedures for removing ballistocardiogram artifacts from EEG data acquired in the MRI scanner. *NeuroImage*, *24*(1), 50–60.

Stanislaw, H. & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137–149.

Stokes, M. (2015). 'Activity-silent' working memory in prefrontal cortex: a dynamic coding framework | Elsevier Enhanced Reader.

Strayer, D. L. & Johnston, W. A. (2001). Driven to Distraction: Dual-Task Studies of Simulated Driving and Conversing on a Cellular Telephone. *Psychological Science*, *12*(6), 462–466.

Stroud, J. M. (1967). The fine structure of psychological time. *Annals of the New York Academy of Sciences*, *138*(2 Interdiscipli), 623–631.

Tenke, C. E. & Kayser, J. (2012). Generator localization by current source density (CSD): Implications of volume conduction and field closure at intracranial and scalp resolutions. *Clinical Neurophysiology*, *123*(12), 2328–2345.

Thornton, I. & Fernandez-Duque, D. (2001). An implicit measure of undetected change. *Spatial Vision*, *14*(1), 21–44.

Tseng, P., Hsu, T.-Y., Chang, C.-F., Tzeng, O. J. L., Hung, D. L., Muggleton, N. G., Walsh, V., Liang, W.-K., Cheng, S.-k., & Juan, C.-H. (2012). Unleashing Potential: Transcranial Direct Current Stimulation over the Right Posterior Parietal Cortex Improves Change Detection in Low-Performing Individuals. *Journal of Neuroscience*, *32*(31), 10554–10561.

Turatto, M. (2002). Looking without seeing the background change: electrophysiological correlates of change detection versus change blindness. *Cognition*, *84*(1), B1–B10.

Uddin, L. Q., Nomi, J. S., Hébert-Seropian, B., Ghaziri, J., & Boucher, O. (2017). Structure and Function of the Human Insula:. *Journal of Clinical Neurophysiology*, *34*(4), 300–306.

Ullsperger, M. & Debener, S. (Eds.). (2010). *Simultaneous EEG and fMRI: recording, analysis, and application*. New York: Oxford University Press.

Ungerleider, S. K. a. L. G. (2000). Mechanisms of Visual Attention in the Human Cortex. *Annual Review of Neuroscience*, *23*(1), 315–341.

van Belle, J., Vink, M., Durston, S., & Zandbelt, B. B. (2014). Common and unique neural networks for proactive and reactive response inhibition revealed by independent component analysis of functional MRI data. *NeuroImage*, *103*, 65–74.

Varakin, D. A., Levin, D. T., & Collins, K. M. (2007). Comparison and Representation Failures Both Cause Real-World Change Blindness. *Perception*, *36*(5), 737–749.

Vogel, E. K. & Luck, S. J. (2000). The visual N1 component as an index of a discrimination process. *Psychophysiology*, *37*(2), 190–203.

Vogel, E. K. & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, *428*(6984), 748–751.

Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature*, *438*(7067), 500–503.

Wallach, H. & Lewis, C. (1966). The effect of abnormal displacement of the retinal image during eye movements. *Perception & Psychophysics*, *1*(1), 25–29.

Walz, J. M., Goldman, R. I., Carapezza, M., Muraskin, J., Brown, T. R., & Sajda, P. (2015). Prestimulus EEG alpha oscillations modulate task-related fMRI BOLD responses to auditory stimuli. *NeuroImage*, *113*, 153–163.

Warbrick, T. & Bagshaw, A. P. (2008). Scanning strategies for simultaneous EEG–fMRI evoked potential studies at 3 T. *International Journal of Psychophysiology*, *67*(3), 169–177.

Watanabe, K. (2003). Differential effect of distractor timing on localizing versus identifying visual changes. *Cognition*, *88*(2), 243–257.

Weiskrantz, L. (1986). *Blindsight: A Case Study and Implications*. Oxford University Press.

Wilenius, M. E. & Revonsuo, A. T. (2007). Timing of the earliest ERP correlate of visual awareness. *Psychophysiology*, *44*(5), 703–710.

Williams, P. & Simons, D. J. (2000). Detecting Changes in Novel, Complex Three-dimensional Objects. *Visual Cognition*, *7*(1-3), 297–322.

Wirsich, J., Bénar, C., Ranjeva, J.-P., Descoins, M., Soulier, E., Le Troter, A., Confort-Gouny, S., Liégeois-Chauvel, C., & Guye, M. (2014). Single-trial EEG-informed fMRI reveals spatial dependency of BOLD signal on early and late IC-ERP amplitudes during face recognition. *NeuroImage*, *100*, 325–336.

Woodman, G. F. & Luck, S. J. (2003). Dissociations among attention, perception, and awareness during object-substitution masking. *Psychological Science*, *14*(6), 605–611.