Multidimensional resource allocation: The role of attentional control in the allocation of visual short-term memory resources

Stuart Bryan Moore

A thesis submitted for the degree of Doctor of Philosophy
Keele University
June 2023

Abstract

Previous research has highlighted the importance of attention in the allocation of visual short-term memory (vSTM) resources. However, much of this research examined performance across a single feature dimension (e.g., colour), disregarding the potential impact of attending to multivalent items which are more typically found in everyday life. Across a series of experiments, vSTM performance in situations wherein the relevant feature dimension either repeated or switched was investigated to assess the role of attentional control in vSTM resource allocation. In Chapter 2, change detection tasks were used, with dimension switches causing an increase in errors and response time relative to dimension repetitions, as well as impacting measures of sensitivity and response bias. This was supplemented by further change detection experiments within Chapter 3 which showed that the extent of the dimension switch cost is not influenced by independent manipulation of cue-stimulus or response-cue intervals. Furthermore, cue switch costs do not contribute to the overall dimension switch cost, suggesting an origin of error distinct to that associated with task switching. Chapter 4 provided a more nuanced insight into the cause of dimension switch costs using a continuous report task alongside mixture modelling. Overall, there was little evidence for dimension switch costs in both behavioural measures and model parameters. In Chapter 5, change localisation tasks were used and six models were generated to provide further insight into the cause of dimension switch costs. Results from these experiments revealed no evidence in support of a dimension switch cost, with formal model comparison revealing that a model wherein all parameters were fixed across sequencing condition was favoured. Taken together, the results of the present research are somewhat inconclusive. While there does appear to be some role of attentional control in vSTM resource allocation, it appears that this is only captured using specific methodologies.

Keywords: visual short-term memory, attentional control, resource allocation

Word count: 60,582

Contents

Acknowledgements

First and foremost, I would like to express sincere gratitude to my primary supervisor Dr Jim Grange for his invaluable support and guidance throughout my PhD. There were occasions during my PhD where I felt somewhat lost or out of my depth, but thanks to Jim, these remained only brief instances of doubt. From the beginning, Jim was always willing to meet and discuss research or help deal with any issues I may have been facing, offering support and encouragement where needed, and for that, I am extremely grateful. I will miss the virtual whiteboard! I would also like to acknowledge Jim's role in helping me to develop as a better researcher in general, from inviting me to collaborate on a paper, introducing me to more Open Research practices, and perhaps most of all, for his help and guidance in learning R. Given I had no previous experience in R before beginning my PhD, it is safe to say that this thesis would not have been written in R without Jim's knowledge and support.

Secondly, I would like to thank my second supervisor Dr Sara Spotorno for her insight, encouragement, and willingness to help regardless of the situation. Unfortunately, the Coronavirus pandemic meant that we were unable to carry out any of the eye-tracking research we had initially planned, but Sara was always willing to assist in whatever way needed, be it helping with participant recruitment or providing invaluable feedback on drafts of my work. For this, I am extremely grateful. Special mention also to former MSc students and research assistants Megan Owen, Natalie Keel, and Peirce Proctor who helped to collect data for the studies within Chapter 3.

I would also like to acknowledge the academic and support staff within the School of Psychology for their support, guidance, and friendship during my PhD. The opportunity to sit on various committees and take on different roles within the School of Psychology has allowed me to develop skills and gain an insight related to aspects of academia I would not have experienced otherwise. For this, I would particularly like to thank Dr Joe Brooks, Professor Nicky Edelstyn, Dr Nick Garnett, Dr Masi Noor, Dr Sue Sherman, and Dr Sara Vestergren.

Getting to this point would have been much more difficult without being surrounded by the people I am fortunate to call friends. In no particular order, I'd like to thank Iwan and Joe for our frequent forays around Newcastle and their unfaltering ability to listen to me complain. Thanks also to Anna, Charlotte, Heather, and Honor for your support (and weekly Zoom quizzes during the initial stages of lockdown!), and thanks to Mitch who was always on hand to lighten the mood. To my fellow PhD candidates at Keele, many thanks for being so welcoming, taking an interest in my research, and offering advice and support.

There are also several people from the University of Central Lancashire (and beyond) I would like to thank. In particular, my MSc supervisory team, Dr John Marsh, Professor Linden Ball, and Dr Reyhan Furman for helping me to develop as a researcher, and for their continued support and encouragement. Thanks also to Dr Beth Richardson, Professor Simon Liversedge, and members of the Perception, Cognition, and Neuroscience Research Group who always make me feel welcome during my trips to Preston, and to Dr Andy Morely and Dr Paul Taylor for their support and banter over the years. Special thanks to my former UCLan office mate Zoe Hughes for always making time to catch up over coffee during my visits. Finally, many thanks to Dr Oliver Kannape, with whom my journey in research began as an undergraduate.

Last but most certainly not least, I would like to thank my family for their support not only during my PhD, but throughout the course of my studies. In particular, my mother Jacqueline and my father Alan for their encouragement and help (quite often financial!) throughout my time at university. Without this, I would not be where I am today so thank you. Also, many thanks to my siblings Craig and Victoria, my siblings-in-law Lesley and Marty, and my niece Emily and nephews Noah and Rory.

Author contribution statement

Some of the research within the present thesis has been communicated externally (see references below), either by way of pre-print, oral presentation, or poster presentation. Stuart B. Moore (SBM; author of the present thesis) conceptualised and planned all the research with support from the primary supervisor Dr James A. Grange (JAG). Data collection was primarily conducted by SBM however, three research assistants, Megan Owen (MO), Natalie Keel (NK), and Peirce Proctor (PP), assisted with the collection of data in Chapter 3. SBM conducted all data analysis and modelling, and wrote the original drafts for each of the communications. All communications were co-authored with the primary supervisor JAG, with one also including the research assistants MO and PP as co-authors. All authors reviewed and edited the final draft for each communication.

Pre-print

Moore, S. B. & Grange, J. A. (2022). Change detection across dimensions: The impact of dimension switching on change detection performance [pre-print]. *PsyArXiv.* doi: 10.31234/osf.io/7rnz2

Oral presentation

Moore, S. B. & Grange, J. A. (March, 2022). *Dimension switching impairs visual short-term memory resource allocation.* Experimental Psychology Society Conference [Keele University]. Available at osf.io/ng76t/.

Poster presentations

Moore, S. B., Owen, M., Proctor, P., & Grange, J. A. (July, 2021). *Reconfiguration or inertia? Identifying a possible origin of switch costs during visual short-term memory resource allocation.* Poster presented at the Experimental Psychology Society Conference [presented online due to Covid-19]. Available at doi.org/10.13140/RG.2.2.19283.50729.

Moore, S. B. & Grange, J. A. (April, 2021). *The impact of dimension switching on visual short-term memory.* Poster presented at the Experimental Psychology Society Conference [presented online due to Covid-19]. Available at doi.org/10.13140/RG.2.2.28885.65761.

Moore, S. B. & Grange, J. A. (July 2020). *Multidimensional change detection: An initial investigation of resource allocation in visual short-term memory.* Poster presented at the Experimental Psychology Society Conference [presented online due to Covid-19]. Available at doi.org/10.17605/OSF.IO/3GYRF.

1 Chapter 1: Introduction and literature review

This chapter outlines the main research question, making reference to visual short-term memory, attentional control, and task switching, before providing a brief background on some of the key research in each of these areas.

1.1 Overview

Throughout daily life, we are bombarded with an abundance of visual sensory input which we must use not only to form an internal representation of our environment, but to also perform various behaviours and functions (see e.g., Berman & Colby, 2009; Mathôt, 2018; Roussy et al., 2021). Central to our ability to perform these behaviours and functions is visual short-term memory (vSTM), the cognitive system responsible for the retention of visual information over brief (on the order of seconds) periods of time. Unlike iconic memory—a large capacity pre-attentive store which holds visual *traces* of previously seen information for less than one second (see e.g., Coltheart, 1980; Rensink, 2014; Sperling, 1960; however, see also Persuh et al., 2012)—vSTM can store only a small amount of information at a given time and can survive saccadic eye movements and spatial shifts of attention (see e.g., Hollingworth et al., 2008). As such, vSTM acts as a bridge between perception and action, allowing for information that would otherwise be lost to be retained in service of an upcoming task or behaviour.

Despite a great deal of research throughout the years, the capacity-limited nature of vSTM still remains debated, with two views coming to the fore. On on hand, there is research which suggests the capacity limit of vSTM is due to the total number of items that can be represented in memory at a given time (see e.g., Luck & Vogel, 1997; Pashler, 1988;

Vogel et al., 2001; Zhang & Luck, 2008). Alternatively, other research suggests that the limitation is due to the allocation of a continuous, finite memory resource (see e.g., Bays et al., 2009; Bays & Husain, 2008; Frick, 1988; van den Berg et al., 2012; Wilken & Ma, 2004). This resource is responsible for the resolution at which items are represented at in memory, with an increase in the total number of items to be retained in memory reducing the fidelity of each representation due to less resource being available for allocation to each item. In relation to the allocation of this finite memory resource, while there are a substantial number of studies which have investigated such allocation (see e.g., Bays et al., 2009; Dube et al., 2017; Emrich et al., 2017; Henderson et al., 2020; Machizawa et al., 2012; Wilken & Ma, 2004), much of this research has examined performance on simple stimuli, defined only by a single *feature dimension* (typically colour). This is atypical of everyday life, wherein stimuli we encounter are often multivalent, containing a substantial number of different feature dimensions (e.g., road traffic signs) or indeed, multiple *feature values* of the same feature dimension (e.g., multicoloured stimuli). Often, only a subset of these features are relevant to current goals or behaviours; thus, if vSTM is limited in capacity, how does the cognitive system ensure that the resources available are allocated to the most relevant visual information?

Attention, or more specifically *attentional control* is likely to play a key role in such allocation. Specifically what attention is still remains ambiguous and clarification of this is outside the remit of this thesis; however Scholl (2001) provides a brief discussion on this matter, concluding with "…there are (possibly several different) types of selective processing—which will collectively be called 'attention'—that play a ubiquitous and important role in visual processing.” (p.5, see also Chun et al., 2011; Oberauer, 2019). As such, attentional control could be defined as a combination of both attentional and inhibitory processes, wherein attention is focused on a specific aspect of a visual scene and interference from irrelevant information is prevented (see e.g., Friedman & Miyake, 2004; Miyake et al., 2000; Yantis, 1998). The way in which the role of attentional control in resource allocation is investigated throughout this thesis is via a combination of tasks which probe vSTM (change detection, Chapters 2 & 3; continuous report, Chapter 4; change localisation, Chapter 5) and *task switching* paradigms. In typical task switching procedures, participants either repeat performance of the same task or switch between performance of two (or more) tasks. A consistent finding from the literature shows that both error rates and response times (RTs) are increased when the task switches relative to when the task repeats, with this detriment to performance referred to as the *task switch cost* (for comprehensive reviews on task switching, see Grange & Houghton, 2014; Kiesel et al., 2010; Monsell, 2003; Vandierendonck et al., 2010). Crucially however, switching requires some degree of attentional control (see e.g., Logan, 1980; D. W. Schneider, 2015).

While different tasks are used in traditional task switching, in the experiments detailed throughout this thesis, participants perform only a single task (e.g., change detection), with the switching element relating to which feature dimension (i.e., colour or orientation) is relevant on a given trial. On some trials, the relevant feature dimension could repeat (e.g., colour–colour), while on others it could switch (e.g., colour–orientation). As such, the cognitive system is required to update the processes and representations from that used on the previous trial when a dimension switch occurs, whereas no such updating would be required when the feature dimension repeats. This increased attentional control demand during dimension switch trials could lead to a reduction in performance relative to dimension repetition trials. The presence of such a cost could be used to further investigate the allocation of vSTM resources which could in turn help inform possible extensions to extant models of vSTM.

1.2 Visual short-term memory capacity

1.2.1 The change detection task

Most of the early research investigating vSTM capacity and the nature of representations in memory made use of the change detection paradigm (see e.g., Alvarez & Cavanagh, 2004; Luck & Vogel, 1997; Pashler, 1988; Phillips, 1974; Purdy et al., 1980; Vogel et al., 2001; Wheeler & Treisman, 2002; Wilken & Ma, 2004) wherein two stimulus displays are presented (e.g., coloured squares), separated by a brief retention interval; the goal of this task is to determine whether a change in some aspect of the stimulus display (e.g., the colour of a square) occurred between the initial *memory* display and subsequent *test* display (see Figure 1). For example, in perhaps one of the most influential papers on vSTM capacity, Luck and Vogel (1997) used change detection tasks in a series of experiments. In one such experiment, coloured squares were used as stimuli with varying set size; participants simply determined whether a change occurred in the colour of one of the squares between the memory and test displays.

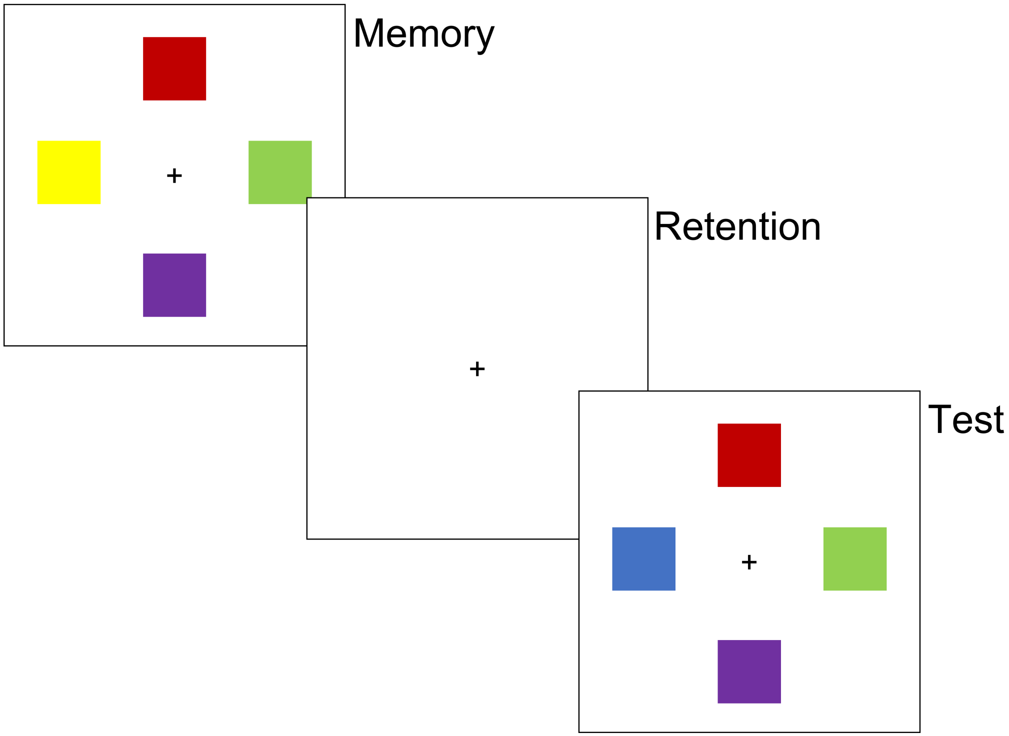

Figure 1

Figure showing an example trial procedure within a change detection task using coloured squares as stimuli.

Luck and Vogel used Pashler's (1988) K *measure* to estimate memory capacity, given as:

$$K = S\left(\frac{HR - FAR}{1 - FAR}\right), \tag{1}$$

where K represents capacity, S set size, HR hit rate, and FAR false alarm rate (see also Cowan et al., 2005; Rouder et al., 2011). Hit and false alarm rates can be obtained from change detection data (among other forms of data, e.g., recognition memory data); a *hit* refers to a change response when a change occurred (i.e., signal present), while a *miss* refers to a no change response when a change occurred. A *correct rejection* refers to a no change response when no change occurred, while a *false alarm* refers to a change response when no change occurred (i.e., signal absent). To obtain the *hit rate* simply divide the total number of hits by the total number of hit and miss trials. For the *false alarm rate* simply divide the total number of false alarms by the total number of correct rejections and false alarms.

Using this measure, Luck and Vogel (1997) found that participants were able to hold approximately four colours in memory. At set sizes of one to four, performance was not substantially different however, from set sizes of four and onward, systematic reductions to performance were observed. This led Luck and Vogel to conclude that vSTM capacity is limited in terms of the total number of items that can be represented in memory simultaneously, with this capacity being around four or five items. In subsequent experiments, Luck and Vogel compared performance on a change detection task using coloured, oriented bars; in one experiment, both the colour and orientation of the bars was variable (i.e., different colours and orientations were used) with three conditions, *colour only* wherein only colour could change, *orientation only* wherein only orientation could change, and *either colour or orientation* wherein either colour or orientation could change. Participants were informed of this at the beginning of each condition. Results showed that performance did not differ between conditions to any great extent, despite an increase in the number of features to be retained in memory in the either colour or orientation condition relative to the single feature conditions. A further experiment was conducted to determine whether feature values for both dimensions were encoded automatically when only a single feature dimension was relevant. In this experiment, the irrelevant feature dimension was held constant (i.e., all black bars if orientation was relevant, all vertical bars if colour was relevant) with results showing that performance between these two experiments did not differ to any great extent.

Furthering these findings, Luck and Vogel (1997) carried out two additional experiments, showing that performance on a change detection task wherein stimuli are defined by four feature dimensions (colour, orientation, size, and presence of a gap) was akin to performance in experiments with individual feature dimensions. This, the authors suggested, was indicative of the ability for vSTM to retain up to 16 features distributed across four objects with the same degree of accuracy as four features distributed across four objects. Finally, Luck and Vogel also ruled out the possibility of independent stores for different feature dimensions. This was achieved by comparing performance on a condition wherein stimuli were made up of different coloured squares (a smaller square overlaid on a larger square) and a condition wherein only one of these squares was presented; performance was found to be at a similar level across these conditions, suggesting eight colours distributed across four objects could be retained as well as four colours distributed across four objects. Luck and Vogel (1997) conclude that vSTM stores integrated objects rather than individual features (see also work by Irwin and colleagues, Irwin, 1991; Irwin, 1992; Irwin & Andrews, 1996), which inherently suggests that each object percept stored in memory is done so at a fixed resolution. Indeed, further support for this conclusion is provided by Vogel et al. (2001).

1.2.2 Objects vs. features

However, a number of studies have provided evidence which conflicts with the findings by Luck and Vogel (1997; see e.g., Horowitz and Wolfe, 1998; Olson and Jiang, 2002;

Rensink, 2000, 2001; Wheeler and Treisman, 2002; Wolfe, 1999; Xu, 2002). For instance, Wheeler and Treisman (2002) attempted to replicate a number of findings from Luck and Vogel (1997), with the aim of further investigating the notion that vSTM stores integrated objects rather than individual features. Change detection tasks similar to those used by Luck and Vogel (1997) were employed by Wheeler and Treisman (2002). In Experiment 1, seven sets of three bicoloured squares were used as stimuli, with a further two control sets made up of three and six squares, each with a single colour. Participants simply determined whether the test display was the same or different than the memory display. Comparison of accuracy data revealed no significant differences between any of the seven bicoloured stimulus sets and the control set consisting of six single-coloured stimuli, while performance on the control condition wherein three single-coloured stimuli were presented was significantly better. In contrast to the conclusions by Luck and Vogel (1997), Wheeler and Treisman (2002) state that memory in this task was limited by the total number of features presented; if stimuli were stored as integrated objects—wherein capacity increases due to *chunking* of feature values—performance in the bicoloured conditions should have been at a similar level to that in the three single-coloured condition.

Subsequent experiments by Wheeler and Treisman (2002; Experiments 3a and 4a) investigated binding memory for features from different dimensions. In Experiment 3a, participants were presented with either three or six coloured squares in different locations on the display; on trials wherein a change occurred between memory and test displays, in the *colour only* condition, two of the stimuli changed colour, in the *location only* condition, the location of two squares would change, in the *either location or colour* condition, on half the trials a colour change occurred, with location changing on the other half, and in the *binding location and colour* condition, the conjunction between colour and location would change on two squares (i.e., the colours of two squares would change to a different location). While some results follow along with that observed by Luck and Vogel (1997; e.g., set size mediated performance in the first three conditions), some results challenged the view of integrated object storage. While Luck and Vogel argued that the similarity between performance on conditions wherein a single feature could change and performance in a condition wherein either feature could change was indicative of integrated object storage, Wheeler and Treisman (2002) highlighted that no binding of feature dimensions was required in the either condition; participants could simply retain lists of the feature values presented in the memory display and compare the relevant list with the feature values presented at test. In the binding condition of Experiment 3a in Wheeler and Treisman (2002), results show that performance was significantly reduced compared to each of the other conditions, indicating that participants struggled to detect changes in the binding of feature dimensions. The authors go on to state that such a result follows with the notion of independent memory stores for different feature dimensions.

Indeed, there is a body of evidence which has highlighted that *information load* or *stimulus complexity* also has some influence over the amount of information that can be retained in vSTM (see e.g., Alvarez & Cavanagh, 2004; Awh et al., 2007; Eng et al., 2005; Fougnie & Alvarez, 2011). For example, using a change detection task, Alvarez and Cavanagh (2004) found that estimated capacity varied across the five sets of stimuli used (Snodgrass line drawings, shaded cubes, random polygons, Chinese characters, letters, and coloured squares). At the top-end of performance, participants were able to retain approximately four coloured squares in memory while only being able to retain approximately two shaded cubes in memory. Given that Luck and Vogel (1997) found that features did not appear to influence capacity, Alvarez and Cavanagh (2004) provided two possible reasons as to why results differed between the two studies. First, they highlighted the potential for independent stores for different feature dimensions, making reference to the finding from Wheeler and Treisman (2002) that multiple feature values from the same dimension on a single object cannot be retained in memory as well as a single feature value. Additionally, while not explicitly stated by Alvarez and Cavanagh (2004), it is possible that some of the stimuli required binding, which Wheeler and Treisman (2002) had previously shown to impair performance. In the case of the shaded cubes for example, these included a white side, a grey side, and a black side, with some viewed from the top and some from the bottom. In order to accurately determine whether a change occurred between the memory and test displays, it is plausible to assume that participants would have to bind the colour with the side of the cube it appeared on (i.e., white on top, grey on the left, black on the right); misremebering where a colour was located on the cube would invariably lead to a change response. The second explanation provided by Alvarez and Cavanagh (2004) is that specific features could be encoded automatically and independent of the demands of the task to form a minimal representation of the object in memory. Encoding of any additional features—such as when stimuli are more complex—then results in a reduction in the total number of objects that can be retained in memory (i.e., a trade-off between stimulus complexity and capacity).

While extensive discussion of object vs. feature storage is outside the remit of this thesis, it is important to highlight as the study by Luck and Vogel (1997) paved the way for the suggestion that vSTM is a *discrete* system, having a set number of *slots* in which to store information (see e.g., Zhang & Luck, 2008). However, an alternative view suggests that a *continuous, finite memory resource* is responsible for encoding items into memory, with the amount of resource allocated to each item indicative of how precisely this item is represented in memory (see e.g., Bays et al., 2009). The following section briefly outlines the development of these theories, including discussion of various measurement models which have been used to estimate the fidelity of representations in memory, as well as indicate sources of error.

1.2.3 The issue with high-threshold models

Despite supporting the conclusions of Luck and Vogel (1997) that vSTM capacity is limited to around four objects with each object stored in memory at a fixed resolution, Vogel et al. (2001) highlighted the possibility that low fidelity representations which exceed this upper limit could be generated. Such a possibility had been considered previously by Frick (1988), who stated that vSTM capacity, or more specifically, the capacity of the *visuospatial sketchpad* (see Baddeley & Hitch, 1974), was associated with a finite amount of *representational medium.* This medium is distributed to items in the immediate visual environment and is responsible for the fidelity of internal representations of these items. As such, when set size is low, high resolution representations can be generated as more of this medium is available to be distributed to each item. Alternatively, when set size is low, less medium is available to be distributed to each item, thus reducing the resolution at which each item is represented at in memory. This notion of vSTM capacity being limited by a *resource* rather than discrete units has led to several more recent instantiations of similar models (Bays, 2014; Bays, 2016; see e.g., Bays et al., 2009; Shin & Ma, 2017; van den Berg et al., 2012), development of which stem from seminal work by Wilken and Ma (2004). Initially, Wilken and Ma highlighted criticisms of the use of *high-threshold* (HT) models which some studies had used to estimate memory capacity (see e.g., Luck & Vogel, 1997 and their use of Pashler's, 1988 K measure), before investigating how well *signal detection theory* (or simply *detection theory*; SDT) accounts fit data obtained from change detection tasks, as well as examining the noise associated with encoding using *continuous report* tasks.

Pashler's (1988) K measure (see Eq. 1) is one example of an HT model owing to the fact it assumes that discrimination may occasionally fail to detect a target (i.e., a signal within noise), but that a non-target can never be mistaken for a target (i.e., noise can never be mistaken as a signal); in other words, the decision threshold is high and cannot be reached by noise alone. Thus, on this and other HT models, the false alarm rate in a task such as change detection simply reflects guessing (see e.g., Burmester & Wallis, 2012). Clearly such a model does not fit well within the context of the human brain which is inherently noisy and as such, Wilken and Ma (2004) set out to investigate how well SDT accounts fit the observed data. Across a series of experiments (1-6), the authors used a change detection task with manipulations of set size (Experiments 1-3) and target number (i.e., the total number of stimuli wherein a change can occur on a given trial; Experiments 4-6) across the feature spaces of colour, orientation, and spatial frequency. The models fit to the data include the SDT models *maximum absolute differences* (MAD) and *sum of absolute differences* (SAD), as well as an HT model (see Wilken & Ma, 2004 for full technical details of each model and the fitting procedures). Primarily, the difference between the MAD and

SAD models and HT models is the way in which items are represented internally; while HT models hold that items are represented in the absence of any internal noise, both the MAD and SAD models hold that each item is (independently) represented in a noisy internal state. Given that representations are held with noise, it is possible for this noise to exceed the threshold required for target (e.g., change) detection even when no target is present (e.g., a change does not occur), thus resulting in a false alarm. Results from the model fitting in relation to set size manipulations showed that both the MAD and SAD models better fit the data than the HT model, with the MAD model providing the best fit overall, with Wilken and Ma highlighting that as set size increased, so to did the noise. Similarly, both MAD and SAD models better fit the data in the target number manipulation relative to the HT model; however, Wilken and Ma state that this is not without some issue, highlighting that assumptions on both SDT accounts may be incorrect.

1.3 Modelling visual short-term memory

1.3.1 The continuous report task

More critically to the present thesis however, are the findings from Experiments 7–9 of Wilken and Ma (2004) which made use of a continuous report task (see Figure 2, see also Prinzmetal et al., 1998). In these experiments, stimuli consisted of coloured squares, Gabor patches, and vertical Gabor patches with added spatial frequencies; set size was varied depending on the stimulus type (colour and spatial frequency $N = 2$, 4, 6, or 8; orientation $N = 2$, 3, 4, or 5). After extinguishing of the memory display, a brief blank period was followed by presentation of a probe display indicating the location of one of the stimuli presented on the memory display. In the colour task, participants reported the colour of the square which appeared in the probed location by clicking on a 360° colour wheel presented in the centre of the display. In the orientation and spatial frequency experiments, participants manipulated a probe Gabor using the right (clockwise) and left (counter-clockwise) keys on the keyboard to match the orientation or spatial frequency of the stimulus which appeared in the probed location in the memory display, confirming their response by pressing spacebar. By calculating the deviation between response and the true value of the target, the authors were able to gain an estimate of how precisely the probed stimulus was represented in memory. The distribution of errors within these experiments revealed that response precision declined as a function of set size, that is to say, as set size increased responses became further from the true value of the target. Wilken and Ma take this as evidence against a fixed capacity vSTM, arguing that if capacity was limited in such a way, performance on this task should be essentially perfect until the upper limit was reached with little (or indeed no) increases in noise. Once maximum memory capacity was reached, noise would then abruptly increase as participants would be forced to guess on some trials. That no such pronounced increase in noise was observed by Wilken and Ma, they suggest that the supposed capacity limit of vSTM suggested by some research (see e.g., Luck & Vogel, 1997; Vogel et al., 2001) is simply artefactual, reflecting an increase in internal noise as set size increases. However, work by Zhang and Luck (2008)—which also employed use of the continuous report task—suggests that vSTM is best explained in terms of a discrete set of representations which retain items in memory at a fixed, yet high resolution. Therefore, if set size exceeds this capacity limit, only a subset of these items will be retained in memory, resulting in guessing if an item not held in memory is probed at test. The task used by Zhang and Luck was similar to the colour version of the continuous report task used by Wilken and Ma (2004), however set size differed (Experiment 1 $N = 3$ or 6; Experiment 2 $N = 1, 2, 3$, or 6; Experiment 3 $N = 4$; Experiment 4 $N = 3$). Using standard estimation, Zhang and Luck (2008) recovered two parameters from the data, P_m which represented the probability an item was held in memory and *s.d.* which represented how precisely the probed item was represented in memory.

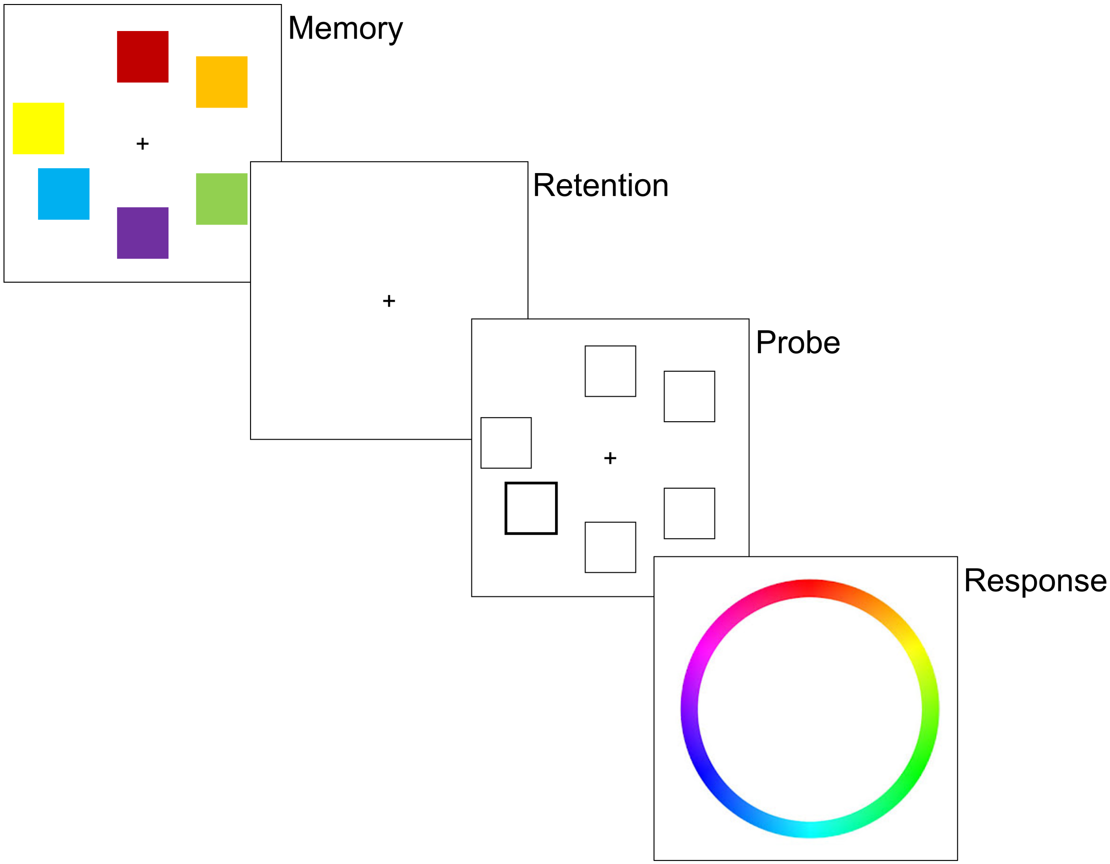

Figure 2
Figure showing an example trial procedure within the continuous (colour) report task.

Results from Experiment 1 revealed that while *s.d.* did not vary to any significant extent, P_m was approximately twice as large when at set sizes of three relative to set sizes of six. Zhang and Luck state that this is evidence to support the notion that a small number of high-resolution representations are stored in memory (i.e., the *slot model*). In Experiment 2, the authors set out to examine predictions made by two differing versions of the slot model, namely the *slots-plus-resource* and *slots-plus-averaging* models. The slots-plus-resource model was developed on the basis of a suggestion that performance could be mediated by both slots and resources, positing that resources are allocated to a single item to generate a high-resolution representation before any remaining resource is allocated to another item. As such, at a certain point there will not be a sufficient amount of resource remaining to generate high-resolution representations of any further items, meaning these items are not held in memory at all. Alternatively, the slots-plus-averaging model was developed based on the suggestion that the process of generating memory representations of items is binary, either succeeding or failing. On this model, while the total number of slots is still low, items can be stored across multiple slots when set size does not exceed capacity. On both the slots-plus-resources and slots-plus-averaging models, precision will be increased when set size does not exceed capacity, with results from Experiment 2 supporting this view. As set size increased from one to three, precision as indicated by the *s.d.* parameter was reduced[1] yet held steady when set size increased to six. Conversely, P_m was found to decline as set size increased from one to three, with a more pronounced reduction as set size increased to six.

In Experiment 3, Zhang and Luck introduced the use of a cue presented simultaneously with the memory display to indicate which of the four coloured squares would be probed. On 70% of trials the cue was valid, on 10% of trials the cue was invalid with one of the uncued squares being probed, and the remaining portion of trials were made up of neutral cues wherein all four squares were probed. The authors stated that on the slots-plus-resources model, most of the available resource would be allocated to the cued

---

[1] Note that a reduction in precision is reflected by an increase in *s.d.* parameter values.

item, resulting in substantial differences in *s.d.* between all cue types whereas little change would occur in P_m. Alternatively, on the slots-plus-averaging model, they state that most of the slots would be allocated to the cued item, resulting in substantial differences in P_m between valid and invalid trials, with *s.d.* also reduced on valid trials relative to neutral trials given the averaging of values in each slot. Results from Experiment 3 revealed that P_m was larger on valid trials relative to invalid trials, with *s.d.* significantly reduced on valid trials relative to neutral trials. Zhang and Luck (2008) take this as evidence against the slots-plus-resources model, stating that these results suggest low-resolution representations cannot be generated through the allocation of a small portion of memory resource. In extension of these results, Experiment 4 investigated the notion that an all-or-none encoding process is required to generate robust, high-resolution representations in memory. By placing visual masks in the positions that the stimuli previously appeared (either 110 or 340ms after stimulus onset), Zhang and Luck mimicked the masking caused by eye movements. They state that if representations increase in resolution as a function of time, early masking should result in low-fidelity representations (i.e., a reduction in *s.d.*). If however, an all-or-none process of encoding is required, masking would only influence P_m. Results from this experiment track with the latter of these hypotheses, with masking at 110ms causing a major decrease in P_m while having no impact on *s.d.* (however, see van Moorselaar et al., 2015), providing further evidence contrary to the slots-plus-resource model—and indeed pure resource models—as well as providing further evidence in support of the slots-plus-averaging model.

1.3.2 The two- and three-component mixture models

Building on these findings, Zhang and Luck (2008) developed a measurement model, referred to as the *two-component model*, wherein responses in a continuous report task are a probabilistic mixture of two processes (hence, two-component), 1) responding to the true value of the target based on a noisy memory representation and 2) random guessing. The two-component model is given formally as:

$$p\left(\hat{\theta}\right) = \left(1 - p_u\right)\phi_\kappa\left(\hat{\theta} - \theta\right) + p_u\frac{1}{2\pi}, \tag{2}$$

where θ represents the value of the target (in radians), $\hat{\theta}$ represents participant response, and p_u the probability of a random guess. ϕ represents the probability density of the von Mises distribution—the circular analogue of the Gaussian (normal) distribution—with mean zero and concentration parameter κ (see Figure 3, see also Grange & Moore, 2022). However, work by Bays and Husain (2008) highlighted methodological issues within the research by Zhang and Luck (2008) which led to the development of this model. Across two experiments, Bays and Husain (2008) investigated the precision of representations in memory as a function of set size, while also controlling for the effects of eye movements, something overlooked by Zhang and Luck (2008). Stimuli consisted of coloured squares (*location* task) and coloured, randomly oriented arrows (*orientation* task). After the memory display and retention interval, the test display was presented wherein a single stimulus that appeared on the memory display reappeared. On the location task, this meant reappearance of a coloured square, displaced along the horizontal axis, while in the orientation task, an arrow reappeared rotated either clockwise or counterclockwise. In Experiment 1 set size was varied ($N = 1$, 2, 4, or 6) while in Experiment 2 set size was held constant ($N = 5$). Four conditions were also implemented in Experiment 1; in the *fixation* condition, participants fixated on a cross which was displaced from the centre of the screen (alternating between left and right on each trial), while in the *saccade* condition, participants made a saccade from the fixation cross to a target stimulus upon sounding of an auditory signal presented 1,000ms after the memory display. Participants were assigned a colour before beginning the task which corresponded to the colour of one of the stimuli displayed. In the *fixation-with-cue* and *saccade-to-cue* conditions, one of the stimuli flashed on and off; in the fixation-with-cue condition, participants were required to remain fixated on the cross, while in the saccade-to-cue condition, the flashing indicated that participants should saccade to the flashing stimulus. The targets in the saccade, fixation-with-cue, and saccade-to-cue conditions was not indicative of which stimulus would be probed on a given trial.

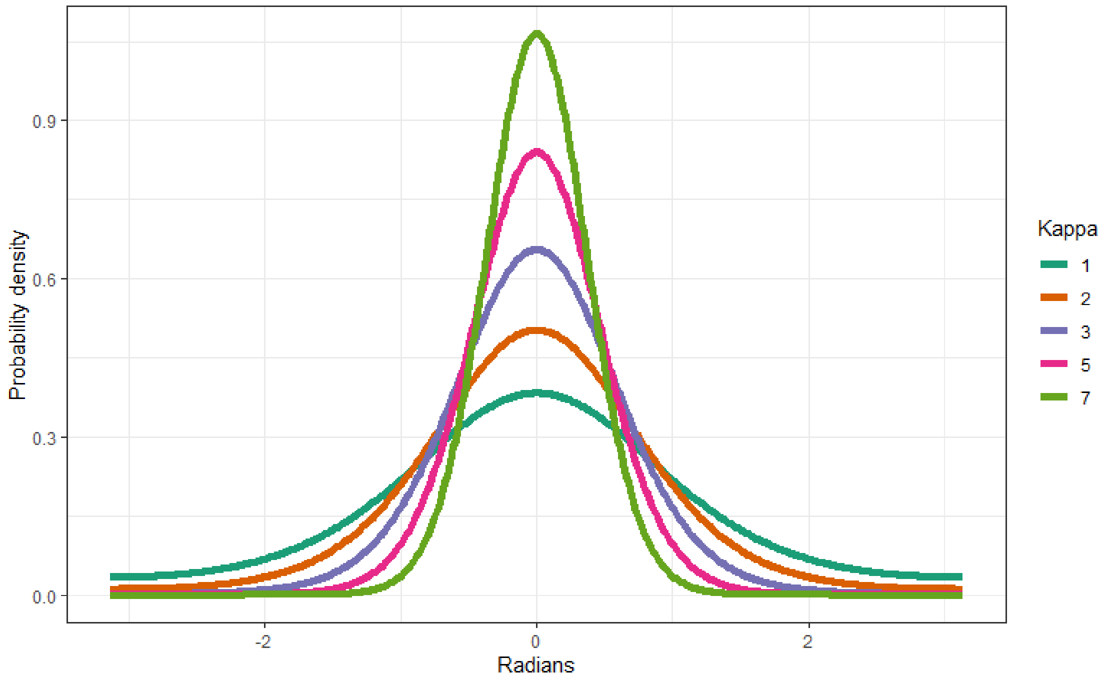

Figure 3

Probability density functions for different values of κ, the concentration parameter of the von Mises distribution. A higher value of κ indicates more precise memory representations.

Results from Experiment 1 revealed that accuracy in both the location and orientation tasks was quite high for single object displays when no eye movement was involved, with performance declining as a function of set size. However, in contrast to the pronounced decline in performance observed by Zhang and Luck (2008) when set size exceeded four, no such reduction was found here. However, performance on single-object displays did result in a degree of bias on the location task when a saccade was introduced; participants tended to report that the direction of change was congruent with the direction of the saccade when in fact, the stimulus had been displaced in the opposite direction. Despite this, the precision of responses when saccades were required was not significantly different from precision when fixation was required, while also observing a similar pattern of poorer performance as set size increased. Indeed, results from both the fixation and saccade conditions revealed that precision declined even at smaller set sizes, with the most substantial reduction in precision between set sizes of one and two, well below the putative capacity limit suggested by previous research. As such, Bays and Husain argued that a model wherein vSTM resources are distributed among items, with precision a declining function of set size, is a more appropriate alternative than a model which invokes a discrete capacity limit. Furthermore, in the saccade condition, it was found that the saccade target was recalled with greater precision than non-targets in both the location and orientation tasks; the authors highlighted this as evidence to suggest some form of preferential treatment for saccade targets, allowing more resource to be allocated to this item, thus increasing the precision of memory representations. Indeed, this is supported by results from the fixation-with-cue and saccade-to-cue conditions, wherein the flashing stimulus was recalled with greater precision when probed relative to when a non-flashing stimulus was probed, suggesting that even covert shifts of attention allow for an increase in precision.

In Experiment 2, Bays and Husain (2008) set out to investigate the dynamic allocation of resources across several saccades; participants made saccades to each item in turn (squares in the location task, arrows in the orientation task), with fixations on each item indicated by a clicking noise. When a saccade to the final target was detected, the screen went blank and was followed by a probe display. Results from this experiment revealed that precision was greatest for the final item in the display in both the location and orientation tasks; interestingly, the final item in each display was not fixated as the screen was blanked upon initiation of the final saccade. The authors also found no significant differences in precision between the last fixated stimulus (item four) and any of the previously fixated stimuli, no significant differences in precision between any of the previously fixated stimuli, and that fixation duration did not have any influence over precision. Based on these results, Bays and Husain (2008) state that the high degree of precision observed for the target of a saccade can only survive a single eye movement.

Extending this work, Bays et al. (2009) employed an almost identical continuous report task to that used by Zhang and Luck (2008), with the exception of different set sizes (N = 1, 2, 4, or 6). Additionally, Bays et al. (2009) also examined performance across a range of stimulus display durations (100, 500, or 2000ms) and also monitored fixations. Results from this experiment revealed that as set size increased—even from one to two—response precision declined, akin to findings from Bays and Husain (2008), but conflicting with findings from Zhang and Luck (2008). More crucially however, given that the probe indicates the position of one of the stimuli in the memory display, Bays et al. (2009) highlighted that performance on this version of the continuous report task is also dependent on memory for *location* as well as simply memory for colour. Participants therefore need to compare the location of the probe with the location of items held in memory, with errors in memory for location causing participants to respond to a non-target item. Indeed, the finding that responses were centred on non-target feature values at a rate higher than what would be expected from chance alone—with such responses increasing as set size increased—provided support for the view of Bays et al. that on some trials, memory for location failed, resulting in a swap between two of the colours presented in the memory display, leading to an erroneous response. As such, Bays et al. (2009) extended the two-component model (see Zhang & Luck, 2008) to the *three-component model* (also referred to as the *swap model*). This extension enabled the model to account for memory for location, wherein responses are based on a probabilistic mixture of three processes, 1) responding to the true value of the target based on a noisy memory representation, 2) responding to a non-target value, and 3) random guessing. This model is given formally as:

$$p\left(\hat{\theta}\right) = (1 - p_u - p_n)\,\phi_\kappa\left(\hat{\theta} - \theta\right) + p_u\frac{1}{2\pi} + p_n\frac{1}{n}\sum_{i}^{n}\phi_\kappa\left(\hat{\theta} - \theta_i^*\right), \tag{3}$$

where p_u, ϕ, and κ are the same as in the two-component model, with p_n representing the probability of making a non-target response and θ_i^* (θ_1^*, $\theta_2^*\ldots$) representing the non-target feature values. Note that both target and non-target items are stored with the same precision (i.e., κ is identical for both target and non-target items) as the probed item is not yet known (see Grange & Moore, 2022).

1.3.3 The role of attention in visual short-term memory resource allocation

As alluded to in the introduction of this chapter, the purpose of the present research is to investigate the way in which vSTM resources are allocated to relevant information when other, irrelevant information is also present. Attentional control is likely to play a role in such allocation and indeed, some studies have investigated this potential (see e.g., Dube et al., 2017; Emrich et al., 2017; Henderson et al., 2020). For example, Emrich et al. (2017) highlighted that much previous research did not account for the possibility of flexible resource allocation across items. Thus, across two experiments, Emrich et al. used continuous report tasks wherein stimuli consisted of coloured squares and set size was varied

(Experiment 1a N = 6; Experiment 1b N = 1 or 4). Predictive spatial cues were presented alongside the memory display, varying in both number and their validity (Experiment 1a N = 1 [100%, 50%, or 33%], 2 [100% or 66%], 3 [100%], or 6 [100%]; Experiment 1b N = 1 [100%] with set size N = 1; Experiment 1b N = 1 [100% or 33%], 2, [100% or 66%], 3 [100%], or 4 [100%] with set size N = 4), with validity indicated at the beginning of each block. By manipulating cue validity, the distribution of attention would be variable within each condition, with cases where non-cued items had a low probability of being probed, enabling the estimation of encoding precision for (potentially) less attended items.

While results revealed that as the total number of items to be retained in memory increased response error also increased, Emrich et al. found that probe likelihood had a greater influence over response error, showing larger effects of doubling probe likelihood relative to doubling set size. Further results from parameter analysis after fitting of the three-component mixture model (see Bays et al., 2009), revealed decreases in precision as set size increased, with decreases observed between the three cue and six cue (100% valid) conditions. This is notable as it directly conflicts with the view adopted by Zhang and Luck (2008); as previously discussed, vSTM can hold approximately four or five item representations (see e.g., Luck & Vogel, 1997; Pashler, 1988; Zhang & Luck, 2008). Based on their results, Zhang and Luck (2008) state that when capacity is exceeded, the precision of representations held in memory should not decrease as all item representations are held at a high and fixed resolution. That Emrich et al. (2017) found decreases in precision as set size increased from three to six items provides evidence against the notion that vSTM capacity should be defined in terms of discrete slots, each being able to store a single item.

Additionally, Emrich et al. tested assumptions of the slots-plus-averaging model (Zhang & Luck, 2008), specifically the notion that items represented in memory could not be represented at a resolution lower than that afforded by the resources available in a single slot. Emrich et al. found that precision decreased on trials wherein uncued items could be probed when compared to trials wherein three (100% valid) cues were presented, suggesting that such a reduction was due to the validity of the cue—and thus, the allocation of attention—mediating the amount of resource allocated to each item.

While this and other research has provided some insight into the role of attention in resource allocation, much research surrounding vSTM in general makes use of simple stimuli, defined by a single feature dimension (typically colour). This is atypical of everyday life wherein stimuli are often multivalent, containing any number of different feature dimensions, or indeed, the same feature dimension with multiple feature values. Furthermore, there are often times when only a subset of the information on a given stimulus in our environment is relevant to a task or behaviour (e.g., road traffic signs containing multiple directions). Thus, if vSTM is limited in capacity in terms of a finite memory resource, how does the cognitive system ensure that the resources we have available are allocated to the most relevant information? The way in which this question will be addressed throughout the present research is through a combination of tasks probing vSTM performance and modified task switching paradigms, enabling investigation of the role of attention in resource allocation when multiple feature dimensions are present.

1.4 Task switching

1.4.1 Overview and initial research

The term *task switching* refers to the unconscious ability to shift attention between different tasks and is one of the executive functions, a set of cognitive processes which facilitate the control of behaviour (see e.g., Diamond, 2013). In typical task switching paradigms, participants either repeat performance of the same task or switch between performance of two (or more) tasks; a consistent finding from the task switching literature shows that when there is a requirement to switch between tasks, performance in terms of accuracy and RT is reduced, relative to when the task repeats. This detriment to performance is known as the *task switch cost* (for comprehensive reviews on task switching, see Grange & Houghton, 2014; Kiesel et al., 2010; Monsell, 2003; Vandierendonck et al., 2010). Various theories have been postulated in an attempt to explain the origin of the task switch cost, with most theories implicating the *task set* as having some responsibility for these costs. Put simply, the term task set refers to the organisation of processes and functions required to perform an upcoming task; however, the composition of a task set varies among researchers (see e.g, Grange & Houghton, 2014; Logan & Gordon, 2001; D. W. Schneider & Logan, 2007). While this rather simplified definition of task set is similar to descriptions given by some researchers (see e.g., Kiesel et al., 2010; Rogers & Monsell, 1995), others have suggested that a task set contains a set of parameters which are programmed by the cognitive system depending on what task is being performed (see e.g., Logan & Gordon, 2001; Logan & Schneider, 2010). Irrespective of definition and constitution, a commonly held view on task sets is that when the task changes, the task set must be updated in some way[2]. Additionally, a task set can be thought of as being made up of two distinct components, an *attentional set*, which facilitates the identification and selection of relevant information, and an *intentional set*, which facilitates assignment of appropriate stimulus-response (S-R) mappings (see e.g., Kopp et al., 2006; Liefooghe & Verbruggen, 2019; Meiran et al., 2000; Rushworth et al., 2005, 2002; Yeung & Monsell, 2003b).

The most widely-cited early research on task switching was that conducted by Jersild (1927; however, see also Ach, 1910–2006), wherein participants were presented with lists of stimuli upon which to perform tasks. For example, lists could be made up of a series of two-digit numbers wherein participants would simply subtract three from each number (i.e., a *pure list*). Alternatively, lists could be made up of alternating two-digit numbers and words, wherein participants would subtract three from each of the numbers and provide an opposite for each of the words (i.e., a *mixed list*). Somewhat counterintuitively, Jersild found that mean RT across two pure lists was slower than RT on a mixed list. While this finding was later replicated by Spector and Biederman (1976; Experiment 1), the authors highlighted that the difference between RTs for pure and mixed lists was minimal (140ms). Additionally, Spector and Biederman also noted some methodological issues[3] that may have facilitated faster performance on mixed lists in Jersild's (1927) study; the lists used by Jersild were presented on a sheet allowing preview of the upcoming item. As such, Spector

---

[2] There are alternative theories which suggest that task sets are not updated at any stage. On such theories, a benefit to performance is conferred on repetition trials due to *repetition priming* of the task set (see e.g., Altmann & Gray, 2008).

[3] A further methodological issue is that concerning the measurement of RT, which was recorded using a stopwatch given either to an assigned partner or to the participant. As such, it is unlikely the RTs obtained by Jersild (1927) provided an accurate measure of performance.

and Biederman (1976) also included a condition wherein each item of the lists used were presented on separate cards, finding a reversal in performance; with card presentation, RTs were slower on mixed lists relative to pure lists. However, Spector and Biederman also noted that while cards prevented preview to a greater extent than sheet presentation, it may not have fully eliminated the ability to preview an upcoming item, with predictability also potentially influencing performance. In Experiment 2, list items were projected individually, preventing preview and items on mixed lists were presented at random, reducing the predictability of an upcoming item. Results from this experiment revealed that RT on pure lists was slightly, though non-significantly, faster than RT on mixed lists.

Perhaps the most crucial of the experiments conducted by Jersild (1927) and then later by Spector and Biederman (1976) was that which used *bivalent stimuli*, wherein multiple tasks could be performed on the same stimulus. Up to this point, each of the items within the mixed lists used by Jersild (1927) indicated to participants which task to perform (i.e., numbers indicated the subtraction task should be performed, while words indicated the opposite word task should be performed). Thus, Jersild investigated performance on mixed lists wherein the items did not indicate which task to perform; these lists contained either two-digit numbers wherein the participant alternated between adding and multiplying or adding and subtracting a given value, or words wherein participants alternated between providing the opposite to the word or providing an object for a verb. Overall, results showed that on these bivalent mixed lists, RTs were slower relative to the pure lists wherein only a single task was performed. Indeed, Spector and Biederman (1976; Experiment 3) also observed slower RTs on bivalent mixed lists wherein participants alternated between adding and subtracting three, relative to pure lists, with this cost to performance being 402ms per item on mixed lists. Furthermore, Spector and Biederman (Experiment 4) also found that the addition of a cue (i.e., +3 or -3 was presented alongside each two-digit number), resulted in marked reductions to RTs relative to those found in Experiment 3; however a cost[4] to performance between mixed and pure lists was still observed, although the authors suggested that this may have reflected the time taken to read the cue indicating which task to perform on mixed lists, with such a requirement unnecessary on pure lists.

1.5 Theories of switch costs

1.5.1 Task set inertia

Despite the methodological issues, Jersild's (1927) work provided the basis for later research into task switching, resulting in the development of new paradigms for use in investigations, as well as the development of various theories attempting to account for the observation of task switch costs. One such theory is that of *task set inertia* (TSI), developed by Allport et al. (1994). Across a series of experiments, Allport et al. used tasks similar to the list procedure employed by Jersild (1927; Experiments 1-4), varied the *response-stimulus interval* (Experiment 5), and used *rapid sequential visual presentation* tasks (Experiments 6 and 7). Perhaps the most critical of these experiments for the development of the TSI account was Experiment 5 wherein the RSI—the duration between response on trial *N-1* and presentation of stimuli on trial *N*—was manipulated, being either 20, 550, or 1,100ms long within a given block. Stimuli were made up of pairs of incongruent Stroop colour words (see

---

[4] Given that the list procedure used by Jersild (1927) does not allow for the estimation of true switch costs, the cost to performance observed here and in studies using the list procedure is referred to as the *alternating list cost* or *global switch cost* (see e.g., Vandierendonck et al., 2010).

e.g., MacLeod, 1991; Stroop, 1935), meaning the word and the colour the word was written in never matched (e.g., the word “red” written in the colour blue). As such, the tasks would be to either name the colour word (i.e., *word* task) or the colour the word was written in (i.e., *colour* task; colours used were red, blue, green, yellow, and pink), with participants informed beforehand which task to perform on which item in each pair (e.g., perform the colour task for the first item in each pair and perform the word task for the second item in each pair). Furthermore, neutral trials were also included wherein stimuli consisted of colour words written in black if the task was to name the colour word, or *xxxxx* in one of the colours used if the task was to name the colour of the word; thus, on neutral trials, the stimuli explicitly cued the participant as to which task to perform. This allowed not only examination of the Stroop effect via comparison of incongruent and neutral trials, but also investigation of the extent to which costs to performance may be influenced by the innate cues present on the neutral stimuli.

Previous research (see e.g., MacLeod, 1991) had shown that performance in the colour task, but not performance on the word task, was dependent upon the congruency of the stimuli. When stimuli were congruent (i.e., the colour word and the colour the word was written in matched), performance on the colour task was improved relative to when the stimuli were incongruent. However, in the word task, performance was found to not differ to any great extent regardless of the congruency of the stimuli. As such, it is held that the word task is automatic, requiring little-to-no cognitive control to perform (i.e., the *dominant* task), whereas the same cannot be said for the colour task (i.e., the *non-dominant* task). Therefore, it would be plausible to assume that switching from the non-dominant to the dominant task would be much quicker than switching from the dominant to non-dominant task. However, contrary to this assumption, results from Experiment 5 of Allport et al. (1994) revealed the opposite effect; switching from the non-dominant to the dominant task resulted in a substantial switch cost relative to the lack of cost when switching from the dominant to non-dominant task (i.e., an *asymmetric switch cost*), independent of the duration of the RSI. This led Allport et al. (1994) to suggest that when a switch in task occurs, the task set relevant for the previous trial (i.e., trial *N-1*) interferes with the task set associated with the now relevant trial (i.e., trial *N*) and must therefore be inhibited to prevent this *proactive interference* (i.e., the TSI account of task switch costs). The required level of inhibition is dependent upon the dominance of the task previously performed, with a switch from a dominant to non-dominant task requiring a greater level of inhibition to prevent interference than the inverse. As such, any subsequent switch from a non-dominant task to a dominant task will require recovery of the relevant task set from a greater level of inhibition.

While some research showed further evidence of this asymmetry (see e.g., Allport & Wylie, 2000; Yeung & Monsell, 2003a, 2003b), other research found no such asymmetrical switch cost (see e.g., Monsell et al., 2000; Reuter et al., 2006; Yeung & Monsell, 2003a), with Wylie and Allport (2000) suggesting that *negative priming* rather than proactive interference was responsible for the asymmetric switch costs observed (see also Waszak et al., 2003, 2005, 2004). Additionally, Bryck and Mayr (2008) observed asymmetrical costs with task repetitions. Stimuli in Experiment 1 of Bryck and Mayr consisted of an arrow pointing either to the left or right, while stimuli in Experiment 2 consisted of incongruent Stroop colour words. In Experiment 1, participants responded with the direction of the arrow presented on screen by pressing either the left or right arrow key on a keyboard; however, response compatibility was manipulated, with the dominant condition mapping responses to the compatible key (e.g., left facing arrow indicated a left arrow key press), while in the non-dominant condition, responses were mapped to the incompatible key (e.g., left facing arrow indicated a right arrow key press). In Experiment 2, participants vocalised either the colour word (i.e., dominant condition) or the colour the word was presented in (non-dominant condition). Crucially, an *alternating-runs task*[5] was used wherein participants performed four trials of one task before switching to the other task (e.g., AAAABBBB), with each run further broken down into pairs of trials (e.g., AA-AA-BB-BB), giving rise to task switches (e.g., AA-BB) and task repetitions (e.g., AA-AA). The RSI between each trial within a pair (i.e., AA or BB) was held constant at 50ms in Experiment 1 and 500ms in Experiment 2, with the longer duration in Experiment 2 due to the requirement of a vocal response; however the RSI between each pair (e.g., AA-AA, AA-BB) was varied to be either 500ms or 5,000ms. Bryck and Mayr (2008) attempted to measure what they called the *selection cost* by comparing performance on the first and second trial within each pair (independent of whether the trial repeated or switched), hypothesising that re-retrieval of the relevant task set from long-term memory would be required at longer RSIs even if the task did not change. The results from these experiments supported this hypothesis, with Bryck and Mayr finding reliable asymmetrical costs for both task switches and task repetitions, leading the authors to suggest that task switches are not a necessary condition to observe an

---

[5] The alternating-runs task was developed by Rogers and Monsell (1995) to account for methodological issues which arose from Jersild's (1927) earlier work, namely the various imbalances between pure and mixed blocks of trials, such as the number of task sets that were required in each type of block or the effort required for performance of each block type.

asymmetrical selection cost.

1.5.2 Task set reconfiguration

Manipulation of the RSI was also involved in the development of a further prominent theory of task switch costs, the *task set reconfiguration* (TSR) account. Across a series of experiments, Rogers and Monsell (1995) used alternating-runs tasks wherein participants performed two trials of one task before switching to the other task (e.g., AABBAA; Experiments 1-5) or performed four trials of one task before switching to the other task (e.g., AAAABBBB). Stimuli consisted of a letter and digit pair (e.g., G7) presented in one of four quadrants of the screen (Experiments 1-5); on each successive trial, the stimulus moved clockwise to the next quadrant, with imaginary vertical and horizontal boundaries indicating a change in task. For example, if the stimulus appeared in the top-left quadrant in the first trial, it would then move to the top-right quadrant on the second trial wherein participants would perform the same task. However, when the stimulus moved to the bottom-right quadrant, a change in task would occur. In Experiment 6, a circular shape was presented, segmented into eight pieces, with a thicker line along either the horizontal or vertical indicating when a change in task was required; stimuli were again presented in a clockwise manner. Overall, these experiments revealed increased error rates and RTs on switch trials relative to repetition trials. While the extent of these switch costs were modulated by the duration of the RSI (i.e., increased RSIs reduced the cost associated with a switch), leading to Rogers and Monsell (1995) suggesting switch costs reflected the time course of task set reconfiguration, there still remained a *residual switch cost*, even when the RSI increased to 1,200ms. On the TSI account, such a cost is easily explained on the basis that interfering processes occur independently of the extent of preparation time available during a switch in task; however, on the TSR account, an additional mechanism to account for such a cost was required. Thus, Rogers and Monsell outline that the process of switching task involves an *endogenous control process* required for updating the task set, with a further *exogenous* process initiated upon presentation of the stimuli required to complete the process of reconfiguration.

1.6 Subsequent research on task switching

1.6.1 The explicit cuing procedure

While the work by Allport et al. (1994) and Rogers and Monsell (1995) led to the development of two distinct theories of switch costs, an issue remained in that the tasks used to develop these accounts do not allow for distinction between the processes that contribute to such costs. Both mechanisms suggested by the TSI and TSR accounts occur during the RSI and as such, neither of the tasks used by Allport et al. (1994) and Rogers and Monsell (1995) could decompose the relative contributions of each of these mechanisms to the switch cost. However, a novel paradigm developed by Meiran (1996; see also Sudevan and Taylor, 1987) which presented cues at the beginning of each trial to indicate which task to perform (i.e., the *explicit cuing procedure*) enabled independent investigation of the contributions of TSI and TSR on switch costs. For example, if a series of single digits are presented as stimuli, the cue at the beginning of some trials could be the word *value* indicating that the task to perform is a higher or lower than five judgement or the word *parity* indicating that the task to perform is an odd or even judgement. Crucially, by presenting a cue at the beginning of each trial, the RSI is segmented into two distinct parts, the *response-cue interval* (RCI), the duration between response on trial *N-1* and cue presentation on trial *N* and the cue-stimulus interval (CSI), the duration between presentation of the cue and stimulus on trial *N*. During the RCI, the cue is not yet known meaning that no reconfiguration of task set can occur; as such, it is during this interval that inhibition and/or *task set dissipation* of the previous task set occurs, in line with the TSI account. Conversely, during the CSI, now that the cue is known, reconfiguration of task set can occur in line with the TSR account. It should be noted that inhibition of the previously relevant task set could still hypothetically occur during the CSI, however this interval is primarily associated with the updating of task set.

In a series of experiments, Meiran (1996) used a 2x2 grid wherein a circle stimulus appeared in one of the four quadrants. The goal of this task was to report the location of the stimulus either in terms of the vertical (i.e., up or down discrimination) or horizontal (i.e., left or right discrimination) by pressing the relevant key on a numberpad; arrows presented alongside the empty grid prior to stimulus presentation indicated which of the two tasks was to be performed. Across all experiments, the CSI (termed the cue-target interval by Meiran) was also manipulated; for instance, in Experiments 2 and 3, the CSI could either be long (1,716ms) or short (216ms). Results from these experiments revealed that increasing the CSI resulted in reductions to—but not elimination of—switch costs. Meiran (1996) highlighted that these results are in line with the notion of advance reconfiguration proposed by the TSR account, however also suggested that the results obtained by Rogers and Monsell (1995) which led to the development of the TSR account could be explained by the proactive interference employed by the TSI account by Allport et al. (1994). While Rogers and Monsell (1995) found a significant interaction between the RSI and task switch costs when

RSI varied within a block of trials only, Meiran (1996) observed a significant interaction between the CSI and task switch costs when CSI varied within a block of trials. Furthermore, the switch costs observed by Meiran were almost entirely eliminated with the use of a substantially shorter interval than that used by Rogers and Monsell (1995). Together, these findings directly contrast two assumptions made by Rogers and Monsell (1995) in relation to conditions that are required for advance reconfiguration to occur, 1) that stimulus presentation must be predictable and 2) that stimulus presentation must be temporally distinct enough from the cue to allow reconfiguration to occur. Meiran (1996) highlighted that one plausible reason for the discrepancy between the results of the two studies is the nature of the cues used. While Rogers and Monsell (1995) used exogenous cues, Meiran (1996) suggested that the use of explicit cues is likely to be more robust against varying interval durations within a block of trials. As such, while the results from Meiran (1996) provided some evidence in support of an advance reconfiguration mechanism, Meiran did not explicitly invoke a endogenous control process as being responsible. Furthermore, later research by Meiran et al. (2000) provided evidence to suggest that the notion of task set dissipation cannot yet be ruled out.

Across a series of experiments Meiran et al. (2000) used a paradigm similar to that used previously in Meiran (1996) wherein a stimulus (smiling face character) appeared in one of the four quadrants of a 2 x 2 grid. The task again was to report the location of the stimulus either in terms of the vertical or horizontal by pressing the relevant key on a numberpad. Arrows presented alongside the empty grid prior to stimulus presentation again indicated the task to be performed on each trial. Critically, within this set of experiments the RCI was manipulated (132ms at shortest, 3,032ms at longest), with results showing that switch costs were reduced—but again, not eliminated—with increases in RCI. Thus, these results suggest that some form of passive dissipation of task set occurs once response is provided, somewhat in line with that suggested by proponents of the TSI account (see e.g., Allport et al., 1994). Furthermore, additional reductions to the extent of the switch cost were observed by increasing the duration of the CSI (again termed the cue-target interval by Meiran et al., 2000). Taken together, Meiran et al. suggest that switch costs reflect an interplay of three different mechanisms, 1) dissipation of the now irrelevant task set, 2) reconfiguration of task set for an upcoming task, and 3) some residual component associated with the switch cost remaining despite sufficient time to prepare for an upcoming switch in task.

1.6.2 Cue switch costs

One shortcoming of the explicit cuing procedure however, relates to the fact that when a task repeats or switches, the cue follows suit. As such, it is plausible that some degree of the task switch cost could be attributed to a switch in cue. A method for combating this confound is to use two cues per task (Logan & Bundesen, 2004; see e.g., Logan & Bundesen, 2003; Mayr & Kliegl, 2003; Monsell & Mizon, 2006), which results in three distinct sequencing conditions: *cue repetitions*, wherein cue and task repeat, *cue switches*, wherein cue switches but task repeats, and *task switches*, wherein both cue and task switch. For example, if single digits are used as stimuli, cues for a magnitude judgement task could be *low* and *high*, while cues for a parity judgement task could be *odd* and *even.* Use of this *2:1 mapping* means that the *true* or *corrected task switch cost* can be found by comparing performance between cue switches and task switches, while the influence of a cue switch can be assessed through comparison of performance between cue repetitions and cue switches. Results from cue switching research have been quite mixed, with some studies showing large cue switch costs with small corrected task switch costs (Logan & Bundesen, 2004; see e.g., Logan & Bundesen, 2003; Monsell & Mizon, 2006), while other studies reported more substantial corrected task switch costs (Mayr & Kliegl, 2003; Monsell & Mizon, 2006). The initial work by Logan and Bundesen (2003) on cue switching led the authors to develop a formal model which stated that a compound representation of both the cue and stimulus is generated, with responses made on the basis of this compound representation. For example, in a task switching paradigm wherein participants perform either a magnitude or parity judgement, if the cue *odd* is presented with the digit *3* presented as the stimulus subsequently, a response (i.e., keypress) will be made on the basis of the compound *odd-3*. This *compound cue encoding* facilitates the retrieval of the appropriate response from long-term memory, meaning that no endogenous reconfiguration process is required as the task remains the same regardless of which of the two judgement tasks are being performed: respond to the compound representation. Indeed, comparison of three models, one wherein task set reconfiguration occurred, one wherein compound cue encoding occurred, and one wherein a combination of both occurred, revealed that the compound cue encoding model provided the best fit to the data. An alternative explanation of cue switching offered by Mayr and Kliegl (2003; see also Mayr and Kliegl, 2000) also implemented a mechanism of retrieval from long-term memory, in this case the retrieval of S-R mappings, with a further process involved in task implementation upon stimulus presentation. Crucially however, neither the TSI or TSR accounts of task switch costs offer an explanation for the influence of a cue switch on the true switch cost, hence the necessity of further models.

1.6.3 Attentional set switching

Perhaps most relevant to the present research is previous task switching work which examined the impact of switching *attentional* but not *intentional* set. To reiterate, attentional set relates to the processes involved in identifying and selecting relevant information, while the intention set refers to the assignment of a S-R mapping that is used to perform the given task (see e.g., Kopp et al., 2006; Liefooghe & Verbruggen, 2019; Meiran et al., 2000; Rushworth et al., 2005, 2002; Yeung & Monsell, 2003b). One such study by Meiran and Marciano (2002) made use of a same or different judgement task wherein stimuli consisted of four feature dimensions, *shape*, *fill*, *size*, and *tilt* however, only shape and fill were ever relevant. Participants were placed in one of three groups, the *dimension* group wherein the task was always the same (e.g., *is the relevant feature different?*), but where the feature dimension could change at random, the *decision rule* group wherein the relevant feature dimension remained the same, but the task could change at random (i.e., *is the relevant feature different?* or *is the relevant feature the same?*), and finally the *response mapping* group wherein both task and feature dimension remained the same, but response mapping could change randomly (e.g., if yes is mapped to the right arrow key and no to the left, these could be reversed). The aim of this study was to examine how switching between feature dimensions was impacted by manipulations of the CSI (again termed the cue-target interval by Meiran & Marciano, 2002). These manipulations meant the CSI was either 170, 470, 1,470, or 2,470ms with results from Experiment 1 showing that increases in CSI duration resulted in reductions to the switch costs observed in both decision rule and response mapping groups, but not in the dimension group. In Experiment 2, Meiran and Marciano sought to extend these findings by comparing the *simultaneous presentation* used in Experiment 1 with *sequential presentation*; while both stimuli were presented concurrently in Experiment 1 following presentation of the cue, in Experiment 2, after presentation of the cue, half of the participants were presented with one of the two stimuli for 300ms before the second stimulus was presented. This was based on previous research which had found costs to performance when switching between simultaneous and sequential presentation modes (see García-Ogueta, 1993), as well as other research which observed reduced interference with sequential presentation relative to simultaneous presentation (see Santee & Egeth, 1980).

Results revealed a significant main effect of presentation, with faster responses observed in sequential presentation (706ms) relative to simultaneous presentation (993ms), with a significant main effect of switching also observed; here RTs were faster on repetition trials (816ms) relative to switch trials (883ms). These results indicate that while there was an impact of switching, a substantial advantage to performance was obtained from sequential presentation in line with results obtained from the previous work by Santee and Egeth (1980). However, Meiran and Marciano (2002) also found a non-significant main effect of CSI as well as a non-significant interaction of CSI and task switching, with Meiran and Marciano highlighting that extended CSIs actually caused slower performance on repetition trials, albeit exclusively found within the simultaneous presentation condition. The authors state that such an underadditive interaction does not reflect a process of reconfiguration, but rather acts as an index of task set dissipation; given that the task set remains the same on a repetition trial, dissipation of the task set would result in reduced performance. Therefore, as RT was not recorded until presentation of the entire stimulus display in the sequential presentation condition (i.e., 300ms later than in the simultaneous presentation condition), this appears to have allowed the task set to dissipate to a greater extent, resulting in participants being more prepared for the upcoming task.

1.7 The present research

The aim of the present research is to examine the role of attentional control in vSTM resource allocation in the presence of multiple feature dimensions (e.g., colour, orientation) using a variety of different tasks which probe vSTM, namely change detection, continuous report, and change localisation. These tasks will be embedded within task switching paradigms which introduce an element of attentional control; critically, in the experiments throughout this thesis, participants will not repeat or switch between different tasks, but rather perform the same task (e.g., change detection) wherein the relevant feature dimension will either repeat (e.g., colour–colour) or switch (e.g., colour–orientation) on a given trial. As the capacity of vSTM is limited, the cognitive system must ensure that the resources available are allocated to the most relevant information, with the present work focusing on how attentional control may be involved in such allocation.

2 Chapter 2: Initial change detection

This chapter outlines an initial series of experiments which aimed to investigate the impact of switching between different feature dimensions by embedding change detection tasks within task switching paradigms.

2.1 Introduction

2.1.1 Overview

One's ability to temporarily retain information from the surrounding visual environment in memory is highly limited, with only a small amount of information able to be stored at a given time. It is postulated that this limitation of visual short-term memory (vSTM) arises as a result of either an upper-limit to the total capacity of the system (see e.g., Luck & Vogel, 1997; Pashler, 1988; Vogel et al., 2001; Zhang & Luck, 2008) or due to the allocation of a finite memory resource (see e.g., Bays et al., 2009; Bays & Husain, 2008; Frick, 1988; Wilken & Ma, 2004). On the former of these models, vSTM is believed to hold fixed and high resolution representations of approximately four items, with no information pertaining to items outside this capacity limit retained (see e.g., Luck & Vogel, 1997). On the latter model, it is believed that a memory resource is allocated to each item in the immediate visual environment, with the amount of resource each item receives responsible for the precision of the internal memory representation (see e.g., Bays et al., 2009).

Despite such explanatory models facilitating advances in our understanding of vSTM, some aspects still remain underdefined. The purpose of the present study is to provide an initial insight into one such aspect, specifically *how* the memory resource outlined by Bays and colleagues (see e.g., Bays et al., 2009; Bays & Husain, 2008; Ma et al., 2014; Wilken &

Ma, 2004) is allocated to relevant visual information when multiple features are present. Most of the existing research investigating resource allocation in vSTM makes use of univalent stimuli (i.e., stimuli defined by a single feature), most typically colour. This is atypical of our everyday environment wherein stimuli we encounter are often multivalent (i.e., stimuli consisting of any number of different features), for example road traffic signs or advertising. Thus, if only a subset of this information is relevant, the cognitive system faces a challenge to ensure that the limited vSTM resources available are allocated to the relevant information only.

The present study aims to provide a first look at the possible role of *attentional control* in such allocation by embedding *change detection* tasks within *task switching* paradigms. Change detection involves presentation of two stimulus displays (e.g., coloured squares) separated by some duration of time; on some trials, the display will remain the same, whereas on others it will change (e.g., a change in colour of one square). Observers simply report whether they believe a change occurred between the first and second displays. Task switching on the other hand involves either repeating performance of the same task or switching between performance of two (or more) tasks; a consistent finding within the literature shows that switching between tasks incurs a penalty to performance relative to repeating a task, known as the *task switch cost.* However, unlike traditional task switching paradigms, participants in the present study perform the same task throughout (i.e., change detection), with the relevant *feature dimension* (i.e., stimulus displays will always contain colour and orientation) either repeating or switching across trials. Therefore, should vSTM resource allocation rely on attentional control, performance on trials wherein there is a requirement to switch feature dimension should be poorer than performance on trials wherein the relevant feature dimension repeats.

While a number of studies have previously investigated the role of attention in vSTM resource allocation (see e.g., Dube et al., 2017; Emrich et al., 2017; Henderson et al., 2020), these studies used univalent stimuli with manipulations such as cue validity or investigations of age differences. It should also be highlighted that most research into resource allocation utilise *continuous report* tasks, wherein participants respond with the feature value of a probed item using a 360° wheel (e.g., a particular orientation). While such tasks are extremely beneficial for deconstructing performance into various parameters through the application of mixture modelling (see e.g., Grange & Moore, 2022), change detection is used in the present study to provide a more systematic account of resource allocation in vSTM; later chapters employ similar methodologies using continuous report (Chapter 4) and change localisation (Chapter 6) tasks.

2.1.2 Background on visual short-term memory capacity

Change detection tasks have perhaps been the most extensively used task in investigations of vSTM capacity (see e.g., Alvarez & Cavanagh, 2004; Luck & Vogel, 1997; Pashler, 1988; Phillips, 1974; Vogel et al., 2001). Perhaps one of the most influential papers on vSTM capacity is that by Luck and Vogel (1997), wherein the authors conducted a series of change detection tasks using univalent and multivalent stimuli. In one such experiment, the authors observed that change detection performance with coloured squares did not differ to any great extent up to set sizes of four; however, as set size increased above four, Luck and Vogel observed systematic reductions in performance. Through the use of Pashler's

(1988) *K measure*, a method for estimating memory capacity, it was found that approximately four items were retained in memory at a given time, a finding supported by later work (see e.g., Vogel et al., 2001; however, see also Olson & Jiang, 2002; Wheeler & Treisman, 2002; Xu, 2002). This led the authors to suggest that vSTM could only retain a small number of items in memory simultaneously, with a later experiment revealing a similar capacity limit even when items contained up to four features; this led Luck and Vogel to also state that vSTM stored integrated item percepts rather than individual features.

Despite the findings of Luck and Vogel (1997) receiving some support, Vogel et al. (2001) highlighted that it was possible that lower fidelity item representations which exceed the capacity limit may be retained, a possibility previously considered by Frick (1988). On Frick's model, the capacity limit of vSTM (termed the *visuospatial sketchpad* by Frick, see e.g., Baddeley & Hitch, 1974), was associated with a finite amount of "representational medium" distributed to items in the immediate visual environment and responsible for the fidelity of internal item representations. Clearly, such a view is in line with that suggested by Bays and colleagues (see e.g., Bays et al., 2009; Bays & Husain, 2008), with low fidelity representations being generated when set size (i.e., the total number of items presented) is high as less medium or resource can be allocated to each item. Additionally, Wilken and Ma (2004) provided evidence against a fixed-capacity nature of vSTM through a series of experiments using continuous report tasks wherein participants were presented with either coloured squares (Experiment 7; set size N = 2, 4, 6, or 8), Gabor patches (Experiment 8; set size N = 2, 3, 4, or 5), or Gabor patches with added spatial frequency (Experiment 9; set size N = 2, 4, 6, or 8).

Results revealed that the precision of participant responses decreased monotonically as set size increased. This led Wilken and Ma to state that if vSTM capacity is limited by the total number of items that can be retained simultaneously, the fidelity of internal representations of probed items should be high and equal until the capacity limit is reached, meaning accuracy of responses should be excellent; upon exceeding capacity, performance should decrease given that the probability of an item not stored in memory increases, resulting in increased variability of responses. However, given that Wilken and Ma observed reductions in precision as set size increased below the supposed capacity limit provided evidence against a fixed-capacity vSTM, with the authors suggesting that the limit may be artefactual due to increasing noise with increases in set size.

2.1.3 Development of measurement models

Later research by Zhang and Luck (2008) however appeared to rule out the possibility of low fidelity representations, suggesting that an "all-or-none" encoding process is required to generate robust item representations in memory. In a series of continuous report tasks, Zhang and Luck used standard estimation techniques to recover the parameters P_m, the probability that an item was represented in memory, and $s.d.$, a measure of how precisely the probed item was represented in memory. Across four experiments with varying set sizes (Experiment 1 $N = 3$ or 6; Experiment 2 $N = 1$, 2, 3, or 6; Experiment 3 $N = 4$; Experiment 4 $N = 3$), results revealed that precision as indicated by the $s.d.$ parameter did not vary to any great extent between set sizes of three and six, however P_m was found to be approximately twice as large at set sizes of three relative to set sizes of six. Such a result indicated that as set size increased past a certain threshold, there was a reduced probability that the probed item would be represented in memory. More critically to the notion of the generation of low fidelity representations however, Zhang and Luck (2008) used masking, a manipulation which mimics the effect of eye movements by disrupting the visual stream (see e.g., Agaoglu et al., 2015). The authors stated that by presenting the mask early, low fidelity representations could be generated, thus causing reductions in precision (i.e., increases in the *s.d.* parameter). Conversely, if an all-or-none encoding process was required, masking would only result in reductions of the probability that the probed item is held in memory (i.e., reductions in the P_m parameter). Indeed, results of this experiment track with these hypotheses, with presentation of a mask 110ms after stimulus presentation resulting in major reductions to P_m, with *s.d.* remaining unaffected (but see van Moorselaar et al., 2015). These results led Zhang and Luck to suggest that in order for robust representations to be generated, an all-or-none process of encoding was required, seemingly negating the claim that low resolution representations could be generated.

This work led to the development of the *two-component mixture model*, a measurement model for performance on continuous report tasks. On this model, responses in this task are a probabilistic mixture of responding to the probed item based on a noisy memory representation and random guessing. However, later work by Bays et al. (2009) highlighted that performance on continuous report tasks are not solely reliant on memory for a given feature value (e.g., the colour red), but also rely on memory for the *location* of that feature in the stimulus display; both the feature value and the spatial location of the feature value must be bound together to provide an accurate response. Bays et al. go on to state that should no information pertaining to location be held in memory it is possible that a non-target value could be reported instead, which on the two-component model would be classified simply as a guess. As such, Bays et al. (2009) extended the two-component model developed by Zhang and Luck (2008) to contain an additional component, with responses on this model a probabilistic mixture of responding to the target based on a noisy memory representation, random guessing, and the probability of responding with a non-target value (i.e., the *three-component mixture model*). Bays et al. (2009) fitted the three-component model to data obtained from a continuous report experiment using coloured squares and manipulations of set size ($N = 1$, 2, 4, or 6). Model fitting revealed that as set size increased, precision decreased and the frequency of non-target responses increased. Perhaps the most crucial finding of this experiment was that precision was reduced even when set size increased from one to two; this conflicts with the account put forward by Zhang and Luck (2008) as precision would not be influenced until capacity was exceeded. Bays et al. suggested on the basis of these results that vSTM capacity is better accounted for by a finite resource deployed across all items in the immediate visual environment, with increases in the number of items to be represented in memory causing reduced fidelity representations.

2.1.4 The role of attentional control

Building upon this work, later research by Marshall and Bays (2013) provided evidence which suggested that this vSTM resource could be allocated unintentionally by investigating the possibility of obligatory encoding of task-irrelevant features when attending to task-relevant features on the same item. In Experiment 1, three conditions were used; a *full-memory* condition wherein two sequential stimulus displays were presented each consisting of two coloured, oriented bars, with participants instructed to remember all eight features, a *cued-absent* condition wherein one stimulus display contained coloured circles and the other oriented bars, with participants instructed to remember all four features, and a *cued-present* condition wherein the stimulus displays were identical to those used in the full-memory condition with a cue indicating which feature should be retained on each of the stimulus displays. Following presentation of the displays, a probe was presented with response provided by turning a dial to change the colour or orientation of the probe to that which was believed to best match the stimulus which appeared in the same location.

It was revealed that error distributions from Experiment 1 showed less variability for the cued-absent condition relative to the full-memory condition, which Marshall and Bays highlight tracks with previous research on the relationship between precision and memory load (see e.g., Bays & Husain, 2008; Zhang & Luck, 2008). In contrast however, error distributions on the cued-present condition were similar to those on the full-memory condition, which the authors take as evidence to suggest that the inclusion of irrelevant features in the cued-present condition—wherein participants were instructed to remember four feature values—resulted in an increase in variability to similar levels as that observed in the full-memory condition—wherein participants were instructed to remember eight feature values. By calculating *precision gain*—a measure of precision improvements on each of the cued conditions relative to the full-memory condition—Marshall and Bays found a significant improvement in precision for both feature dimensions in the cued-absent condition but not the cued-present condition. The authors suggested that despite instructions, task-irrelevant features in the cued-present condition were obligatorily encoded into memory, with this encoding consuming vSTM resources to a similar extent as the encoding of actively attended features (i.e., full-memory condition).

In Experiments 2a and 2b, Marshall and Bays (2013) investigated if simply attending to an item with no instruction to remember it facilitated storage of irrelevant features. In Experiment 2a, four conditions were used; all conditions began with presentation of an initial stimulus display consisting of two coloured, oriented bars, with a requirement to remember the orientation of both bars. This was followed by a brief interval where the display was blanked. In the *full-memory* condition, a second stimulus display was then presented consisting of two more coloured, oriented bars with a requirement to remember the orientations of both bars. One of the four orientations could be probed for recall with equal probability. In the *match-feature-absent* condition, the second stimulus display consisted of two coloured circles, with instructions to press a button if the colours matched or withhold response if different. In the *match-different-feature* condition, the second stimulus display contained two coloured bars with instructions to press a button if the bars matched in colour. Finally, in the *match-same-feature* condition, the second stimulus display consisted of two coloured, oriented bars, with instructions to press a button if the bars matched in orientation. Participants were then also probed to report the orientation of one of the two stimuli presented on the initial display. Experiment 2b simply reversed the order of presentation for the memory and matching displays, with everything else remaining identical.

Results from Experiment 2a revealed that recall precision was higher in the match-feature-absent condition—wherein participants attended to non-oriented stimuli while retaining orientation information in memory—relative to the full-memory condition. This is in line with findings from Experiment 1 and previous research showing reductions in precision with increases in memory load. Contrastingly, in the match-different-feature and match-same-feature conditions—wherein participants attended to oriented stimuli during a matching task while also retaining orientation information in memory—recall precision was found to be comparable to the full-memory condition; Marshall and Bays highlighted that this increase in variability appeared to be due to simply attending oriented stimuli during the matching task, independent of whether the feature dimension was relevant to the matching task. Results from Experiment 2b—wherein comparison occurred in the first stimulus display with the second stimulus display being retained in memory—showed that recall precision in all three matching conditions was significantly higher than that observed in the full-memory condition. The authors state that this provides evidence to suggest that while irrelevant features are encoded into memory obligatorily, consuming vSTM resources, these resources can be redistributed when new information is presented, resulting in recall precision akin to that observed when no competing information is presented.

2.1.5 Using task switching to investigate attentional control

This study by Marshall and Bays (2013) clearly highlights a role for attentional control in the allocation of resources to visual information containing multiple feature dimensions. As can be seen from the results, when competing information is present, resources can be automatically allocated to task-irrelevant information, thus reducing the resources available for task-relevant information. Indeed, later studies have also examined the role of attention in vSTM resource allocation (see e.g., Dube et al., 2017; Emrich et al., 2017; Henderson et al., 2020), however these studies use univalent stimuli. As stated previously, the current study will embed change detection tasks within task switching paradigms. In traditional task switching, participants either repeat performance of the same task (e.g., Task A–Task A) or switch between performance of different tasks (e.g., Task A–Task B). Consistent findings show that when there is a requirement to switch between tasks, performance in terms of response times (RT) and error rate is poorer relative to when the task repeats, referred to as the task switch cost.

In order to perform a given task, one must activate a *task set* in memory. What precisely a task set is varies with researcher (see Grange & Houghton, 2014, pp. 4–6 for discussion on the differing definitions), with some suggesting that a task set is a grouping of processes and representations needed to perform an given task (see e.g., Rogers & Monsell, 1995), while others have suggested a task set contains a set of parameters which are programmed by the cognitive system depending on the task to be performed (see e.g., Logan & Gordon, 2001; Logan & Schneider, 2010). Despite the variance, one assumption is common across most definitions of task set—when the task changes, a new task set must be activated in order to correctly perform the task (however, see Altmann & Gray, 2008 for an alternative view). The structure of a task set can also be broken down into two distinct components, the *stimulus* or *attentional set*, which deals with the identification and selection of information relevant to the task, and the *response* or *intentional set*, which deals with the assignment of appropriate stimulus-response (S-R) mappings (see e.g., Kopp et al., 2006; Liefooghe & Verbruggen, 2019; Meiran, 2000; Rushworth et al., 2005, 2002; Yeung & Monsell, 2003b).

At present, two accounts of switch costs are most prominent, the *interference* and *reconfiguration* accounts. Initial research on the interference view stated that switch costs are believed to arise as the result of *proactive interference*. If a task switch occurs after response on trial *N-1*, a new task set will be required to correctly perform the new task on trial *N*; however, the *task set inertia* account (see e.g., Allport et al., 1994) states that the persistence of the task set from trial *N-1* may interfere with the newly relevant task set on trial *N* if it is not appropriately suppressed or if it has not fully dissipated. In contrast, initial research on the reconfiguration view stated that the switch cost reflects the time course of activating the newly relevant task set; activation of task sets is a time consuming process and as such, on the *task set reconfiguration* account (see e.g., Rogers & Monsell, 1995), increased RTs on switch trials are indicative of this process of activation, while errors are reflective of a failure to activate the task set appropriately.

Independent of the origin of task switch costs—which have also been postulated to be a mixture of both interference and reconfiguration (see e.g., Rubinstein et al., 2001; Vandierendonck et al., 2010)—the act of switching requires some degree of attentional control. Put simply, attentional control refers to a combination of attentional and inhibitory processes, with the former allocating the focus of attention to relevant information and the latter preventing interference from competing, irrelevant information (see e.g., Friedman & Miyake, 2004; Miyake et al., 2000; Yantis, 1998). Therefore, task switching provides a somewhat simplistic way in which to investigate the role of attentional control in vSTM resource allocation, which in the present study is achieved by requiring attendance to different feature dimensions on different trials on a change detection task. As such, only the attentional set will change when a switch in task occurs, with the intentional set remaining the same throughout (i.e., detect changes), enabling a measure of vSTM performance to be obtained under situations where attentional control is required to select the relevant information.

2.1.6 The present study

In the present study, multivalent stimulus displays will be used; in some cases stimuli will be multivalent containing both colour and orientation, while in others, stimuli will be univalent (i.e., either coloured or oriented), however stimulus displays will always consist of both features. Participants will perform a change detection task on each trial, however crucially, the feature dimension relevant on a given trial will either repeat (e.g., colour–colour) or switch (e.g., colour–orientation). Should a cost to performance be observed on dimension switch trials relative to dimension repetition trials in the present set of experiments, this would suggest that attentional control is required for appropriate resource allocation, with switching potentially resulting in a misallocation of resources to the irrelevant feature dimension.

2.2 Experiments 1 and 2: General method

2.2.1 Participants

45 participants completed Experiment 1a, with 57 participants completing Experiment 1b. 48 participants completed Experiment 2a, with 45 participants completing Experiment 2b. As these experiments made up the initial, exploratory set, data was collected from as many participants as possible within a predetermined timeframe. Except for Experiment 1a, a binomial test was conducted on each experiment to identify participants who did not perform significantly greater than chance. Change performance in a change detection task is 0.5 (50%), with performance at this level indicative of guessing. As a result, 13 participants were excluded from data analysis in Experiment 1b (final $N = 44$),

with four participants excluded in Experiment 2a (final $N = 44$). No participants were removed from Experiment 2b. Participants were recruited via Prolific (prolific.co), with unique participants recruited for each experiment. Participants were aged between 18 and 60 years of age (inclusive) and self-reported normal or corrected-to-normal visual acuity and normal colour vision. Recruitment was limited to the United Kingdom (UK) and the United States of America (USA). Participants were paid a small fee for taking part and ethical approval for Experiments 1 and 2 (and all subsequent experiments within this thesis) was obtained from the Psychology Ethics Committee at Keele University.

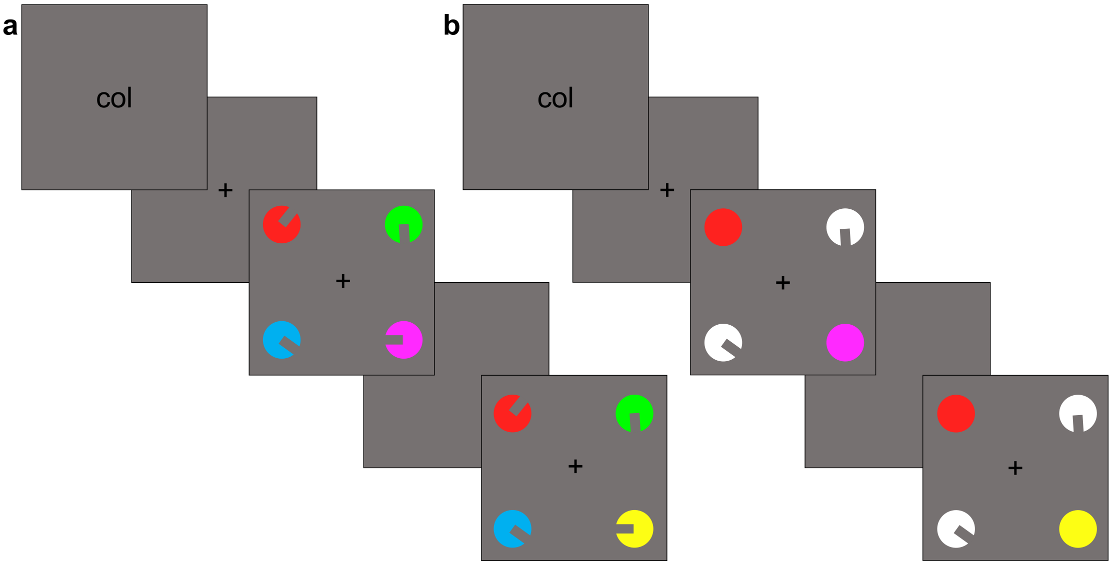

Figure 4

Figure showing stimuli used in Experiments 1 and 2. Note that in Experiments 1b and 2b a change also occurred in the irrelevant feature dimension.

2.2.2 *Stimuli and materials*

Stimuli in Experiment 1 consisted of four bivalent circular shapes which displayed both colour and orientation (see Figure 4a). In Experiment 2, four univalent stimuli were used, two of which were the same circular shapes as those used in Experiment 1 with the exception that they were always white (i.e., colourless), with coloured circles making up the remaining two stimuli (see Figure 4b). In Experiment 2, participants were reminded at the beginning of the task and throughout that white would not be tested as a colour. The RGB (red, green, blue) coordinates for the colours used in these experiments were: black (0, 11, 16), blue (65, 105, 225), cyan (20, 253, 255), green (0, 250, 3), purple (255, 41, 255), red (255, 54, 31), and yellow (253, 254, 21). The RGB coordinates for the white oriented stimuli in Experiment 2 were 255, 255, 255. RGB coordinates for the background were 127, 128, 128, with both cues and fixation crosses in Experiment 1 being white. Given the use of white stimuli in Experiment 2, the colour of the fixation crosses was changed to dark grey (89, 89, 89). Cues consisted of the word “col” if colour was relevant, with the word “ori” presented if orientation was relevant. Stimuli were presented within a dark grey square frame positioned in the centre of each screen. Each experiment was created and run using Gorilla Experiment Builder (Anwyl-Irvine et al., 2020), with restrictions limiting participation in each experiment to those using desktop or laptop computers only.

2.2.3 Procedure

In each experiment, participants made a change or no change judgement in pure blocks wherein the relevant feature dimension repeated (i.e., all colour *or* all orientation trials; *pure repetition*) and mixed blocks, wherein the relevant feature dimension could repeat (e.g., colour–colour; *mixed repetition*), or switch (e.g., colour–orientation; *mixed switch*). Practice blocks in each experiment consisted of a total of 40 trials (10 pure orientation, 10 pure colour, 20 mixed), with a total of 450 trials making up the main experimental section (75 pure orientation, 75 pure colour, 300 mixed). These were separated into cycles following the same ordering pattern of 25 pure orientation, 25 pure colour, and two blocks of 50 mixed trials. Self-paced breaks were given every 50 trials.

At the beginning of each trial, participants were presented with a fixation cross for 750ms, followed by a 500ms cue indicating which feature dimension was relevant for the current trial. If colour was relevant, the word “col” was presented, while the word “ori” was presented if orientation was relevant. A cue-stimulus interval (CSI) of 500ms followed, with subsequent presentation of the memory display for 200ms. A retention interval of 1,500ms was followed by presentation of the test display for 200ms. Responses were provided by pressing one of two keys on a keyboard; pressing “M” indicated a change response, while pressing “Z” indicated a no change response. On change trials in Experiments 1a and 2a, a change occurred in the relevant feature dimension only, whereas in Experiment 1b and 2b, a change occurred in the irrelevant feature dimension on every trial. Participants were instructed to respond as quickly and as accurately as possible, with accuracy and response time (RT) recorded on each trial. Responses were not time-limited and were followed by a 250ms intertrial interval (ITI).

2.2.4 Statistical analysis

Bayesian analysis of variance (ANOVA) tests were used to examine the effect of sequencing condition (pure repetition, mixed repetition, mixed switch) on change detection accuracy, RT, and two measures from detection theory, d' (“*dee-prime*”), a measure of sensitivity defined in terms of the inverse of the normal distribution function z (Macmillan & Creelman, 2004), with h representing hits and f representing false alarms:

$$d' = z(h) - z(f), \tag{4}$$

and *criterion* (c), a measure of response bias:

$$c = \frac{1}{2}[z(h) + z(f)] \tag{5}$$

Detection theory measures were implemented using the R pacakge *sdtt* (see Moore, 2022). Bayesian paired-samples t-tests were used if the Bayes Factor (BF_{10}) was 10 or greater which indicates strong evidence in support of the alternative hypothesis; interpretation of BFs follows that outlined by Lee and Wagenmakers (2013). All Bayesian analysis was conducted using the R package *BayesFactor* (see Morey & Rouder, 2021) using package default priors, namely a Cauchy distribution centred on zero effect with scale parameter $r = 0.707$. Prior to analysis, the first trial in each block was removed as this is neither a repetition or switch trial. Additionally, the trial immediately following an error was removed prior to analysis of accuracy, sensitivity, and response bias data; as the nature of an error cannot be determined, this can make the sequencing condition of the subsequent trial inaccurate, thus having the potential to produce unreliable results. Prior to RT analysis, RTs less than 150ms and 2.5 standard deviations (SDs) above the mean were trimmed per participant, per condition using the R package *trimr* (see Grange, 2022). Error trials and trials immediately following an error were also removed prior to RT analysis due to the potential for post-error slowing. All analyses were conducted using R[6]. Anonymised data and analysis code for each experiment and all experiments within this thesis can be found on the Open Science Framework at the following link:

https://osf.io/dnqxz/?view_only=e68968defbcc471dbcd1b1561ec515f6.

2.3 Experiment 1a

Experiment 1a served as the initial investigation into the impact of dimension switching on change detection performance. While some previous studies have utilised multivalent stimuli (see e.g., Luck & Vogel, 1997), these studies did not systematically investigate the effect of switching between detecting changes in different feature dimensions. As such, the role of atttentional control with respect to change detection performance and vSTM resource allocation more broadly remains unclear. The set size of four used in the present experiment—and in many subsequent experiments—is based on previous research suggesting vSTM has a capacity limit of approximately four items (see e.g., Luck & Vogel, 1997; Vogel et al., 2001; Zhang & Luck, 2008). Therefore, should a disruptive impact to performance be observed in these experiments, this cannot be attributed to exceeding this supposed upper limit.

---

[6] Analyses in all experiments within this thesis used R (Version 4.2.0; R Core Team, 2022) and the R-packages *afex* (Version 1.1.1; Singmann et al., 2022), *BayesFactor* (Version 0.9.12.4.3; Morey & Rouder, 2021), *coda* (Version 0.19.4; Plummer et al., 2006), *cowplot* (Version 1.1.1; Wilke, 2020), *dplyr* (Version 1.0.9; Wickham, François, et al., 2022), *effsize* (Version 0.8.1; Torchiano, 2020), *forcats* (Version 0.5.1; Wickham, 2021), *ggplot2* (Version 3.3.6; Wickham, 2016), *ggpubr* (Version 0.4.0; Kassambara, 2020), *lme4* (Version 1.1.29; Bates et al., 2015), *magick* (Version 2.7.3; Ooms, 2021), *Matrix* (Version 1.4.1; Bates et al., 2022), *mixtur* (Version 1.2.0; Grange & Moore, 2021), *papaja* (Version 0.1.1; Aust & Barth, 2020), *plyr* (Wickham, 2011; Version 1.8.7; Wickham, François, et al., 2022), *purrr* (Version 0.3.4; Henry & Wickham, 2020), *readr* (Version 2.1.2; Wickham, Hester, et al., 2022), *sdtt* (Version 0.2.0; Moore, 2022), *stringr* (Version 1.4.0; Wickham, 2019), *tibble* (Version 3.1.7; Müller & Wickham, 2022), *tidyr* (Version 1.2.0; Wickham & Girlich, 2022), *tidyverse* (Version 1.3.1; Wickham et al., 2019), *tinylabels* (Version 0.2.3; Barth, 2022), and *trimr* (Version 1.1.1; Grange, 2022).

2.3.1 Results

Means for each of the measures used in Experiments 1a and 1b can be found in Figure 5. Analysis of accuracy data revealed a $\mathrm{BF}_{10} = 0.14$ indicating moderate evidence for the absence of an effect of sequencing condition. Analysis of RT data revealed a $\mathrm{BF}_{10} = 1.43$ indicating anecdotal evidence for the presence of an effect of sequencing condition. Analysis of d' data revealed a $\mathrm{BF}_{10} = 0.14$ indicating moderate evidence for the absence of an effect of sequencing condition. Finally, analysis of c data revealed a $\mathrm{BF}_{10} = 0.27$ indicating moderate evidence for the absence of an effect of sequencing condition. Note that the log-linear correction (see Hautus, 1995) was used to calculate hit and false alarm rate due to extreme values.

2.3.2 Discussion

Clearly, the results from Experiment 1a suggest that sequencing condition had no impact on change detection performance. However, it is possible that a methodological oversight may have contributed to the overwhelming absence of switch costs on this task. On change trials in this experiment, a change only occurred in the relevant feature dimension; as such, participants would have been able to attend to the entire stimulus display (i.e., both colour and orientation) rather than selectively attending to the relevant feature dimension. Such a *whole-display monitoring* technique renders the cue non-essential for correct task performance and does not require updating of attentional set as the task set would remain the same on each trial (i.e., monitor the entire display and detect any change), negating any potential switch costs that may have arisen. Therefore, the aim of Experiment 1b was to rectify this oversight by making the cue essential for correct performance of the task.

2.4 Experiment 1b

Experiment 1b employed the same methodology as Experiment 1a with the exception that on every trial a change occurred in the irrelevant feature dimension. Having a change in both relevant and irrelevant feature dimensions means that adoption of the whole-display monitoring technique—which may have contributed to the results observed in Experiment 1a—is less effective. Rather, in the present experiment, it would be more beneficial to performance to attend to only the cued feature dimension. While some research has suggested that performance on change detection tasks can be negatively impacted by a change in the irrelevant feature dimension (i.e., *the irrelevant change effect*, see e.g., Jaswal & Logie, 2011), a number of studies have demonstrated that this effect appears to present primarily at shorter retention intervals Bocincova et al. (2017). Based on this research, the 1,500ms retention interval used in the present experiment means that a change in the irrelevant feature dimension should not be as disruptive—or at least be minimally disruptive—to performance.

2.4.1 Results

Analysis of accuracy data revealed a BF_{10} = 1,816.37 indicating extreme evidence in support of an effect of sequencing condition. A Bayesian paired-samples t-test between pure repetition (M = 0.80, SD = 0.08) and mixed repetition (M = 0.78, SD = 0.09) data revealed a BF_{10} = 0.49 indicating anecdotal evidence in support of the absence of a mixing cost. A further Bayesian paired-samples t-test between mixed repetition and mixed switch (M = 0.75, SD = 0.08) data revealed a BF_{10} = 51.22 indicating very strong evidence for the presence of a switch cost. The relationship between pure repetition and mixed switch data is theoretically uninformative therefore no analysis was conducted between these conditions. Analysis of RT data revealed a $\text{BF}_{10} = 49.94$ indicating very strong evidence in support of an effect of sequencing condition. A Bayesian paired-samples t-test between pure repetition (M = 958.38, SD = 193.04) and mixed repetition (M = 979.58, SD = 191.33) data revealed a $\text{BF}_{10} = 0.38$ indicating anecdotal evidence in support of the absence of a mixing cost. Further analysis between mixed repetition and mixed switch (M = 1,023.58, SD = 209.30) data revealed a $\text{BF}_{10} = 24.46$ indicating strong evidence in support of a switch cost.

Analysis of d' data revealed a $\text{BF}_{10} = 3{,}196.53$ indicating extreme evidence in support of an effect of sequencing condition. A Bayesian paired-samples t-test between pure repetition (M = 1.79, SD = 0.56) and mixed repetition (M = 1.71, SD = 0.64) data revealed a $\text{BF}_{10} = 0.35$ indicating anecdotal evidence in support of the absence of a mixing cost. A further Bayesian paired-samples t-test between mixed repetition and mixed switch (M = 1.46, SD = 0.52) data revealed a $\text{BF}_{10} = 46.61$ indicating very strong evidence in support of a switch cost. Analysis of c data revealed a $\text{BF}_{10} = 3.95$ indicating moderate evidence for the presence of an effect of sequencing condition. Note that the log-linear correction (see Hautus, 1995) was used to calculate hit and false alarm rate due to extreme values.

2.4.2 Discussion

Contrasting the results of Experiment 1a, the results of Experiment 1b appear to show that switching between different feature dimensions does result in a cost to change detection performance (i.e., a *dimension switch cost*). While both accuracy and RT showed reduced performance on mixed switch relative to mixed repetition trials, perhaps the most notable result was the finding that sensitivity was impacted by dimension switching. This suggests that on switch trials, participants were less sensitive to a change in the relevant feature dimension, perhaps due to the need to update attentional set from that previously used or due to interference from the previously active attentional set. Finally, while there is evidence to suggest dimension switching has little effect on response bias, it is perhaps interesting to note that bias became more liberal across sequencing conditions as can be seen in Figure 5d. This may be indicative of an increased failure to attend to the relevant feature dimension during mixed blocks of trials, resulting in performance of the task on the wrong feature dimension.

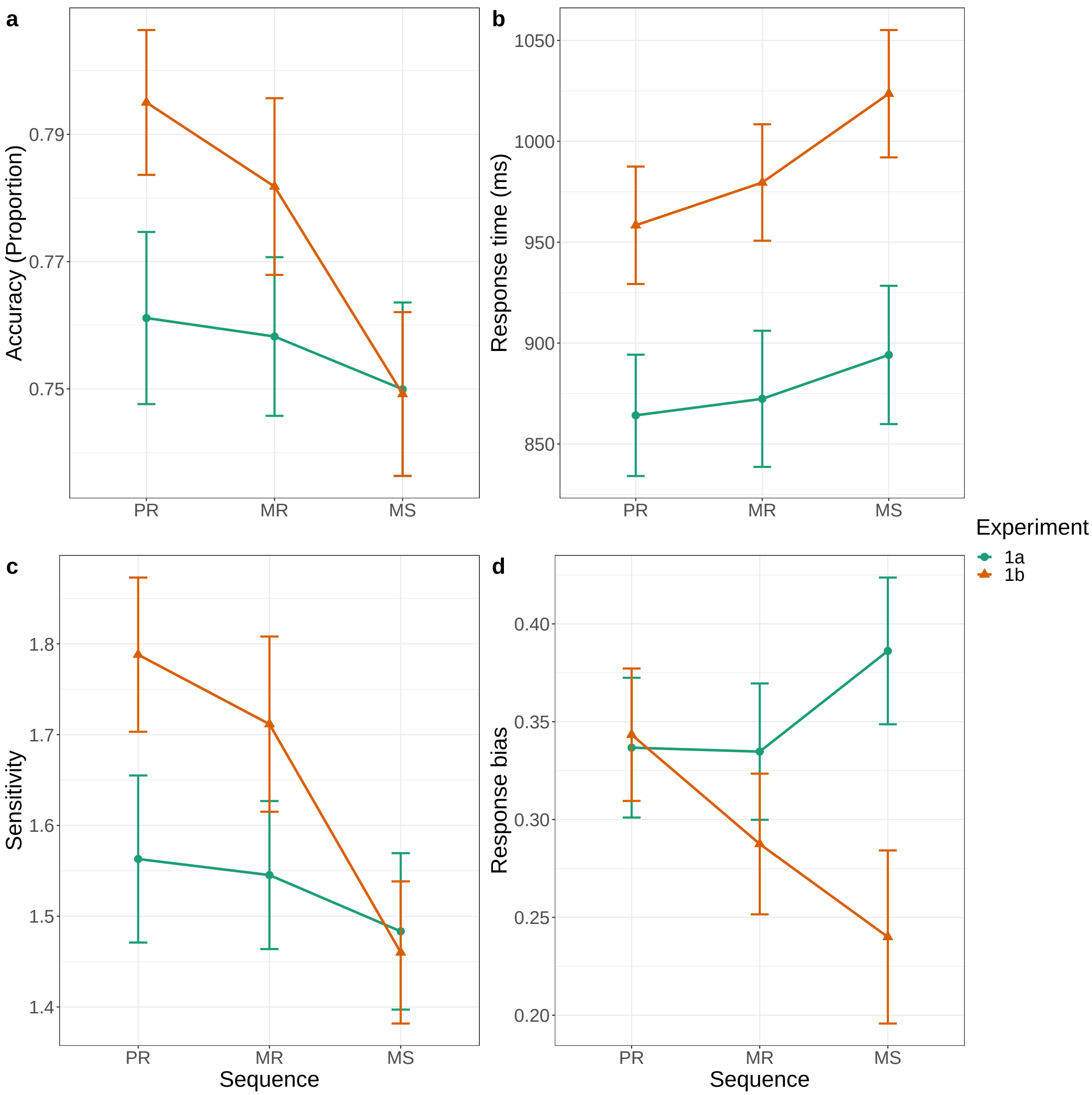

Figure 5

Plots displaying mean a) accuracy, b) response time, c) sensitivity (d′), and d) response bias (c) across each sequencing condition in Experiment 1 (PR = pure repetition, MR = mixed repetition, MS = mixed switch). Error bars represent the standard error of the mean.

2.5 Experiment 2a

While Experiment 1 provided an initial look at the impact of dimension switching on change detection performance with multivalent stimuli, Experiment 2 sought to investigate the impact of dimension switching with univalent stimuli. On a given trial within Experiment 2, stimulus displays consisted of two of the circular shapes in Experiment 1, however these were always white (i.e., colourless), with the remaining two stimuli being coloured circles. Previous research concerning resource models of vSTM suggest that irrelevant features of an item are encoded automatically when attending to a relevant feature on the same item (see Marshall & Bays, 2013), with such encoding utilising a portion of the available memory resource. Based on this research, attending to univalent stimuli—such as those used in Experiment 2—will use less resource relative to attending multivalent stimuli—such as those used in Experiment 1. This suggests that performance on the present task will be markedly better than that observed in Experiment 1 as no obligatory encoding of irrelevant features will occur (although this is not tested statistically here). In keeping with the methodology employed in Experiment 1, changes in Experiment 2a only occur in the relevant feature dimension.

2.5.1 Results

Analysis of accuracy data revealed a $\mathrm{BF}_{10} = 6.48$ indicating moderate evidence in support of an effect of sequencing condition. Analysis of RT data revealed a $\mathrm{BF}_{10} = 0.17$ revealing moderate evidence in support of the absence of an effect of sequencing condition. Analysis of d' data revealed a $\mathrm{BF}_{10} = 0.72$ indicating anecdotal evidence in support of the absence of an effect of sequencing condition. Finally, analysis of c data revealed a $\mathrm{BF}_{10} =$

0.63 indicating anecdotal evidence in support of the absence of an effect of sequencing condition. Note that the log-linear correction (see Hautus, 1995) was used to calculate hit and false alarm rate due to extreme values.

2.5.2 Discussion

Means for each of the measures used in Experiments 2a and 2b can be found in Figure 6. The results of Experiment 2a reveal a similar pattern in that switching between feature dimensions does not impact change detection performance. Most notably, a low degree of evidence was observed for an effect of sequencing condition in accuracy, although from Figure 6a, it appears that this effect is likely due to the difference between pure repetition and mixed switch trials, a theoretically uninformative relationship. The general lack of evidence for an effect of sequencing condition may be again be due to the cue not being essential for correct performance of the task, with participants potentially attending to the whole display. Indeed, this would be a much easier technique to adopt in Experiment 2a than in Experiment 1a as only four features were present on four stimuli in Experiment 2a relative to the eight total features on four stimuli in Experiment 1a.

2.6 Experiment 2b

Experiment 2b used the same methodology as Experiment 2a with the exception that a change in the irrelevant feature dimension occurred on every trial. This ensured that the cue was essential for correct performance of the task.

2.6.1 Results

Analysis of accuracy data revealed a $BF_{10} = 1.69$ indicating anecdotal evidence in support of an effect of sequencing condition. Analysis of RT data revealed a $BF_{10} = 278.31$

indicating extreme evidence in support of an effect of sequencing condition. A Bayesian paired-samples t-test between pure repetition (M = 882.52, SD = 220.08) and mixed repetition (M = 855.19, SD = 219.49) data revealed a BF_{10} = 5.87 indicating moderate evidence in support of a *mixing benefit*. A further Bayesian paired-samples t-test between mixed repetition and mixed switch (M = 902.28, SD = 245.58) data revealed a BF_{10} = 5,858.72 indicating extreme evidence in support of a switch cost. Analysis of d' data revealed a BF_{10} = 0.24 indicating moderate evidence in support of the absence of an effect of sequencing condition. Finally, analysis of c data revealed a BF_{10} = 133,214.05 indicating extreme evidence in support of an effect of sequencing condition. A Bayesian paired-samples t-test between pure repetition (M = 0.20, SD = 0.22) and mixed repetition (M = 0.05, SD = 0.18) data revealed a BF_{10} = 111.19 indicating extreme evidence in support of a mixing cost. A further Bayesian paired-samples t-test between mixed repetition and mixed switch (M = -0.02, SD = 0.23) data revealed a BF_{10} = 1.17 indicating anecdotal evidence in support of a switch cost. Note that the log-linear correction (see Hautus, 1995) was used to calculate hit and false alarm rate due to extreme values.

2.6.2 Discussion

Results from Experiment 2b present a somewhat different pattern of results from those observed in Experiment 1b. While RT showed a high level of evidence for the presence of an effect of sequencing condition, RT was faster on mixed repetition trials relative to mixed switch trials resulting in a mixing benefit rather than a mixing cost. As such, this faster performance on mixed repetition trials may have contributed to the high level of evidence supporting the presence of a switch cost. Perhaps a further notable observation from this study is the fact that similar to Experiment 1b, mixed blocks caused response bias to become more liberal. This suggests that while switching between feature dimensions did not have any impact on bias, the rate at which participants fail to attend to the relevant feature dimension increases on mixed blocks of trials.

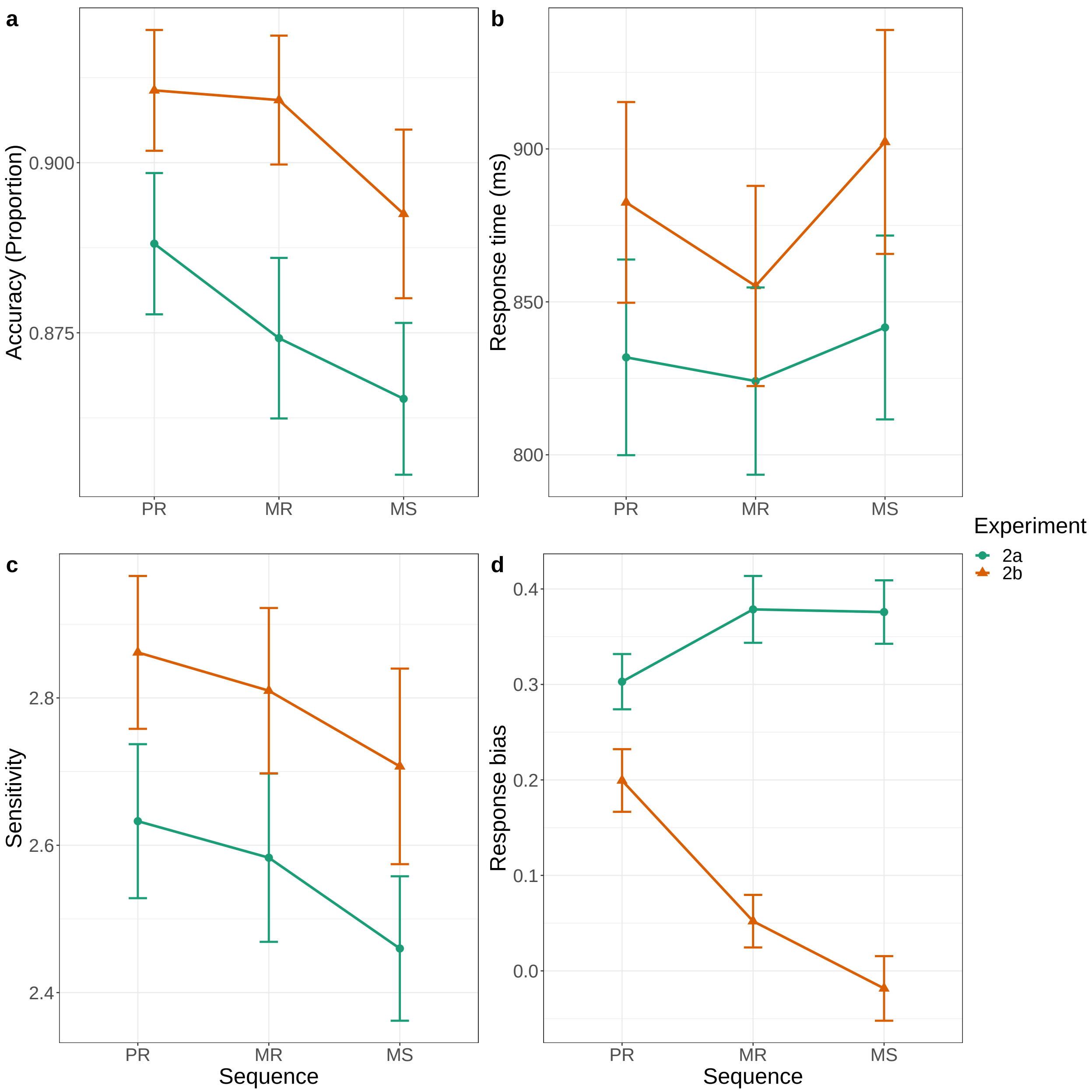

Figure 6

Plots displaying mean a) accuracy, b) response time, c) sensitivity (d'), and d) response bias (c) across each sequencing condition in Experiment 2 (PR = pure repetition, MR = mixed repetition, MS = mixed switch). Error bars represent the standard error of the mean.

2.7 Experiment 3 : General method

While Experiments 1 and 2 provided insight into the impact of dimension switching on change detection performance, Experiment 3 aimed to provide a more nuanced examination of switch costs, as well as examine performance under other experimental manipulations. As such, in Experiment 3, pure repetition trials were removed meaning more mixed trials could be included. Experiments 3a and 3b were similar to Experiments 1b and 2b respectively, with the exception that stimuli were isosceles triangles, being coloured and oriented in Experiment 3a and coloured or oriented in Experiment 3b. Experiment 3c simply doubled the total number of univalent stimuli from Experiment 3b from four to eight, while in Experiment 3d coloured circles were overlaid onto white (i.e., colourless) oriented isosceles triangles. These experiments are based on previous work by Markov et al. (2019) who investigated feature vs. object-based storage in vSTM.

2.7.1 Participants

A sequential Bayes Factor design as outlined by Schönbrodt and Wagenmakers (2018) was implemented in Experiment 3 (see Sequential Bayes Factor analysis for more detail). This resulted in a total of 47 participants completing Experiment 3a, 73 in Experiment 3b, 72 in Experiment 3c, and 75 in Experiment 3d. A binomial test to identify participants who did not perform significantly greater than chance (i.e., 50%) resulted in the exclusion of two participants from Experiment 3a (final $N = 45$), three participants from Experiment 3b (final $N = 70$), two participants from Experiment 3c (final $N = 70$), and five participants from Experiment 3d (final $N = 70$). Participants were recruited via a combination of the SONA Participant Recruitment System at Keele University and Prolific (prolific.co). Participants were aged between 18 and 60 years (inclusive) and reported normal or corrected-to-normal visual acuity and normal colour vision. Unique participants were recruited for each experiment, with those who had taken part in Experiments 1 and 2 prevented from participating in Experiment 3. Recruitment was limited to the UK and USA and participants were paid a small fee for taking part (Prolific) or awarded course credit (SONA).

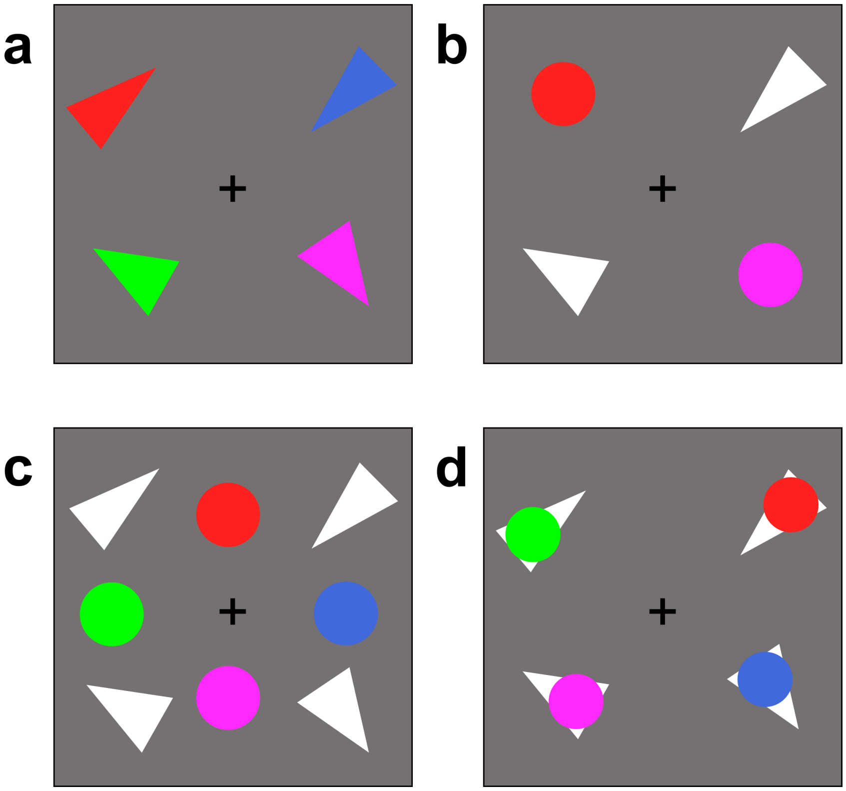

Figure 7
Figure showing stimuli used in Experiment 3.

2.7.2 Stimuli and materials

Stimuli in Experiment 3a consisted of four isosceles triangles containing both colour and orientation. In Experiment 4b, four univalent stimuli were used, two of which were the same isosceles triangles used in Experiment 3a with the exception that they were always white (i.e., colourless), with two coloured circles making up the remaining stimuli. Stimuli in

Experiment 3c were the same as that in Experiment 3b just with a doubling of set size (i.e., four colourless isosceles triangles and four coloured circles). Finally, stimuli in Experiment 3d consisted of four colourless isosceles triangles with four coloured circles overlaid (see Figure 7). The RGB coordinates for the colours used in Experiment 3 were identical to those used in Experiments 1 and 2, with the fixation cross presented in dark grey (89, 89, 89) across all experiments. Each experiment was created and run using Gorilla Experiment Builder (Anwyl-Irvine et al., 2020), with restrictions limiting participation in each experiment to those using desktop or laptop computers only.

2.7.3 Procedure

The procedure for Experiment 3 was the same as the procedure for Experiments 1 and 2 with the exception that only mixed blocks of trials were used and that a constant change in the irrelevant feature dimension was employed in all experiments. Thus, participants made a change or no change judgement in blocks wherein the feature dimension could repeat (e.g.., colour–colour; mixed repetition) or switch (e.g., colour–orientation; mixed switch). Practice blocks again contained 40 trials (10 pure orientation, 10 pure colour, 20 mixed) which allowed participants to familiarise themselves with the task. The main experimental section consisted of a total of 400 trials, separated into blocks of 50. Self-paced breaks were given every 50 trials.

2.7.4 Sequential Bayes Factor analysis

A sequential Bayes Factor analysis with maximal sample size was implemented within each experiment. A minimum sample size of 20 was required for each experiment with a maximum of 70. The critical test was a Bayesian paired-samples t-test conducted on mixed repetition and mixed switch accuracy values. Once data had been collected for the minimum sample size, the test was performed; if the BF_{10} was greater than 10 or less than 0.1, indicating strong evidence in support of the alternative and null hypotheses respectively (see Lee & Wagenmakers, 2013), data collection was terminated. If neither limit was reached, data collection continued, with the critical test performed every five participants until the maximum sample size was reached at which point data collection was terminated.

2.7.5 Statistical analysis

Bayesian paired-samples t-tests were used to analyse performance between mixed repetition and mixed switch data across all measures. As in Experiments 1 and 2, all Bayesian analysis was conducted using the R package *BayesFactor* (see Morey & Rouder, 2021), using package default priors (a Cauchy distribution centred on zero effect with scale parameter $r = 0.707$). All remaining data analysis (i.e., calculation of detection theory measures) and data preparation (i.e., error removal, RT trimming) was identical to that used in Experiments 1 and 2.

2.8 Experiment 3a

Experiment 3a sought to confirm the findings from Experiment 1b that dimension switching impacts change detection performance. The same change detection switching task was used, however the stimuli in Experiment 3a were coloured, oriented isosceles triangles. By using different stimuli here, should switch costs be observed, this would demonstrate that this effect can occur across a range of stimulus types.

2.8.1 Results

A Bayesian paired-samples t-test on accuracy between mixed repetition (M = 0.76, SD = 0.08) and mixed switch (M = 0.74, SD = 0.08) data revealed a BF_{10} = 31.85 indicating very strong evidence in support of the presence of a switch cost. Analysis of RT between mixed repetition (M = 1,036.01, SD = 267.64) and mixed switch (M = 1,104.30, SD = 312.91) data revealed a BF_{10} = 162.06 indicating extreme evidence in support of the presence of a switch cost. Analysis of sensitivity between mixed repetition (M = 1.58, SD = 0.54) and mixed switch (M = 1.37, SD = 0.56) data revealed a BF_{10} = 85.07 indicating very strong evidence in support of the presence of a switch cost. Finally, analysis of response bias between mixed repetition (M = 0.29, SD = 0.54) and mixed switch (M = 0.22, SD = 0.26) data revealed a BF_{10} = 3.21 indicating moderate evidence in support of the presence of a switch cost.

2.8.2 Discussion

Results from Experiment 3a are very similar to those observed in Experiment 1b, with dimension switching impacting performance primarily on accuracy, RT, and sensitivity measures. Therefore, it appears that switching between feature dimensions when stimuli are bivalent results in a cost to performance.

2.9 Experiment 3b

Experiment 3b aimed to confirm the findings of Experiment 2b using univalent stimuli, with two white (i.e., colourless) isosceles trianges and two coloured circles.

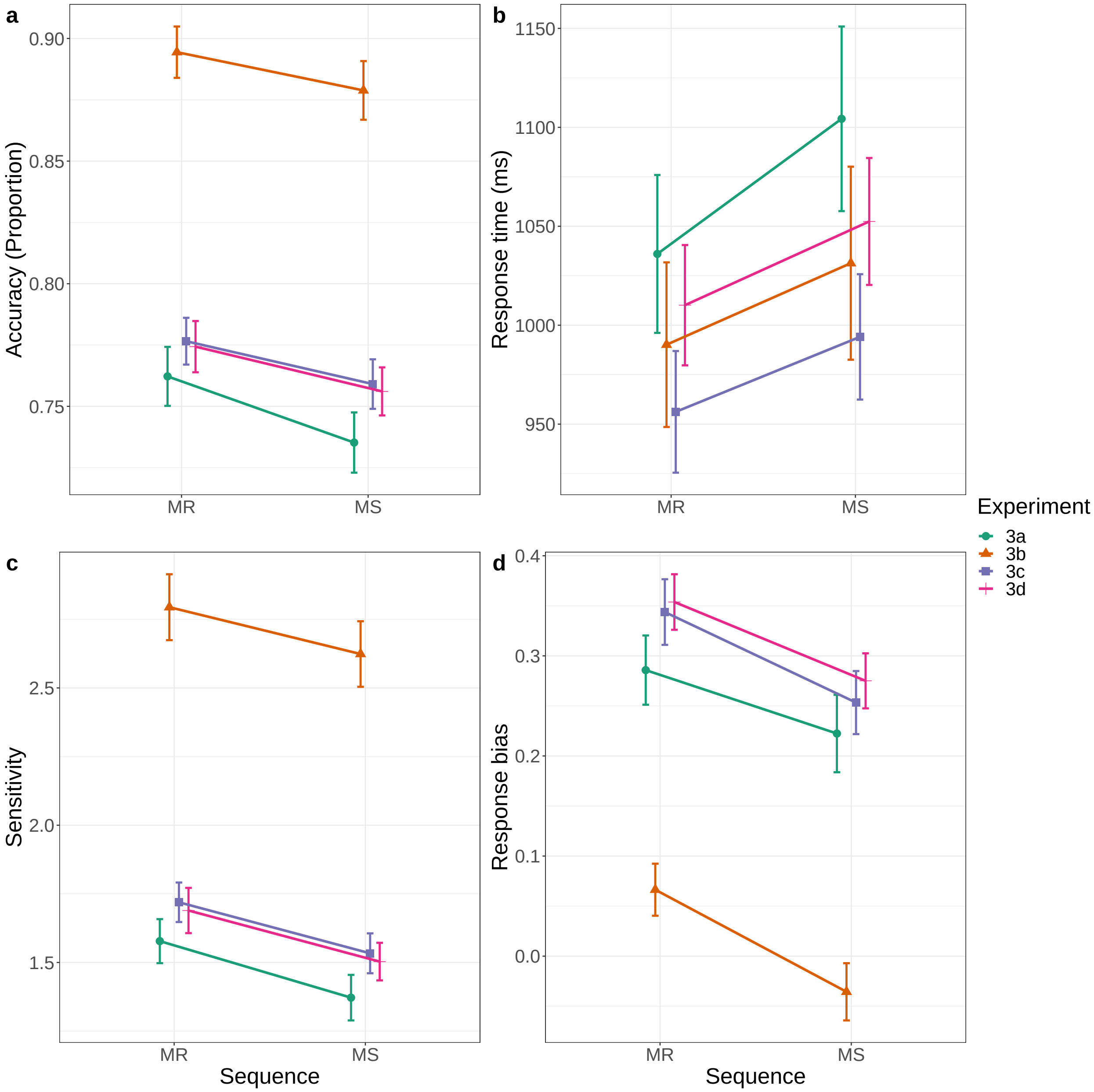

Figure 8

Plots displaying mean a) accuracy, b) response time, c) sensitivity (d′), and d) response bias (c) *across each sequencing condition in Experiment 3 (MR = mixed repetition, MS = mixed switch). Error bars represent the standard error of the mean.*

2.9.1 Results

A Bayesian paired-samples t-test on accuracy between mixed repeititon (M = 0.89, SD = 0.09) and mixed switch (M = 0.88, SD = 0.10) data revealed a BF_{10} = 11.05 indicating strong evidence in support of the presence of a switch cost. Analysis of RT between mixed repetition (M = 1,036.01, SD = 267.64) and mixed switch (M = 1,104.30, SD = 312.91) revealed a BF_{10} = 8.10 indicating moderate evidence in support of the presence of a switch cost. Analysis of sensitivity between mixed repetition (M = 2.79, SD = 1.00) and mixed switch (M = 2.62, SD = 1.00) data revealed a BF_{10} = 4.10 indicating moderate evidence in support of the presence of a switch cost. Finally, analysis of response bias between mixed repetition (M = 0.07, SD = 0.22) and mixed switch (M = -0.04, SD = 0.24) data revealed a BF_{10} = 14.88 indicating strong evidence in support of the presence of a switch cost. Note that the log-linear correction (see Hautus, 1995) was used to calculate hit and false alarm rate due to extreme values.

2.9.2 Discussion

Overall, the results from Experiment 3b appear to deviate somewhat from those observed in Experiment 2b, however such a statement should be taken with a degree of caution. The evidence in Experiment 2b suggested a lack of an effect of sequencing condition on accuracy, whereas a high level of evidence in support of a switch cost was found for accuracy in the present experiment. Additionally, the finding of a mixing benefit for RT in Experiment 2b may have contributed to the high level of evidence in support of a switch cost, whereas the level of evidence supporting a switch cost in Experiment 3b was not nearly as high. Finally, while the low level of evidence for a switch cost in sensitivity somewhat follows the general lack of evidence in Experiment 2b, strong evidence for a switch cost in response bias was found in Experiment 3b relative to the lack of evidence for such an effect in Experiment 2b. This result shows that response bias became more liberal on mixed switch relative to mixed repetition trials which is perhaps indicative of an increased failure to attend to the relevant feature dimension on switch trials; as a change in the irrelevant feature dimension occurred on each trial in Experiment 3b, it follows that more change responses would be made if attending to the irrelevant feature dimension.

2.10 Experiment 3c

Experiment 3c aimed to extend the findings from Experiment 3b by doubling the number of stimuli. As such, four white (i.e., colourless) isosceles triangles were used along with four coloured circles. This resulted in a total set size of eight. Thus, while only two stimuli were relevant on a given trial in Experiment 3b, four stimuli were relevant on each trial within Experiment 3c.

2.10.1 Results

A Bayesian paired-samples t-test on accuracy between mixed repetition ($M = 0.78$, $SD = 0.08$) and mixed switch ($M = 0.76$, $SD = 0.08$) data revealed a $\mathrm{BF}_{10} = 4.59$ indicating moderate evidence in support of the presence of a switch cost. Analysis of RT between mixed repetition ($M = 956.21$, $SD = 257.15$) and mixed switch ($M = 994.10$, $SD = 265.08$) data revealed a $\mathrm{BF}_{10} = 18{,}845.37$ indicating extreme evidence in support of the presence of a switch cost. Analysis of sensitivity between mixed repetition ($M = 1.72$, $SD = 0.60$) and mixed switch ($M = 1.53$, $SD = 0.61$) data revealed a $\mathrm{BF}_{10} = 509.30$ indicating extreme evidence in support of the presence of a switch cost. Finally, analysis of response bias between mixed repetition ($M = 0.34$, $SD = 0.27$) and mixed switch ($M = 0.25$, $SD = 0.26$) data revealed a $\mathrm{BF}_{10} = 40.81$ indicating very strong evidence in support of the presence of a switch cost. Note that the log-linear correction (see Hautus, 1995) was used to calculate hit and false alarm rate due to extreme values.

2.10.2 Discussion

Results from Experiment 3c show that dimension switching greatly impacted performance across most measures; while little evidence was observed for a switch cost in accuracy, high levels of evidence were observed across RT, sensitivity, and response bias. Thus, it appears that switching was more detrimental in this experiment relative to Experiment 3b potentially due to the doubling of set size despite the fact that only four stimuli were relevant on a given trial. Given that substantial evidence in support of a switch cost was observed for both sensitivity and response bias measures, this suggests that dimension switching resulted in participants becoming less able to detect changes (in the relevant feature dimension) on switch trials, as well as attending to the irrelevant feature dimension more often on switch trials relative to repetition trials.

2.11 Experiment 3d

Experiment 3d again attempted to extend findings from the previous experiments by using spatially overlapping stimuli. This consisted of four coloured circles overlaid onto four white (i.e., colourless) isosceles triangles. The purpose of such a manipulation was so that the same number of features (eight) were presented as in Experiment 3c, however, based on the Gestalt principles of grouping, it would be expected that these spatially overlapping stimuli would be perceived as complex objects made up of multiple features.

2.11.1 Results

A Bayesian paired-samples t-test on accuracy between mixed repetition (M = 0.77, SD = 0.09) and mixed switch (M = 0.76, SD = 0.08) data revealed a BF_{10} = 4.48 indicating moderate evidence in support of the presence of a switch cost. Analysis of RT between mixed repetition (M = 1,010.11, SD = 254.34) and mixed switch (M = 1,052.42, SD = 268.36) data revealed a BF_{10} = 3,310.76 indicating extreme evidence in support of the presence of a switch cost. Analysis of sensitivity between mixed repetition (M = 1.69, SD = 0.69) and mixed switch (M = 1.50, SD = 0.57) data revealed a BF_{10} = 71.23 indicating very strong evidence in support of the presence of a switch cost. Finally, analysis of response bias between mixed repetition (M = 0.35, SD = 0.23) and mixed switch (M = 0.28, SD = 0.23) data revealed a BF_{10} = 142.00 indicating extreme evidence in support of the presence of a switch cost. Note that the log-linear correction (see Hautus, 1995) was used to calculate hit and false alarm rate due to extreme values.

2.11.2 Discussion

Interestingly, the results of Experiment 3d are very similar to the results from Experiment 3c. Again little evidence in support of a switch cost was observed for accuracy, however high levels of evidence in support of a switch cost was found for all other measures, suggesting that on switch trials, participants became less able to detect changes in the relevant feature dimension and attended to the irrelevant feature dimension more regularly. It could be suggested that this is indicative of participants viewing the spatially overlapping stimuli used in Experiment 3d as separate items, more akin to the set size of eight univalent stimuli used in Experiment 3c rather than the set size of four bivalent stimuli used in

Experiment 3a. Indeed, the finding that RT was faster overall in Experiment 3d relative to Experiment 3a may lead to the suggestion that participants were able to more effectively suppress information pertaining to the irrelevant feature dimension.

2.12 General discussion

2.12.1 Summary of results

The purpose of the present study was to provide an initial look at how attentional control may be used to allocate vSTM resources to relevant information. This was achieved by introducing an element of cognitive control into a task that has been extensively used in investigations of vSTM capacity and performance, namely the change detection task. However, while typical change detection tasks only assess performance on a single feature dimension, here the task was embedded within a task switching paradigm wherein participants performed the same change detection task or the same or different feature dimensions, thus introducing an element of attentional control to change detection.

In Experiments 1 and 2 it was shown that disruption on this task only occurred when the cue was essential for correct performance of the task. In both Experiments 1a and 2a, no substantial impact to performance was observed which would suggest that dimension switching does not impact change detection performance. However, given that a change only occurred in the relevant feature dimension, participants could simply monitor the entire display for any change; this would therefore negate the requirement to switch between relevant feature dimensions. As such, a constant change in the irrelevant feature dimension in Experiments 1b and 2b made the cue essential for correct performance of the task, with sequencing condition having an effect in most measures across these experiments. Despite this, some results from Experiment 2b should be interpreted with caution. For instance, the mixing benefit observed for RT is difficult to account for and may have contributed to the substantial evidence observed in support of a switch cost. Interestingly however, from Figure 6 it can be seen that in Experiment 2a, a minor benefit to performance on mixed repetition relative to mixed switch trials was also observed, suggesting that this benefit was increased by making the cue essential for correct task performance.

Experiment 3 sought to both confirm and extend findings from Experiments 1 and 2 by removing pure repetition trials and increasing the number of mixed trials. The methodology employed in Experiments 3a and 3b were essentially the same as that used in Experiments 1b and 2b respectively, with the exception that instead of coloured, oriented circular shapes, the stimuli in Experiment 3 were isosceles triangles. Results from Experiment 3a seemingly confirmed findings from Experiment 1b, with dimension switch costs observed for accuracy, RT, and sensitivity, while little evidence of such an effect was observed for response bias. In relation to Experiment 3b, some results deviated from those observed in Experiment 2b, perhaps notably an increased level of evidence supporting the presence of a dimension switch cost for accuracy and reduced evidence in support of a dimension switch cost for RT. Thus, it is difficult to state that this experiment confirmed findings from Experiment 2b, however, it does appear that dimension switching does impact change detection performance with univalent stimuli.

Extending these results, Experiment 3c doubled the number of univalent stimuli in Experiment 3b from four to eight while Experiment 3d used spatially overlapping stimuli, with coloured circles overlying white isosceles triangles. Perhaps somewhat interestingly, the results of these experiments were similar; while little evidence in support of the presence of a switch cost for accuracy was observed in both experiments, substantial levels of evidence for such a cost were found across all other measures. In terms of Experiment 3c, this may be simply due to the increased number of stimuli, whereas in Experiment 3d, it may be the case that participants viewed the spatially overlapping stimuli as distinct items.

2.12.2 Dimension switch costs as a result of resource misallocation

While the present set of experiments only provides an initial, behavioural view concerning the impact of dimension switching on change detection performance, the finding that switching between different feature dimensions is disruptive across various measures and experimental manipulations suggests that such switching impacts some aspect of trial processing. The dimension switch costs observed in some of the experiment detailed within this chapter are indeed akin to the task switch cost found consistently within the task switching literature (see e.g., Hsieh, 2012) and indeed, are extremely similar to costs found in research wherein switches of attentional set were required (see e.g., Meiran & Marciano, 2002). Despite the critical switching element in the present study being feature dimension and not task, the theories outlined to account for switch costs could be co-opted to account for the disruption by dimension switching observed here.

While definitions of task set are variable across researchers, the requirement for configuration of cognitive processes and representations relevant to a given task (i.e., the task set) is consistent (see e.g., Logan & Gordon, 2001; Mayr & Kliegl, 2000; Meiran, 2010; D. W. Schneider & Logan, 2014). The task set underlies goal-directed behaviour, facilitating not only the selection of task-relevant information, but also the way in which to respond to such information in the form of a stimulus-response (S-R) mapping (see e.g., Dreisbach et al., 2007). For instance, on a colour trial within one of the experiments in the present study, a task set would direct attention to the relevant feature dimension (colour) and how to respond ("if there is a change in colour, press the 'M' key, if there is no change in colour, press the 'Z' key). While the generation of a robust task set facilitates goal-directed action for a specific task, some degree of attentional control is required to update it in service of a new task or goal, as well as guard against interference from task-irrelevant information which may in some instances afford the same response (see e.g., Cooper et al., 2005; Goschke, 2000; Houghton & Tipper, 1994). However, as evidenced by numerous task switching studies, as well as in cases of frontal lobe damage (see e.g., Lhermitte, 1983), such control is disrupted, resulting in increased error and RTs in the case of task switching, and performance of actions unrelated to current goals in the case of some frontal lobe damage.

In terms of the present research, if the generation of a reliable task set facilitates goal-directed action, the appropriate allocation of vSTM resources therefore relies on a task set consisting of the correct processes and representations for performance of a given task, and as such, relies on attentional control to update the task set when necessary. Using the example of a colour trial, task-relevant representations held in memory should bias attention towards the colours of the presented stimuli (see e.g., Carlisle et al., 2011; Desimone & Duncan, 1995; Duncan, 1984), with vSTM resources subsequently allocated to the colour feature of each stimulus; however, if this colour trial was preceded by an orientation trial, attentional control would be required to update the currently relevant task set, as well as guard against interference from the previously relevant task set. The task-relevant representation of orientation on trial *N-1* would no longer be sufficient for detecting changes in colour in trial *N* and would therefore require updating to correctly perform the new task which would then lead to a misallocation of some (or all) resources to the incorrect feature dimension. More specifically, attentional control in this instance is required to update the attentional set—the element of the task set which identifies and selects relevant information—but not the intentional set—the element of the task set which implements the appropriate S-R mapping—as the relevant feature dimension switched but the task remained the same.

While the process of updating task set takes time, invariably leading to increased RTs, a failure of attentional control to update task set would result in attention being biased towards the irrelevant feature dimension, which would subsequently impact upon the allocation of resources. This would therefore lead to increased error rates, reductions in sensitivity to change, and an increase in change responses (provided there is a constant change in the irrelevant feature dimension; i.e., participants would simply perform the incorrect task). Results from the present set of experiments provides some support for such a view; while evidence for dimension switch costs varies across measure and experimental manipulation, in general it appears that accuracy is most affected when stimuli are multivalent (see Experiments 1b and 3a). This may be indicative of a failure of attentional control and a subsequent misallocation of vSTM resources to the irrelevant feature dimension, driven by the integration of relevant and irrelevant features on the same stimuli. Sensitivity to changes appeared to be most impacted by dimension switching when stimulus displays are more populated, either by multivalent stimuli or an increased number of univalent stimuli (see Experiments 1b, 3a, 3c, and 3d). Thus, it could be suggested that the impact of dimension switching on discrimination was exacerbated by the increased noise associated with these stimulus displays. Finally, response bias appears to have only been impacted to any great extent with either spatially separated (see Experiment 3c) or spatially overlapping (see Experiment 3d) univalent stimuli.

2.12.3 Further research

One area of further research that is explored in the following chapter is to determine whether the dimension switch costs observed throughout the present study is the result of task set inertia or reconfiguration. Through simple trial timing manipulations, the impact of interference from a previously relevant task set or the impact of configuring a new task set will be investigated. Previous research within the task switching literature suggests that by increasing the duration of the response-cue interval (RCI)—the period between response on trial *N-1* and presentation of the cue on trial *N*—facilitates the decay of the task set used on trial *N-1*, subsequently reducing switch costs; shortening the duration of the RCI on the other hand prevents such decay from occurring (see e.g., Meiran, 2000; however, see also Horoufchin et al., 2011a, 2011b). In contrast, increasing the cue-stimulus interval (CSI)—the period between presentation of the cue and presentation of the stimulus on trial *N*—is believed to provide more time for configuration of the newly relevant task set on switch trials, resulting in reduced switch costs, while shortening of the CSI reduces the time available to reconfigure task set (see e.g., Meiran, 1996).

Similarly, potentially confounding variables will also be investigated in the following chapter to ensure that the disruption by dimension switching observed here was not due to some uncontrolled factor; the factor in question is that of the use of a single cue. In the task used throughout the present study, a dimension repetition was paired with a cue repetition, while a dimension switch was paired with a cue switch. Such *cue switching* has received some attention within the task switching literature, having been shown to contribute to the overall switch cost (see e.g., Logan & Bundesen, 2003; Mayr & Kliegl, 2003). As such, further experiments within the next chapter will use two cues per task as a way to overcome this potential confound. By using two cues per task, sequencing consists of *cue repetitions*, *cue switches*, and *task switches*, with the true cost of a task switching the difference between a cue switch and a task switch.

2.12.4 Conclusion

In summary, the present study provides an initial look at how attentional control may be used to allocate vSTM resources. Despite findings varying across measure and experimental manipulation, it appears that the requirement to switch between attending to different feature dimensions while performing a change detection task has a disruptive impact on performance. While speculative, it is suggested that such disruption arises as the result of a failure of attentional control to appropriately update the attentional set required to identify and attend to the relevant visual information, thus causing a misallocation of vSTM resources. Subsequent investigations aim to determine whether the origin of such dimension switch costs is associated with task set inertia or reconfiguration, as well as account for the potential confound caused by using a single cue per task.

3 Chapter 3: Intervals and cues

This chapter outlines a series of experiments which builds upon those in Chapter 2 by investigating the influence of cue-stimulus and response-cue intervals and cue switching on dimension switch costs.

3.1 Introduction

3.1.1 Overview

Attentional control is central in the ability to perform goal-directed behaviour and is critical in everyday life, particularly in situations wherein different tasks are performed in quick succession. Such *task switching* is demanding and the environment (i.e., stimuli) can often afford performance of more than one task (e.g., writing a manuscript on a computer or replying to emails). This requires a delicate balance between attendance to the relevant task and inhibition of goal-irrelevant behaviour, all while maintaining the ability to change tasks when goals change (see e.g., Goschke, 2000). Experimentally, task switching has been studied using various paradigms wherein participants either repeat the same task or switch between performance of two (or more tasks, for reviews see e.g., Grange & Houghton, 2014; Kiesel et al., 2010; Monsell, 2003; Vandierendonck et al., 2010), for instance, the *list procedure* (see e.g., Jersild, 1927; Spector & Biederman, 1976), the *alternating runs* paradigm (see e.g., Rogers & Monsell, 1995), the *explicit cuing* paradigm (Meiran, 2000; see e.g., Meiran, 1996; Sudevan & Taylor, 1987), and the *voluntary task switching* paradigm (see e.g., Arrington & Logan, 2004b; Arrington & Weaver, 2015). Performance on these tasks is measured by error rates (i.e., accuracy) and response times (RTs), with a consistent finding showing that both error rates and RTs are increased when the task switches relative to when the task repeats, referred to as a *switch cost*. A widely held assumption is that each task is associated with a *task set*, the definition of which varies with researcher. Simply put, a task set consists of a set of processes and representations required to perform a given task and must be updated when the task changes (however, for an alternative view, see Altmann & Gray, 2008).

While the observation of switch costs is a consistent finding within the task switching literature, debate remains regarding the origin of such costs. On one hand, some research suggests that switch costs refelct interference from a previously relevant task set (i.e., *task set inertia*; TSI, hereafter referred to as the *interference view*); when a switch occurs, the task set associated with the now-irrelevant task persists (i.e., *proactive interference*) and must be inhibited to prevent interference with the now-relevant task set (see e.g., Allport et al., 1994; Allport & Wylie, 2000; Wylie & Allport, 2000). An alternative account suggests that an endogenous control process reconfigures the task set for performance of the now-relevant task (i.e., *task set reconfiguration*; TSR, hereafter referred to as the *reconfiguration view*); when a switch occurs, the process of updating the task set is time-consuming, thus, the reconfiguration view states that the increased RT on switch trials reflects this process of updating (see e.g., Rogers & Monsell, 1995; Rubinstein et al., 2001). Further accounts suggest that switch costs may arise due to a combination of both interference and reconfiguration (Meiran, 2000; see e.g., Meiran, 1996). Additionally, research has shown that in certain task switching paradigms, a switch in cue can contribute greatly to the observed switch cost, referred to as the *cue switch cost* (Logan & Bundesen, 2004; see e.g., Logan & Bundesen, 2003; Mayr & Kliegl, 2003; Monsell & Mizon, 2006; D. W. Schneider & Logan, 2005, 2011).

The present study aims to investigate the contributions of interference, reconfiguration, and cue switching to the *dimension switch cost*. In the previous chapter, it was shown that by embedding a change detection task within an explicit cuing paradigm—wherein participants either detected changes in colour or orientation—performance was impacted across a range of measures and experimental manipulations when the relevant feature dimension switched relative to when it repeated. These results led to the speculation that the dimension switch cost arises as a result of misallocation of visual short-term memory (vSTM) resources (see e.g., Bays et al., 2009; Bays & Husain, 2008) on switch trials, with experiments here seeking to investigate this possibility further. Results from the present study do not rule out this possibility, but highlight the potential for a distinct origin of dimension switch costs, as interference, reconfiguration, and cue switching were shown to contribute little to the cost of a dimension switch. First, a brief background on the literature surrounding task switching and the various origins of switch costs is presented, as well as a concise outline of the process of resource allocation in vSTM. The cue-stimulus interval and response-cue interval will be manipulated independently in Experiments 1a and 1b respectively, while in Experiment 2, the contribution of cue switching to the overall dimension switch cost will be investigating using two cues per task. Results are discussed in terms of their impact on previous theories while also attempting to adapt such theories to account for the observed results.

3.1.2 Origins of interference and reconfiguration

The seminal work by Allport et al. (1994) and Rogers and Monsell (1995) was responsible for the development of the TSI and TSR accounts respectively (and indeed, the revival of interest in task switching). While both studies employed different experimental paradigms, both observed switch costs, and in some experiments, the same experimental manipulation was performed. Perhaps one of the key manipulations was that of preparation time; by extending the duration between response on trial *N-1* and presentation of the stimulus on trial *N* (i.e., the *response-stimulus interval*; RSI), participants would have additional time to prepare for an upcoming switch, either in terms of dissipating interference or enhanced inhibition on the TSI account, or increased time for reconfiguration on the TSR account. Interestingly, while both studies did find a reduction in the switch cost with extended preparation time, switch costs were never completely eliminated, even at the longest durations (1,110ms and 1,200ms respectively for Allport et al., 1994; Rogers & Monsell, 1995).

While Allport et al. (1994) hypothesised that the switch cost would not be eliminated by increased preparation time, highlighting interference as the cause, explanation of the residual switch cost was more difficult on the TSR account. Rogers and Monsell (1995) had stated that the switch cost observed in their previous experiments reflected the time course of reconfiguration of the task set and as such, extending the time for preparation could reduce the switch cost until it reached zero. However, based on results from the preparation time experiment, Rogers and Monsell suggested that while preparation facilitated an *endogenous component* of reconfiguration, a second, *exogenous component* was required for the completion of reconfiguration; this additional component was triggered upon presentation of the stimulus meaning that the RSI had no influence over this aspect of reconfiguration. Indeed, this suggestion of a two-stage process has some degree of crossover with theories based on interference, such as that associated with response selection/execution (see e.g., Philipp et al., 2007), further highlighting the potential for a theory on the origin of switch costs encompassing both the interference and reconfiguration views.

Despite the theoretical advances made by Allport et al. (1994) and Rogers and Monsell (1995), a major drawback of these studies was the inability to distinguish between which theory best accounted for the origin of switch costs; the mechanisms associated with both the TSI and TSR accounts occurred during the RSI and as such, neither of the paradigms used were able to decompose the influence of interference or reconfiguration on switch costs. It was the application of explicit cuing to a task switching procedure by Meiran (1996) that facilitated a more detailed look at the true nature of switch costs. In explicit cuing, a cue is presented at the beginning of each trial which indicates which task to perform on the upcoming stimuli. For example, if a single digit is presented on each trial, the cue 'value' could indicate the task to be performed is a higher/lower judgement, while the cue 'parity' could indicate an odd/even judgement. The most important aspect of this paradigm however, was that the addition of a cue partitioned the RSI into two distinct intervals: the response-cue interval (RCI) and the cue-stimulus interval (CSI). This meant that the impact of interference from the previously relevant task set and the impact of reconfiguration could be independently investigated. During the RCI, given that the cue has not yet been presented for the upcoming trial, this time period allows for dissipation of the previously relevant task set, but not advance preparation. During the CSI, the cue is now known and while dissipation of the previously relevant task set can still theoretically occur within this interval, this duration is primarily associated with advance reconfiguration of the now-relevant task set.

Meiran (1996) found that when the CSI (termed the *cue-target interval* by Meiran) was extended (1,716ms; Experiments 2 and 3), switch costs were lower relative to when a shorter CSI was used (216ms). Meiran highlighted that the results observed in these experiments were not confounded with task set dissipation as the RCI was held constant, stating that these were the first results which could not be explained in terms of proactive interference. Instead, Meiran suggested that the reduction of switch costs at longer CSIs reflected a process of advance reconfiguration. However, Meiran also outlined disagreement with the theoretical account put forth previously by Rogers and Monsell (1995), stating that it is possible the results obtained by Rogers and Monsell could be explained by the TSI account. Rogers and Monsell outlined two assumptions for advance reconfiguration to occur; the first stated that target onset must be predictable and the second that there should be sufficient time to allow for the completion of reconfiguration. These assumptions were used by Rogers and Monsell to account for the finding that manipulating RSI within a block did not result in any significant interaction with sequencing condition (i.e., target onset was unpredictable with such a manipulation). However, in contrast to this finding, Meiran (1996) observed a significant interaction between CSI duration and sequencing condition, stating that this provided evidence that reconfiguration occurred irrespective of whether target onset was predictable.

While these results suggested that the process of reconfiguration was responsible for switch costs, later research by Meiran et al. (2000) showed that interference still played a role. Results of this study showed that when the RCI was varied (132–3032ms), switch costs were reduced at longer intervals but again, not abolished. This led the authors to suggest that three distinct components contribute to the overall switch cost: 1) dissipation of the previously relevant task set, 2) reconfiguration of the newly relevant task set, and 3) a residual cost that appears to be unaffected by extentions to preparation time (i.e., both CSI and RCI). Taken together, the initial research surrounding the origin of switch costs suggests no consensus was reached; while results are similar in some instances, competing theories have emerged (e.g., Allport et al., 1994; Rogers & Monsell, 1995), with some reconciliation (Meiran et al., 2000; see e.g., Meiran, 1996). Indeed, in a review of task switching by Vandierendonck et al. (2010), the authors discuss the interplay of both interference and reconfiguration, closing with a discussion of theoretical integration of both accounts (see also Kiesel et al., 2010).

3.1.3 Cue switch costs

Despite aiding in the development of theories surrounding the origin of switch costs, the explicit cuing task, like many of the task switching paradigms which had come previously, suffered from a major issue. In a typical task cuing procedure, task repetitions result in cue repetitions, while task switches result in cue switches due to the use of a single cue per task. Thus, it was plausible to assume that a degree of the switch cost in the explicit cuing paradigm could originate from cue switching processes. In order to combat this issue, Logan and Bundesen (2003) and Mayr and Kliegl (2003; see also Logan and Bundesen, 2004; Monsell and Mizon, 2006) employed two cues per tak (hereafter referred to as *2:1 mapping*), which resulted in three sequencing conditions: 1) *cue repetitions*, wherein both the cue and task repeat, 2) *cue switches*, wherein the cue switched but the task repeated, and 3) *task* *switches*, wherein both the cue and task switched. Thus, the *true* or *corrected* task switch cost lies in the difference between performance on cue switch and task switch trials, while the impact of a cue switch is the difference between cue repetition and cue switch trials.

Results from these studies showed that cue switches do indeed contribute to the *uncorrected* task switch cost, with both Mayr and Kliegl (2003) and Logan and Bundesen (2003) finding large cue switch costs. Given that neither the interference or reconfiguration accounts offered explanation for the impact of a cue switch, Mayr and Kliegl (2003) suggested that advance preparation consisted of the retrieval of stimulus-response (S-R) mappings from long-term memory (LTM), with a further process involving the implementation of the task upon presentation of the stimuli. Based on this model, the cost of a cue switch arose as a result of a change in retrieval path in LTM on cue switch trials, whereas on cue repetition trials, a benefit was conferred due to priming of the path. However, Logan and Bundesen (2003) presented an alternative view, suggesting that participants encode both the cue and stimulus (i.e., *compound-cue encoding*, see also Arrington & Logan, 2004a; Logan & Bundesen, 2004; D. W. Schneider & Logan, 2005) and use this joint cue as a method for retrieval of the correct response from LTM. As such, on this view, the cost of a cue switch reflects the process of encoding a new cue-stimulus compound, with cue repetitions conferring a priming benefit.

Later work by Arrington et al. (2007) attempted to dissociate between the encoding of cues and processing of stimuli by having participants respond to the cue prior to target presentation and then respond to the stimulus. This was based on the assumption that the processing of cues and stimuli during the explicit cuing procedure is serial. Thus, if dissociated correctly, cue responses should only be impacted by factors that influence cue encoding, while processing of stimuli should only be impacted by factors that influence stimulus processing. Upon successful dissociation Arrington et al. observed substantial switch costs in stimulus responses, independent of cue encoding. The authors take this as evidence supporting a priming account of cue switch costs, suggesting that some form of representation is generated once the cue is encoded. This is in direct conflict with the view that compound-cue encoding could account for the entirety of observed switch costs in the absence of an endogenous control process (see e.g., Logan & Schneider, 2006; D. W. Schneider & Logan, 2005).

These two theories of cue switch costs have been further distinguished in an attention switching study manipulating *cue transparency* (see Grange & Houghton, 2010), wherein participants searched for an oval target among other ovals containing different features (e.g., border, shaded). Note that in this case, cue transparency is defined as being the extent to which the cue exogenously aids in generation of representations required for task performance; that is to say, a square with a thick border (i.e., an iconic cue) is a more transparent cue than the word “border”. Across two experiments, the authors showed the presence of cue switch costs with less transparent cues, which was eliminated when iconic cues were presented (Experiment 1), while also showing that the absence of a cue switch cost was not due to *cue collapsing*, wherein pairs of iconic cues could be treated as a single cue (Experiment 2; see e.g., Monsell and Mizon (2006)). Grange and Houghton state that the cue switch cost observed for word cues is indicative of a control process required to form representations in working memory. Interestingly, when both the cue and target switched (referred to as an *attention switch* by Grange and Houghton), costs to performance were equivalent, with the authors highlighting that such a difference between cue and attention switching shows that the cost associated with each arise from different origins. This research further supports the view that an endogenous control process is required when the cue changes, refuting the suggestion that a compound cue can account for performance on 2:1 mapping tasks (Logan & Bundesen, 2004; see e.g., Logan & Bundesen, 2003).

3.1.4 Visual short-term memory resource allocation

The purpose of the present study is to adopt these methods for investigation of switch costs in an attempt to determine the underlying nature of resource allocation in vSTM. At present, two theories of vSTM capacity are dominant within the literature; on one hand, capacity is believed to be limited to a small number of high resolution representations (approximately four), with no further information about other items stored above this capacity limit (see e.g., Luck & Vogel, 1997; Vogel et al., 2001; Zhang & Luck, 2008). Alternatively, another theory suggests that capacity is limited by a finite memory resource, distributed to each item in the immediate visual environment, with increasing set szie reducing the precision of each internal representation (see e.g., Bays et al., 2009; Bays & Husain, 2008; Wilken & Ma, 2004). While more recently these theories have been examined via the use of *continuous report* tasks, initial development of theories surrounding vSTM capacity utilised *change detection* tasks which will be used here in an attempt to provide a more comprehensive investigation of resource allocation.

Previously (see Chapter 2), a change detection paradigm was embedded within a explicit task cuing procedure with stimulus displays consisting of either bivalent, coloured, oriented shapes, or univalent shapes either coloured or oriented (note that colour and orientation was present on every stimulus display). A cue was presented at the beginning of each trial indicating which feature dimension was relevant, meaning that in some instances participants attended to the same dimension on consecutive trials (e.g., colour–colour), while on others, switched between dimensions (e.g., colour–orientation). It was found that switching between attending to different feature dimensions relative to attending the same feature dimension impacted performance across a range of measures to varying degrees (potentially due to the experimental manipulations used); accuracy becoming poorer, RTs longer, while sensitivity to change signals was also reduced and the number of change responses increased. It was also observed that these dimension switch costs do not present when there is a change in the relevant feature dimension only; such a method means that participants are free to monitor the entire display for *any* change, rendering the cue non-essential and negating any switch cost that may have occurred.

While these dimension switch costs could simply be a manifestation of “standard’ ’ task switch costs discussed previously, in order to fully develop a theory surrounding vSTM resource allocation, it is important to verify the underlying nature of the process. For instance, given that much of the research discussed previously implicates an additional, rather unspecified process which is initiated upon presentation of a stimulus and which occurs outside the usual endogenous control process associated with reconfiguration, it is plausible to suggest such a process may in fact be (or contain) vSTM resource allocation.

3.1.5 The present study

To investigate the origin of these dimension switch costs, the same embedded change detection switching task used in Chapter 2 (see Experiment 1b) was employed here. In Experiment 1, the CSI and RCI was manipulated independently (Experiments 1a and 1b respectively), with the short interval set to 100ms and the long interval to 900ms. If the dimension switch cost reflects a reconfiguration process, performance on switch trials should be poorer when the CSI is short, with improvements when the CSI is long. Conversely, if the dimension switch cost reflects interference, poorer performance on switch trials should be observed with a shorter RCI, with improvements at longer RCIs. In Experiment 2, two cues were used per task to investigate the contribution of cue switching to the dimension switch cost. As such, Experiment 2a had three sequencing conditions, cue repetitions, cue switches, and task switches. In Experiment 2b cue repetitions were removed.

3.2 Experiment 1: General method

3.2.1 Participants

A sequential Bayes Factor design (see Schönbrodt & Wagenmakers, 2018) was implemented in Experiment 1 (see Sequential Bayes Factor analysis for more detail). This resulted in a total of 70 participants completing Experiment 1a and 69 participants completing Experiment 1b. A binomial test was used to identify participants who did not perform significantly greater than chance (i.e., 50%); this resulted in the exclusion of six participants from data analysis in Experiment 1a (final $N = 64$) and nine participants from Experiment 1b (final $N = 60$). Participants were recruited via a combination of the SONA Participant Recruitment System at Keele University and Prolific (prolific.co). All participants were aged between 18 and 60 years (inclusive) and reported normal or corrected-to-normal visual acuity and normal colour vision. Unique participants were recruited for each experiment, with participants who had completed any of the experiments in Chapter 2 also excluded form participation. Recruitment was limited to the UK and USA and participants were paid a small fee for taking part (Prolific) or awarded course credit (SONA).

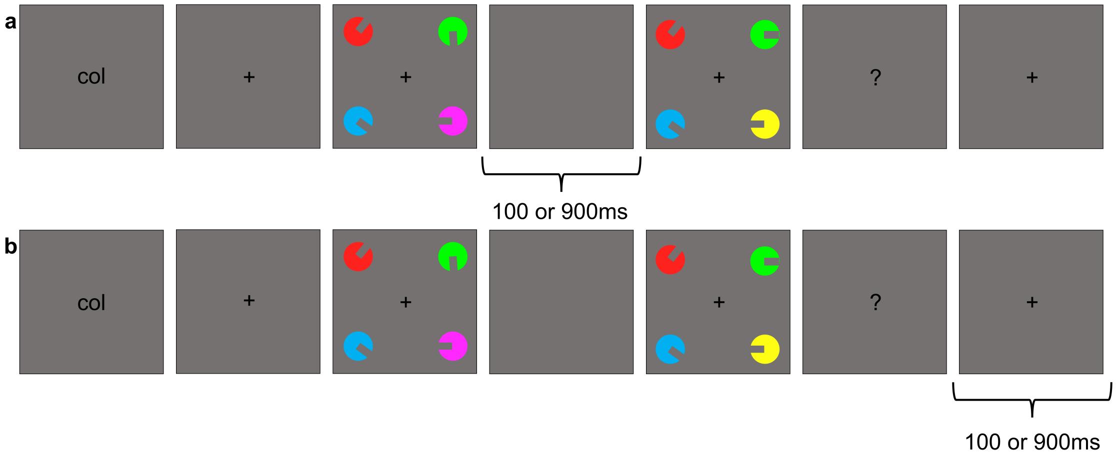

Figure 9
Figure showing example trial procedures in a) Experiment 1a and b) Experiment 1b. Note that a change occurred in both the relevant and irrelevant feature dimensions. The response screen (indicated with a question mark) was blank during the each experiment.

3.2.2 Stimuli and materials

Stimuli in Experiment 1 consisted of four circular shapes which displayed both colour and orientation (see Figure 9). The RGB (red, green, blue) cooridnates for the colours used in Experiment 1 were: black (0, 11, 16), blue (65, 105, 225), cyan (20, 253, 255), green (0, 250, 3), purple (255, 41, 255), red (255, 54, 31), and yellow (253, 254, 21). The background colour was light grey (127, 128, 128), with both cues and fixations presented in dark grey (89, 89, 89). Stimuli were presented within a dark grey frame positioned in the centre of the screen. The cues were “col” if colour was the relevant dimension and “ori” if orientation was the relevant dimension. Both Experiment 1a and 1b were created and run using Gorilla Experiment Builder (Anwyl-Irvine et al., 2020) with restrictions limiting participation to those using desktop or laptop computers only.

3.2.3 Procedure

In each experiment, participants made a change or no change judgement in mixed blocks of trials wherein the relevant feature dimension could repeat (e.g., colour–colour; *repetition*) or switch (e.g., colour–orientation; *switch*). On every trial, an additional change occurred in the irrelevant feature dimension which participants were instructed to ignore. Practice blocks in each experiment consisted of 20 pure repetition trials (10 colour only and 10 orientation only), followed by 20 mixed trials. Accuracy feedback was given at the end of each trial during the practice section only. The main section of the experiment consisted of eight blocks of 50 trials for a total of 400 trials. Self-paced breaks were given every 50 trials.

In Experiment 1a, the duration of the CSI was manipulated, with the duration of the RCI manipulated in Experiment 1b. In each of the experiments, the short interval was 100ms, with the long interval 900ms. In Experiment 1a, the RCI was constant at 500ms across all trials, with the CSI in Experiment 1b also constant at 500ms across all trials. At the beginning of each trial in Experiment 1a, participants were presented with a fixation cross for 250ms, followed by presentation of the cue for a further 250ms. If colour was relevant, the word “col” was presented, with the word “ori” presented if orientation was relevant. Upon extinction of the cue, a CSI of either 100ms or 900ms followed the cue, with subsequent presentation of the memory display for 200ms. After a retention interval of

1,500ms, the test display was presented for 200ms. Upon extinction of the test display, participants could provide response. Pressing “M” indicated a change response while pressing “Z” indicated a no change response. Once a response was provided, an intertrial interval of 250ms followed. The trial procedure for Experiment 1b was identical to that of Experiment 1a with the following exceptions; the fixation cross presented at the beginning of each trial was either presented for 50ms (short) or 450ms (long), the CSI was held constant at 500ms, and the intertrial interval was either 50ms (short) or 450ms (long).

3.2.4 Sequential Bayes Factor analysis

In each experiment, a sequential Bayes Factor analysis with maximal sample size was implemented, with a minimum sample size of 20 and a maximum sample size of 60. Once data was collected from the initial 20 participants, a two-stage procedure began. First, a Bayesian paired-samples t-test was conducted to determine the presence of an overall switch cost. If the BF_{10} was greater than 10 or less than 0.1, a second test was then performed; if not, data collection continued. The second test examined the difference between the switch cost *values* for long and short intervals. This was again a Bayesian paired-samples t-test, with data collection terminated if the BF_{10} was greater than 10 or less than 0.1; if neither threshold was reached, data collection continued. After collection of the initial sample, this two-stage procedure was performed every five participants. On one occasion the BF_{10} for the first-stage test in Experiment 1b exceeded 10, however remained within the limits in the second-stage test. Due to time-constraints placed on data collection in Experiment 1a, prior to running the final two-stage procedure for samples 55–60, four additional samples were collected due to possible attrition (i.e., exclusion due to the result of the binomial test). As the BF_{10} at sample size of 60 did not exceed either limit, the additional four samples have been included in the analysis.

3.2.5 Statistical analysis

Bayesian ANOVAs were used to examine the impact of sequencing condition (repetition, switch) and trial type (long, short) on accuracy, RT, and two measures from signal detection theory d', a measure of sensitivity given as:

$$d' = z(h) - z(f), \tag{6}$$

and *criterion* (c), a measure of response bias given as:

$$c = \frac{1}{2}[z(h) + z(f)], \tag{7}$$

with z representing the inverse of the normal distribution function, h representing hits, and f false alarms (see Macmillan & Creelman, 2004). Detection theory measures were implemented using the *sdtt* package in R (see Moore, 2022). All Bayesian analysis was conducted using the R package *BayesFactor* (see Morey & Rouder, 2021) using package default priors (a Cauchy distribution centred on zero effect with scale parameter $r = 0.707$). Prior to analysis, the first trial in each block was removed as this is neither a repetition or switch trial. Additionally, the trial immediately following an error was removed prior to analysis of accuracy, sensitivity, and response bias data; as the nature of an error cannot be determined, this can make the sequencing condition of the subsequent trial inaccurate, thus having the potential to produce unreliable results. Prior to RT analysis, RTs less than 150ms and 2.5 standard deviations (SDs) above the mean were trimmed per participant, per condition using the R package *trimr* (see Grange, 2022). Error trials and trials immediately following an error were also removed prior to RT analysis due to the potential for post-error slowing.

3.3 Experiment 1a

In Experiment 1a, the duration of the CSI was manipulated in an attempt to investigate the contributions of task set reconfiguration to dimension switch costs. Should reconfiguration be an origin of such costs, reducing the CSI to 100ms should result in greater switch costs as there is less time for the cognitive system to reconfigure the task set. Conversely, increasing the CSI to 900ms should lead to reductions to the extent of the switch cost as there will be more time for the process of reconfiguration to occur; given previous research, even should the switch cost be reduced, there is an expectation that residual switch costs will still be present (see e.g., Rogers & Monsell, 1995).

3.3.1 Results

See Table 1 for BFs for all models in Experiment 1a. Means for each measure in Experiment 1 can be found in Figure 10. Analysis of accuracy data revealed the model containing sequence only as the best model ($\text{BF}_{10} = 8.06$), with comparison of this model against the next best model (sequence + interval, $\text{BF}_{10} = 1.34$) showing the data are 6.01 times more likely under the model containing sequence only. All other models had a BF_{10} less than one, therefore model comparison was not conducted. Analysis of RT data also revealed the model containing sequence only as the best model ($\text{BF}_{10} = 10{,}857.69$).

Table 1

Table displaying Bayes Factors (BF_{10}) for each model across accuracy, response time, sensitivity, and response bias analyses in Experiment 1a. The BF_{10} for each model is in comparison to the null model.

Measure	Model	BF	Error
Accuracy	Sequence	8.056	0.037
	Interval	0.167	0.019
	Sequence + Interval	1.34	0.015
	Sequence + Interval + Interaction	0.237	0.023
Response time	Sequence	10857.693	0.013
	Interval	0.172	0.007
	Sequence + Interval	2120.144	0.038
	Sequence + Interval + Interaction	6678.776	0.025
Sensitivity	Sequence	10.174	0.057
	Interval	0.141	0.011
	Sequence + Interval	1.37	0.015
	Sequence + Interval + Interaction	0.287	0.022
Response bias	Sequence	0.333	0.013
	Interval	0.453	0.007
	Sequence + Interval	0.152	0.014
	Sequence + Interval + Interaction	0.033	0.026

Comparison of this model against the next best model (sequence + interval + interaction, BF_{10} = 6,678.78) showed the data are 1.63 times more likely under the model containing sequence only and 5.12 times more likely when compared to the third best model (sequence + interval, BF_{10} = 2,120.14). Analysis of sensitivity data revealed the model containing sequence only as the best model (BF_{10} = 10.17), with comparison of this model against the next best model (sequence + interval, BF_{10} = 1.37) showed the data are 7.43 times more likely under the model containing sequence only. Analysis of response bias data revealed a BF_{10} less than one for all models. Note that the log-linear correction (see Hautus, 1995) was used to calculate hit and false alarm rate due to extreme values.

3.3.2 Discussion

While the results of Experiment 1a revealed varying degrees of evidence for a dimension switch cost across all measures, the extent of the switch cost does not appear to be influenced by the length of the CSI. These results suggest that advance reconfiguration does not contribute to the dimension switch cost observed here and in previous experiments using the same task (see Chapter 2). It is possible that the origin of the dimension switch cost lies with interference from the previously relevant task set which will be investigated in Experiment 1b.

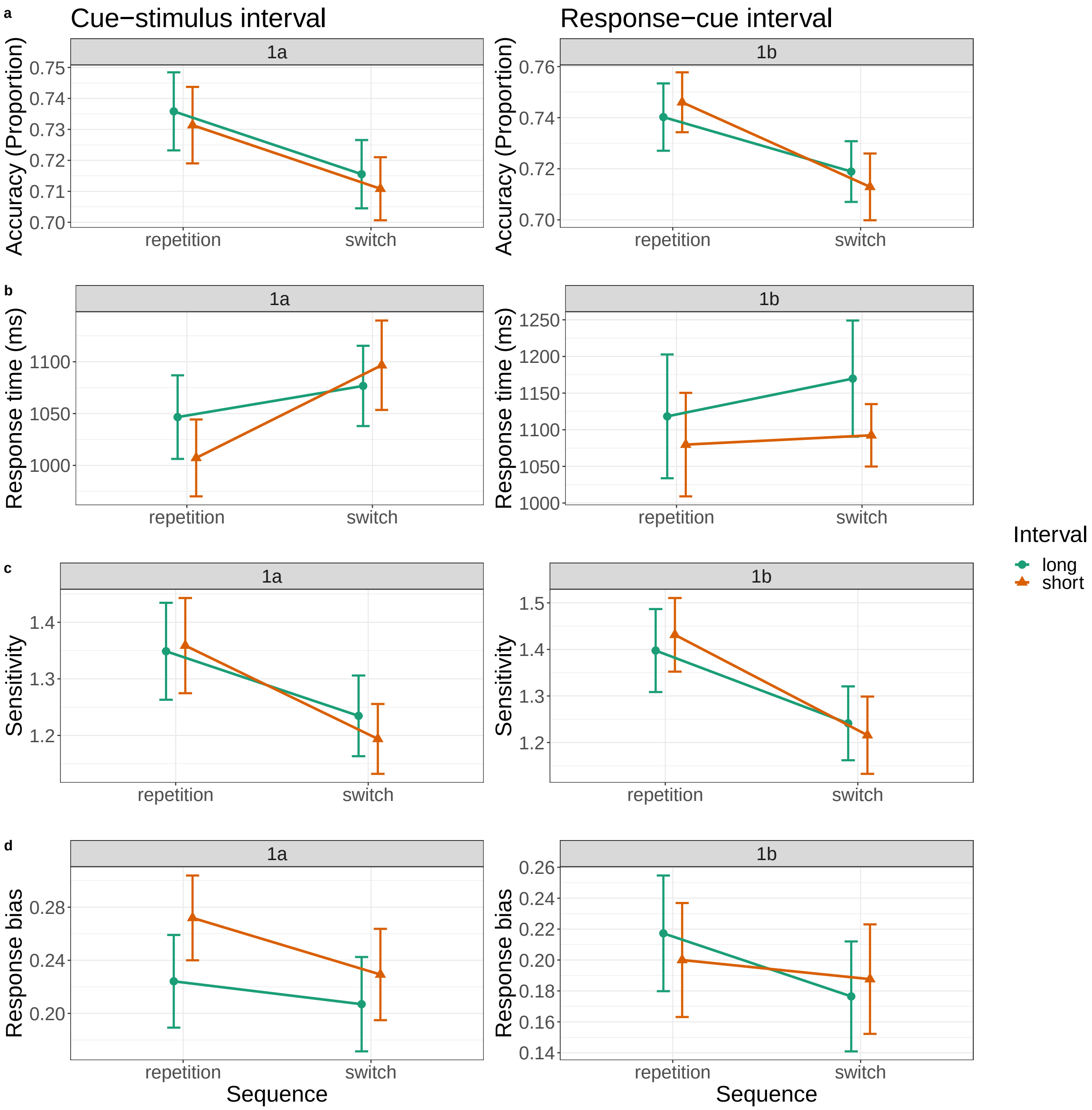

Figure 10

Plots displaying mean a) accuracy, b) response time, c) sensitivity (d'), and d) response bias (c) across sequencing condition and interval duration for Experiment 1a and Experiment 1b. Error bars represent the standard error of the mean.

3.4 Experiment 1b

In Experiment 1b, the duration of the RCI was manipulated to investigate the contributions of TSI (i.e., interference, inhibition). As in Experiment 1a, the short duration was set to 100ms with the long duration set to 900ms. At the shorter RCI, should TSI contribute to the dimension switch cost, switch costs will increase as there will be less time for the previously relevant task set to dissipate or be inhibited. Conversely, increasing the duration of the RCI should provide more time for these processes to occur, thus reducing the extent of the switch cost. Again, even if reductions to the switch cost are observed, it is still expected that residual switch costs will be present (see e.g., Allport et al., 1994).

3.4.1 Results

See Table 2 for BFs for all models in Experiment 1b. Analysis of accuracy data revealed the model containing sequence only as the best model ($BF_{10} = 152.78$). Comparison of this model against the next best model (sequence + interval, $BF_{10} = 21.26$) showed the data is 7.18 times more likely under the model containing sequence only. A further comparison showed the data were 28.59 times more likely under the model containing sequence only when compared to the third best model (sequence + interval + interaction, $BF_{10} = 5.34$. Analysis of RT data revealed that all models had a BF_{10} less than one. Analysis of sensitivity data revealed the model containing sequence only as the best model ($BF_{10} = 441.76$). Comparison with the next best model (sequence + interval, $BF_{10} = 58.47$) revealed the data are 7.56 times more likely under the model containing sequence only and 32.78 times more likely compared to the third best model (sequence + interval + interaction, $BF_{10} = 13.48$). Analysis of response bias data revealed that all models had a BF_{10} less than

Table 2

Table displaying Bayes Factors (BF_{10}) for each model across accuracy, response time, sensitivity, and response bias analyses in Experiment 1b. The BF_{10} for each model is in comparison to the null model.

Measure	Model	BF	Error
Accuracy	Sequence	152.782	0.037
	Interval	0.141	0.019
	Sequence + Interval	21.264	0.014
	Sequence + Interval + Interaction	5.345	0.024
Response time	Sequence	0.248	0.013
	Interval	0.938	0.007
	Sequence + Interval	0.259	0.038
	Sequence + Interaval + Interaction	0.058	0.026
Sensitivity	Sequence	441.765	0.056
	Interval	0.139	0.011
	Sequence + Interval	58.468	0.015
	Sequence + Interval + Interaction	13.475	0.022
Response bias	Sequence	0.291	0.013
	Interval	0.137	0.008
	Sequence + Interval	0.04	0.014
	Sequence + Interval + Interaction	0.009	0.025

one. Note that the log-linear correction (see Hautus, 1995) was used to calculate hit and false alarm rate due to extreme values.

3.4.2 Discussion

While evidence for dimension switch costs was observed for both accuracy and sensitivity in Experiment 1b, it appears that the extent of these costs was not influenced by the duration of the RCI. As such, it appears that the extent of dimension switch costs are not determined by interference or inhibition. Taken together, the results of Experiment 1 suggest that the origin of dimension switch costs is either entirely distinct from that observed in traditional task switching or that processes associated with the task set on switch trials have minimal influence on resource allocation, making these differences barely detectable. Indeed, it is also plausible that the dimension switch cost is simply a cue switch cost, which will be investigated in Experiment 2.

3.5 Experiment 2: General method

3.5.1 Participants

A total of 84 participants took part in Experiment 2a with 100 participants completing Experiment 2b. In Experiment 2a[7] a sequential Bayes Factor design was used (see Sequential Bayes Factor analysis for more detail). A binomial test was used to identify participants who did not perform significantly better than chance (i.e., 50%). As such, 14 participants were excluded from analysis in Experiment 2a (final $N = 70$), while 15 participants were excluded from analysis in Experiment 2b (final $N = 85$). Participants were recruited via a combination of the SONA Participant Recruitment System at Keele University and Prolific (prolific.co). All participants were aged between 18 and 60 years (inclusive) and reported normal or corrected-to-normal visual acuity and normal colour vision. Participants who had previously completed any of the experiments within Chapter 2 or Experiments 1a or 1b within the current chapter were prevented from participating. Recruitment was limited to the UK and USA and participants were paid a small fee for taking part (Prolific) or awarded course credit (SONA).

---

[7] Data from Experiment 2b was obtained as part of a separate study within our group which investigated the association between depression symtomatology and vSTM in our switching design.

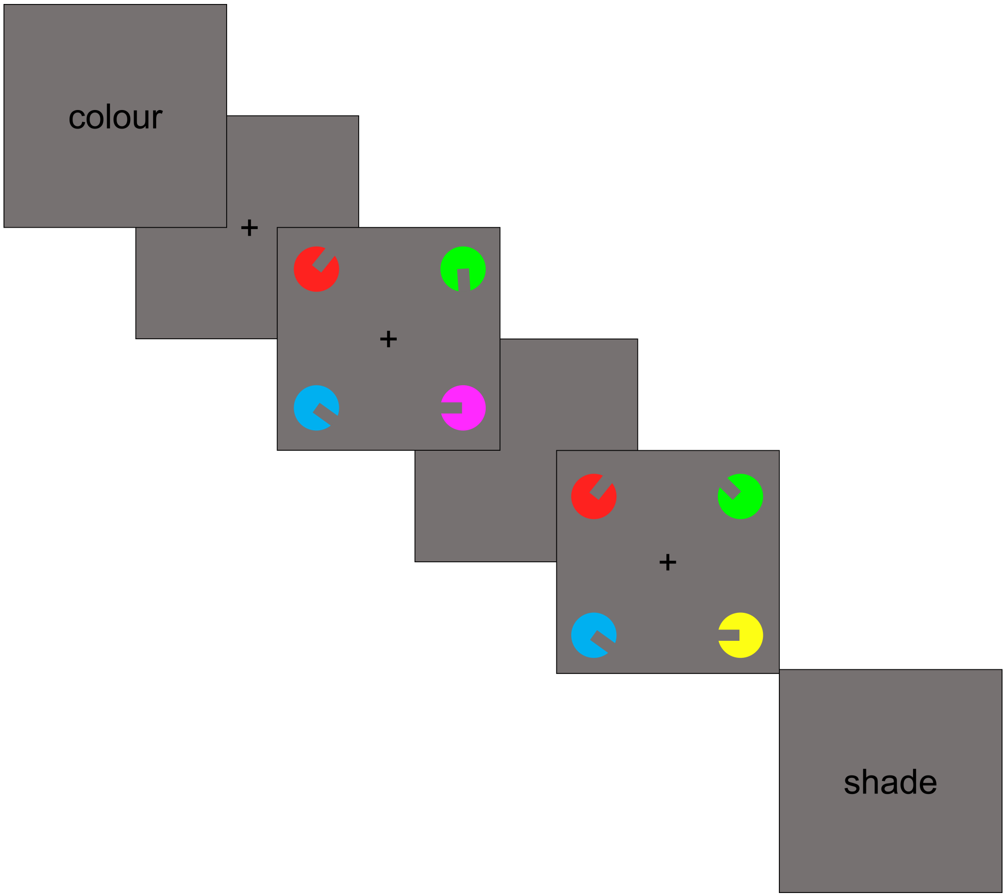

Figure 11
Figure showing example trial procedure in Experiment 2. Note that a change occurred in both the relevant and irrelevant feature dimensions.

3.5.2 Stimuli and materials

The stimuli used in Experiment 2 were identical to those used in Experiment 1 (see Figure 11. The addition in Experiment 2 was that a total of four cues were used. If colour was relevant, the cue could either be the word “colour” or the word “shade”, whereas if orientation was relevant, the cue could either be the word “orientation” or the word “angle”. Both experiments were created and run using Gorilla Experiment Builder (Anwyl-Irvine et al., 2020), with participation limited to those using desktop or laptop computers only.

3.5.3 Procedure

In each experiment, participants made change or no change judgements; in Experiment 2a, both the dimension and cue could repeat (i.e., a cue repetition trial), the dimension could repeat but the cue could switch (i.e., a cue switch trial), or both the dimension and cue could switch (i.e., a dimension switch trial). In Experiment 2b, cue repetition trials were removed. On every trial, an additional change in the irrelevant feature dimension occurred which participants were instructed to ignore. Practice blocks in each experiment consisted of 20 pure repetition trials (10 colour only and 10 orientation only) followed by 20 mixed trials. Accuracy feedback was provided at the end of each trial in the practice section only. In the main section of the experiments participants completed eight blocks of 50 trials for a total of 400 trials. Self-paced breaks were given every 50 trials.

At the beginning of each trial, participants were presented with a fixation cross for 750ms, followed by presentation of the cue indicating which feature dimension was relevant for 500ms. If colour was relevant, the cue could either be the word “colour” or “shade”, while if orientation was relevant, the cue could either be the word “orientation” or “angle”. A CSI of 500ms then followed, with subsequent presentation of the memory display for 200ms. A retention interval of 1,500ms was followed by presentation of the test display for 200ms. Once the test display was extinguished, participants provided response by pressing one of two keys; pressing “M” indicated a change response, while pressing “Z” indicated a no change response. Responses were not time limited and were followed by a 250ms intertrial interval.

3.5.4 Sequential Bayes Factor analysis

A sequential Bayes Factor analysis with maximal sample size was implemented in Experiment 2a, with a minimum sample size of 20 and a maximum of 70. The critical test was a Bayesian repeated-measures ANOVA to determine the impact of sequencing condition (cue repetition, cue switch, dimension switch) on accuracy data. If the BF_{10} was greater than 10 or less than 0.1 (indicative of strong evidence in favour of the alternative and null respectively; see Lee and Wagenmakers (2013)), data collection was terminated. If neither threshold was reached, data collection continued with the test performed every five participants until the maximum sample size was reached.

3.5.5 Statistical analysis

Bayesian repeated-measures ANOVAs were used to examine the effect of sequencing condition (cue repetition, cue switch, dimension switch) on accuracy, RT, sensitivity, and response bias measures in Experiment 2a. If the BF_{10} was greater than or equal to 10, a Bayesian paired-samples t-test was conducted to determine where the effect lay. In Experiment 2b, Bayesian paired-samples t-tests were used to examine the difference between performance on cue switch and dimension switch trials for accuracy, RT, sensitivity, and response bias. Analysis methods and procedures for data trimming and error removal were identical to those used in Experiment 1.

3.6 Experiment 2a

In Experiment 2a, the general task was identical to that used in Experiment 1 except that the trial timings were constant across all trials. The manipulation in this experiment concerned the use of two cues per task; when colour was relevant, the word “colour” or “shade” were presented at the beginning of the trial, whereas when orientation was relevant, the word “orientation” or “angle” was presented. Thus, this resulted in three sequencing conditions, cue repetitions, cue switches, and dimension switches; the contributions of a cue switch can be observed independently from that concerned with dimension switching, enabling a *corrected dimension switch cost* to be obtained.

3.6.1 Results

Means from each measure in Experiment 2 can be seen in Figure 12. Analysis of accuracy data revealed a $\mathrm{BF}_{10} = 2.52$ indicating anecdotal evidence in support of an effect of sequencing condition. Analysis of RT data revealed a $\mathrm{BF}_{10} = 47.93$ indicating very strong evidence in support of an effect of sequencing condition. A follow-up Bayesian paired-samples t-test between cue repetition (M = 1,011.15, SD = 289.90) and cue switch (M = 1,007.65, SD = 264.02) data revealed a $\mathrm{BF}_{10} = 0.14$ indicating moderate evidence in support of the absence of a cue switch cost, while further analysis between cue switch and dimension switch (M = 1,050.17, SD = 298.04) data revealed a $\mathrm{BF}_{10} = 35.32$ indicating very strong evidence in support of the presence of a switch cost. Analysis of sensitivity data revealed a $\mathrm{BF}_{10} =$ 9.66 indicating moderate evidence in support of an effect of sequencing condition. Given this result was essentially at the threshold for post-hoc comparisons as stated in the method, comparison of sensitivity between cue repetition (M = 1.33, SD = 0.61) and cue switch (M = 1.29, SD = 0.62) data revealed a $\mathrm{BF}_{10} = 0.17$ indicating moderate evidence in support of the absence of a switch cost. Comparison of cue switch and dimension switch (M = 1.16, SD = 0.62) data revealed a $\mathrm{BF}_{10} = 2.82$ indicating anecdotal evidence in support of a switch cost. Finally, analysis of response bias data revealed a $\mathrm{BF}_{10} = 1.59$ indicating anecdotal evidence in support of an effect of sequencing condition. Note that the log-linear correction (see Hautus, 1995) was used to calculate hit and false alarm rate due to extreme values.

3.6.2 Discussion

Dissociation of the contributions of a cue switch to the uncorrected dimension switch cost in Experiment 2a provide some evidence that cue switches do not contribute to dimension switch costs to any great extent. Despite a general lack of evidence supporting an effect of sequencing condition for most measures, the observation of a substantial corrected switch cost for RT and lack of cue switch cost support the notion that dimension switch costs originate from a source other than cue switching. Experiment 2b provides a more robust estimate of the corrected dimension switch cost.

3.7 Experiment 2b

The task in Experiment 2b is identical to that used in Experiment 2a with the exception that cue repetition trials were removed. This resulted in an increase in cue switch and dimension switch trials which will allow a more robust estimate of the corrected dimension switch cost to be obtained.

3.7.1 Results

Analysis of accuracy between cue switch (M = 0.72, SD = 0.10) and dimension switch (0.70, SD = 0.09) data revealed a $\text{BF}_{10} = 2.57$ indicating anecdotal evidence in support of the presence of a switch cost. Analysis of RT data between cue switch (M = 1,015.30, SD = 347.44) and dimension switch (M = 1,051.68, SD = 423.40) data revealed a $\text{BF}_{10} = 3.01$ indicating moderate evidence in support of the presence of a switch cost. Analysis of sensitivity between cue switch (M = 1.37, SD = 0.71) and dimension switch (M

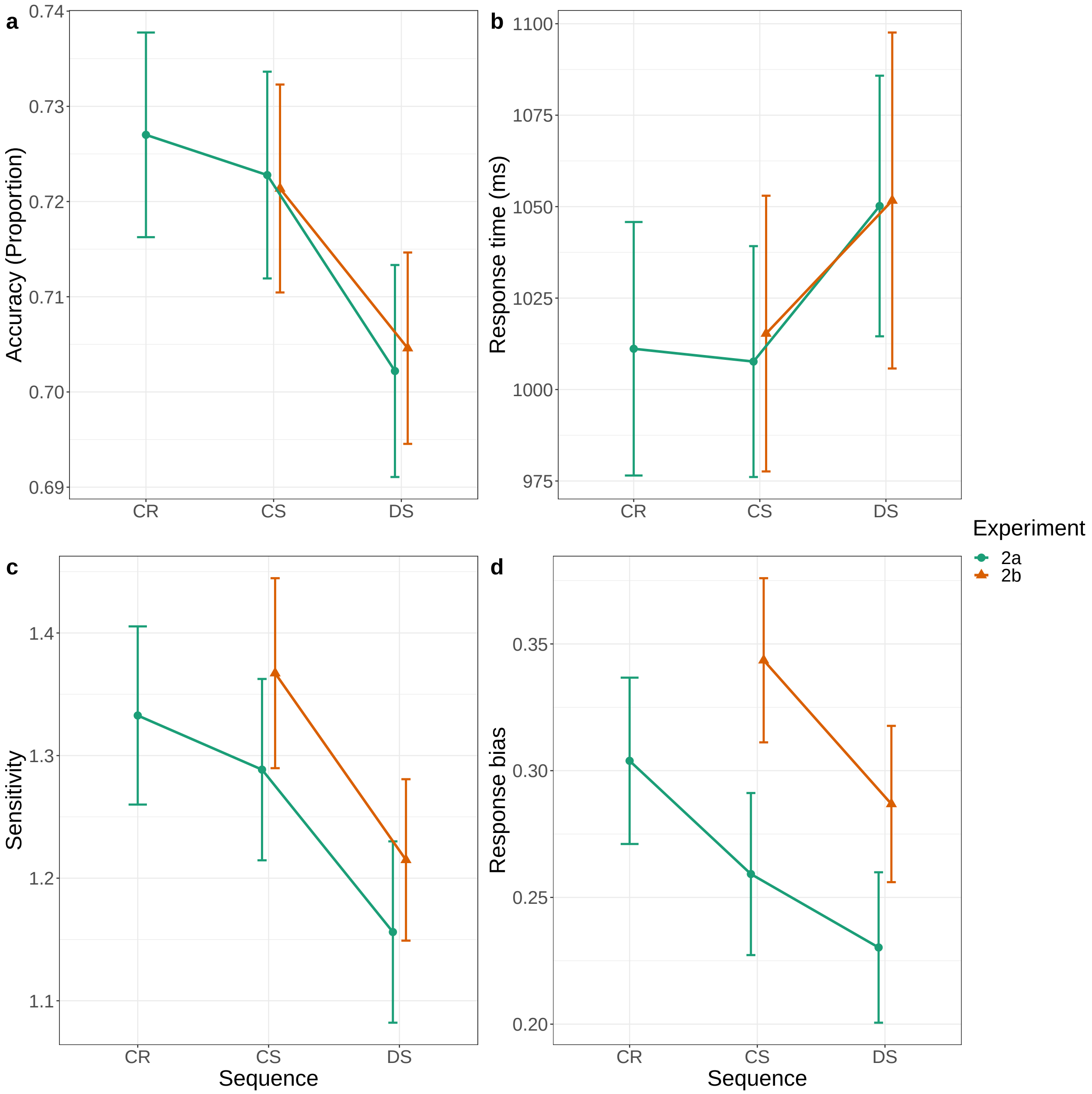

Figure 12

Plots displaying mean a) accuracy, b) response time, c) sensitivity (d′), and d) response bias (c) across sequencing condition (CR = cue repetition, CS = cue switch, DS = dimension switch) in Experiment 2. Error bars represent the standard error of the mean.

= 1.22, SD = 0.61) revealed a BF_{10} = 24.89 indicating strong evidence in support of the presence of a switch cost. Finally, analysis of response bias between cue switch (M = 0.34, SD = 0.30) and dimension switch (M = 0.29, SD = 0.28) data revealed a BF_{10} = 5.29 indicating moderate evidence in support of the presence of a switch cost.

3.7.2 Discussion

The aim of Experiment 2b was simply to provide a more robust estimate of the corrected dimension switch cost. While there was a marked reduction in evidence supporting the presence of a switch cost for RT relative to that found in Experiment 2a, the most notable result from Experiment 2b is perhaps the substantial evidence observed for a dimension switch cost in sensitivity, suggesting that participants found discrimination of change signals more difficult on dimension switch trials.

3.8 General discussion

3.8.1 Summary of results

To recap, Experiment 1 sought to determine the contributions of reconfiguration (Experiment 1a) and interference (Experiment 1b) to dimension switch costs. Overall, the results showed that dimension switch costs did not differ as a function of interval duration; sequencing condition had the greatest impact on performance, with lower levels of evidence supporting an interaction effect between sequencing condition and interval duration. Experiment 2 aimed to determine contributions of a cue switch to the dimension switch cost. Primarily, results from Experiment 2a showed little evidence for an effect of sequencing condition overall; however, most notably, substantial evidence was found for a dimension switch cost in RT whereas there was little evidence supporting a cue switch cost. Experiment

2b simply provided a more robust estimate of the corrected cue switch cost.

3.8.2 Theoretical impact

As alluded to in the introduction, the switch costs found previously in Chapter 2 may reflect a misallocation of vSTM resources; on switch trials, these resources may be allocated to the incorrect feature dimension due to incomplete reconfiguration or persisting interference of previously relevant task sets. However, the results of the present study appear to suggest that dimension switch costs have an origin distinct to those found in more traditional task switching paradigms. Considering the manipulations in Experiment 1, previous research showed that the extent of the task switch cost can be attenuated by extending preparation time (see e.g., Allport et al., 1994; Rogers & Monsell, 1995), with this additional time facilitating either the process of reconfiguration or dissipation/inhibition of a previously relevant task set. Given the lack of evidence to suggest interval duration impacted the extent of the dimension switch cost for both CSI and RCI manipulations, this may suggest that dimension switching relies on a process (or processes) distinct from those associated with traditional task switching. In addition, it is also possible that there may be an effect of CSI or RCI on the dimension switch cost, but that the experiments employed here were not sensitive enough to capture this. This may have been a power issue, however it may be more plausible to assume that the use of online testing may have impacted the results of this study more so than sample size.

One possibility is that processes associated with dimension switching are contained within the exogenous component suggested by Rogers and Monsell (1995) to account for residual switch costs on the TSR account. Rogers and Monsell suggested that extending preparation time (i.e., the RSI) facilitates completion of the endogenous component during a task switch; however, as switch costs were still observed at longer intervals, the authors hypothesised that an additional, exogenous component was required to complete reconfiguration. As this exogenous component is triggered by presentation of the stimuli, it is unaffected by manipulations of preparation time. While others have suggested alternative processes distinct to the endogenous component occur prior to stimulus presentation (see e.g., Meiran, 1996), it would be plausible to assume that a process of resource allocation is initiated upon presenation of the stimuli and as such, the dimension switch costs observed in Experiment 1 may directly reflect issues within this exogenous component, namely a misallocation of resources that cannot be attenuated through increased time for reconfiguration. Indeed, given that the task itself remained the same throughout (i.e., detect a change in the relevant feature dimension), it could be suggested that performance on the paradigm used in Experiment 1 reflects *only* the exogenous component, with disruption to associated processes responsible for the observed switch costs.

Perhaps a more refined view can be garnered based on the *executive control theory of visual attention* (ECTVA, see e.g., Logan & Gordon, 2001; Logan & Schneider, 2010). Within this model, certain parameters have been adopted to distinguish between different aspects of a task set. The parameters β and π represent a measure of response bias and attentional weighting of stimuli respectively which, broadly speaking, can be considered as representations of *intentional set* and *attentional set* (see also Meiran, 2000). In the present study, while the attentional set changed when a switch in feature dimension occurred, the intentional set remained the same on every trial. Thus, it could be suggested that the switch cost observed reflects the cost associated with the updating of only the π parameter in terms of the ECTVA model; in more general terms, dimension switch costs may in fact originate from processes pertaining to the stimuli (e.g., resource allocation, stimulus encoding), and not those associated with interference or reconfiguration. While these stimulus-based processes may still be under executive control (see e.g., Logan & Schneider, 2010), given that they are only initiated upon presentation of the stimuli, extended preparation time would have little effect on the disruption of these processes as the result of a switch.

In relation to cue switch costs, while previous research employing a 2:1 mapping (see e.g., Logan & Bundesen, 2003; Mayr & Kliegl, 2003; Monsell & Mizon, 2006) has shown substantial contributions of cue switching to the task switch cost, no such cost was observed in the present study. Rather, there was a general lack of evidence supporting an effect of sequencing condition in Experiment 2a however, notably substantial evidence was observed for a dimension switch cost for RT relative to the lack of evidence for a cue switch cost. Further developing the line of reasoning discussed previously which suggested that only the attentional set changed, it could be posited that the cue switch cost relates to changes in intentional set which is why little evidence was found for such a cost in the present study. In traditional task switching experiments the processes and representations (i.e., the task set) necessary for performance of the upcoming task require activation when a switch in task occurs; the timecourse of such activation likely initiated upon sufficient encoding of the cue. However, in the present study, the cue simply indicated which feature dimension was relevant, helping to weight attention towards the most relevant information. It may therefore be plausible that cue switch costs more directly reflect the requirement to update intentional set despite the fact that the task itself may remain the same.

In general, the results from Experiment 2, however weak, appear to show that cue switching does not contribute to dimension switch costs to any great extent. As such, the results observed may fall more in line with those found by Grange and Houghton (2010), wherein word cues produced a cue switch cost, but iconic cues did not; yet an equivalent attention switch cost (here this would be the dimension switch cost) was observed for both word and iconic cues. These results are indicative of a dissociation between costs—and therefore potentially processes—associated with switching cues and switching attention. Indeed, further experimentation is required before drawing any robust conclusions, however it is pertinent to provide at least some theoretical speculation given that this is one of the first investigations of cue switching in a dimension switching context.

3.8.3 Conclusion

The present study aimed to determine the relative contributions of manipulations of the CSI and RCI as well as cue switching to the dimension switch cost. Overall, these results suggest that dimension switch costs are unaffected by manipulations which have been shown to influence task switch costs in traditional task switching experiments. Despite the presence of dimension switch costs, there was little evidence to suggest that such costs are impacted by the duration of both the CSI or RCI. It is postulated that these results may reflect an additional, exogenous component associated with the attentional set and unaffected by reconfiguration and proactive interference. Further results from the present study also show that cue switching does not contribute to the overall dimension switch cost; while perhaps more speculative, it is suggested that this may be due to the lack of requirement to generate a new intentional set when the task or cue switches however, additional research should be undertaken before making any robust claims. Taken together, these results highlight an origin of disruption distinct to that found in traditional task swtiching however, given the prevalence of these effects within the task switching literature (in particular the CSI), it is crucial that further, laboratory-based research is conducted.

4 Chapter 4: Continuous report

This chapter outlines an experiment which made use of a continuous report task embedded within a task switching paradigm to further investigate the impact of switching between different feature dimensions. A more nuanced look at potential origins for the dimension switch cost was achieved via mixture modelling.

4.1 Introduction

4.1.1 Overview

Visual short-term memory (vSTM) is inextricably linked with other forms of cognition such as attention (see e.g., Awh & Jonides, 2001) thus, our understanding of this system can help to better understand other functions within the brain and disorders that may arise. Through the years, a number of different theories regarding the underlying nature of vSTM have been developed, with two most prominent. One theory suggests that vSTM is limited in capacity by the total number of high resolution item representations that can be retained at a given time (see e.g., Zhang & Luck, 2008, 2011), while an alternative theory suggests that vSTM is limited by a finite memory resource, responsible for the precision at which items are represented at in memory (see e.g., Bays et al., 2009; Bays & Husain, 2008; Wilken & Ma, 2004); on this resource account, as the total number of items to be retained increases, the precision of each internal item representation decreases as less resource can be allocated to each item. More recently, the role of attention in the allocation of such vSTM resources has been investigated (see e.g., Dube et al., 2017; Emrich et al., 2017; Henderson et al., 2020), with the suggestion that attention serves to allocate resources to relevant visual information. To further examine this suggestion, the present chapter makes use of a novel *continuous report* task; in standard versions of this task, participants are presented with stimuli (typically coloured shapes) and report the colour of a probed stimulus using a 360° response wheel. Here, participants instead reported either the colour or orientation of a probed stimulus, with the relevant feature either repeating (e.g., colour–colour) or switching (e.g., colour–orientation) on consecutive trials. This addition of switching introduces an element of attentional control to a task otherwise devoid of such control. Furthermore, the fitting of a mixture model to behavioural data will allow a more nuanced look into the origin of any observed cost to performance.

4.1.2 Initial research on visual short-term memory precision

While much earlier research alluded to a memory resource being responsible for the fidelity of internal representations (see e.g., Frick, 1988), perhaps the most influential paper challenging the notion of a fixed-capacity vSTM was that by Wilken and Ma (2004) (specifically Experiments 7–9). Previous research (see e.g., Luck & Vogel, 1997; Vogel et al., 2001) had suggested that only a small number high resolution item representations could be retained in memory at a given time (around 3–4). However, Wilken and Ma (2004) highlighted issues with the type of *high-threshold* (HT) model used to arrive at such a conclusion, namely that items are encoded into memory in the absence of noise. Initial experiments (Experiments 1–6) examined how well two signal detection theory accounts fit data obtained from change detection tasks with various manipulations (e.g., set size, target number), relative to a HT model. However, while both the *maximum absolute differences* (MAD) and *sum of absolute differences* (SAD) models provided a better fit to the data than the HT model, Wilken and Ma used a continuous report task to more directly estimate the noise associated with stimulus encoding. Rather than presenting a second stimulus array, wherein comparisons are made between stimuli in the first and second displays, Wilken and Ma presented a probe display, highlighting the location of one of the stimuli in the initial stimulus display; participants then simply attempted to reproduce the feature value in the corresponding position from the first stimulus display.

Stimuli consisted of coloured squares (Experiment 7), Gabor patches (Experiment 8), and Gabor patches with additive spatial frequency (Experiment 9). Set size varied with $N =$ 2, 4, 6, or 8 for the colour and spatial frequency tasks and $N =$ 2, 3, 4, or 5 for the orientation task. Following presentation of the stimuli, a probe display was presented, highlighting the location of one of the stimuli; participants responded by either clicking a location on a 360° colour wheel they believed best matched the colour of the stimulus that appeared in the probed location (colour task) or manipulated a probe Gabor using the left and right arrow keys to the position they believe best matched the stimulus that appeared in the probed location (orientation and spatial frequency tasks), pressing spacebar to confirm response. By calculating the angular deviation—the distance between response and the true value of the target—Wilken and Ma were able to obtain an estimate of how precisely the probed item was represented in memory, with error distributions revealing that precision declined as set size increased. Critically, no substantial reduction in precision was observed as set size exceeded the supposed capacity limit suggested by previous research (approximately four items, see e.g., Luck & Vogel, 1997; Pashler, 1988; Vogel et al., 2001). Wilken and Ma (2004) take this as evidence against a fixed-capacity vSTM, stating that should capacity be fixed in such a way, performance on this continuous report task should be excellent until the upper limit was reached as, according to previous research, all items would be stored in memory with a high degree of precision (see e.g., Luck & Vogel, 1997). Once the capacity limit was breached, this would lead to a marked increase in noise, with participants forced to guess on some trials as only a subset of items would be represented in memory.

4.1.3 Development of the two-component mixture model

Later work by Zhang and Luck (2008) however, seemingly ruled out vSTM models wherein a larger number of items (than the previously supposed fixed capacity limit) could be stored at resolutions determined by the total number of items to be encoded. Like Wilken and Ma (2004), Zhang and Luck (2008) used a continuous report task wherein stimuli were coloured squares with varying set size (Experiment 1 $N = 3$ or 6; Experiment 2 $N = 1$, 2, 3, or 6; Experiment 3 $N = 4$; Experiment 4 $N = 3$). Through the use of standard estimation techniques, Zhang and Luck extracted two components from the behavioural data obtained from these experiments P_m representing the probability that an item was represented in memory and *s.d.* representing how precisely the probed item was represented in memory. This model was later formalised as into what is referred to as the *two-component* mixture model:

$$p\left(\hat{\theta}\right) = \left(1 - p_u\right)\phi_\kappa\left(\hat{\theta} - \theta\right) + p_u\frac{1}{2\pi}, \tag{8}$$

where θ represents the value of the probed stimulus (in radians), $\hat{\theta}$ represents the participant's response, with p_u representing the probability of a random guess. ϕ represents the probability density of the von Mises distribution with mean zero and concentration parameter κ (see e.g., Grange & Moore, 2022). Experiment 1 revealed that while *s.d.* (i.e., precision) did not vary to any great extent between set sizes of three and six, P_m was found to approximately half when set size increased from three to six. The authors took this as evidence to rule out other models which suggested that vSTM wherein all items are stored with resolution a function of the total number of items to be stored. However, Zhang and Luck also highlighted the possibility that both limited capacity (i.e., *fixed slots*) and a finite resource could contribute to performance. To investigate this notion, in Experiment 2 the authors tested two differing models, the *slots-plus-resources* and *slots-plus-averaging* models. On the slots-plus-resources model, while capacity is still limited to a small number of items, the resource can be flexibly allocated across these items, meaning a single item could receive more resource than others, increasing its resolution. The slots-plus-averaging model on the other hand posits that each of the slots contains a set amount of resource; while each slot can store a single item, a single item can also be stored in multiple slots, thus increasing its resolution. Results from Experiment 2 revealed that *s.d.* increased as set size increased from one to three, remaining constant as set size increased to six, a prediction made by both models; as set size is below the supposed upper limit of vSTM capacity, additional resource is available to be allocated across items. Further results from Experiment 2 revealed that P_m gradually declined as set size increased from one to three before declining dramatically as set size increased to six, providing further evidence against models wherein all items would be stored with precision dependent on set size.

In order to discern which of the slots-plus-resources and slots-plus-averaging models Experiment 3 made use of a line in the stimulus display cuing participants to a particular stimulus. On 70% of trials the cued stimulus was probed (i.e., valid trials), while each of the remaining stimuli were probed on 10% of trials (i.e., invalid trials), with neutral cues wherein all four stimuli were probed also included. Zhang and Luck stated that the slots-plus-resource model predicted that on such a task, the majority of the resource would be allocated to the cued stimulus causing a substantial difference in *s.d.* between the different trial types, but having only a minimal effect on P_m. In contrast, the slots-plus-averaging model posited that most of the slots would be allocated to the cued item, resulting in substantial differences in P_m between valid and invalid trials; on valid trials, the probed stimulus would be represented in memory having been allocated multiple slots however, on invalid trials, the slots would have been allocated to the wrong stimulus. Furthermore, the authors stated that *s.d.* would be reduced on valid trials relative to neutral trials due to the averaging of slots, whereas there would be no difference between neutral and invalid trials as each stimulus would receive either one or two slots. The results from Experiment 3 revealed that P_m was vastly increased on valid trials relative to invalid trials, suggestive of participants allocating more slots to the cued stimulus, with *s.d.* reduced on valid trials relative to neutral trials within the range that would be expected from the averaging of slots. Finally, *s.d.* was essentially the same on invalid and neutral trials, which the authors suggest is indicative of resources not being directed to the cued stimulus only. Zhang and Luck go on to state that these results suggest that it is unlikely that a lower resolution representation can be generated using minimal resources, favouring the slots-plus-averaging model over the slots-plus-resources model.

Following from this, Experiment 4 aimed to investigate whether an “all-or-none” encoding process was required to generate robust memory representations. By presenting a masks to interfere with stimulus encoding (110 or 340ms after onset of the stimulus display), Zhang and Luck sought to determine whether low resolution representations can be generated; should generation of such a representation be possible, presentation of the early mask would result in a decrease in *s.d.*. If not, decreasing the onset of the mask would only influence P_m, indicating that an all-or-none process is required to generate representations that are able to survive subsequent sensory input. The results of Experiment 4 revealed that *s.d.* did not differ between masking intervals whereas P_m was substantially reduced (but see van Moorselaar et al., 2015). This led Zhang and Luck to suggest that generation of representations requires this all-or-none process, seemingly ruling out pure resource models and models wherein resources can be allocated flexibly across a fixed number of items (i.e., the slots-plus-resources model). However, later research (see e.g., Bays et al., 2009; Bays & Husain, 2008) countered this by highlighting that performance on the continuous report task could not simply be deconstructed into responses based on a noisy memory representation and random guessing, but that another critical piece of information was required.

4.1.4 Development of the three-component model

Across a series of experiments, Bays and Husain (2008) used tasks wherein participants were tested on their memory for orientation and location. In the orientation tasks, stimuli consisted of oriented arrows, while in the location tasks, stimuli consisted of coloured squares and set size was varied with experiment (Experiment 1 N = 1, 2, 4, or 6; Experiment 2 N = 5). The goal was to report the direction of change in a probed stimulus (i.e., clockwise or counter-clockwise in the orientation task; left or right in the location task). In Experiment 1, a series of conditions relating to whether an eye movement was required were also implemented and controlled via the use of eye-tracking. In the *fixation* condition, participants simply fixated on a cross presented in the centre of the display, while in the *saccade* condition, participants made a saccade to one item in the display after 1,000ms. In two further conditions, after 1,000ms one of the stimuli in the display began to flash; in the *saccade-to-cue* condition this acted as a cue to saccade to the flashing stimulus, whereas in the *fixation-with-cue* condition, this acted as a distractor, with participants required to continue fixating on the cross. Note that the flashing item in each of these conditions did not indicate which stimulus would be probed on a given trial. In Experiment 2, the dynamic allocation of resources was investigated, wherein participants fixated on each of the five stimuli one at a time, with the display extinguished upon saccade to the final item; four of the stimuli were presented surrounding the fifth which acted as the final saccade target and saccade onset was cued with an auditory click.

Results from Experiment 1 revealed that in both the orientation and location tasks, performance was high with set sizes of one and in the absence of eye movement, with performance declining as a function of increasing set size. However, unlike the notable reductions in performance that have been observed in previous studies when set size has increased past a supposed capacity limit (see e.g., Luck & Vogel, 1997; Vogel et al., 2001; Zhang & Luck, 2008), no such pronounced reduction was observed here. Comparing response precision between fixation and saccade conditions, it was found that there were no substantial difference and interestingly, precision was found to decline even at lower set sizes in both conditions; the most substantial reduction was observed between set sizes of one and two, clearly conflicting with previous research suggesting an fixed capacity to vSTM around four items (see e.g., Luck & Vogel, 1997). Given these results, Bays and Husain (2008) suggested that a model wherein a resource is distributed across items, with the precision of internal representations a function of set size, is a more appropriate model relative to a model positing a fixed upper limit to vSTM capacity. Indeed, the finding that in the saccade condition, the target of the saccade was recalled with greater precision when probed than non-targets in both orientation and location tasks. This is further supported by results from the saccade-to-cue and fixation-with-cue conditions, wherein the flashing stimulus was recalled with greater precision when probed relative to when the probe was a non-flashing stimulus. This suggests that additional resources can be allocated to the target of a saccade, but also that covert shifts of attention are sufficient to allocate resources. Results from Experiment 2 revealed that in both the orientation and location tasks, precision was highest for the final stimulus in the display despite this stimulus not being fixated as the display was extinguished upon initiation of the saccade to the final stimulus. This is again indicative of the ability for covert shifts of attention to facilitate resource allocation, with further results showing that there were no substantial differences between the final fixated stimulus (stimulus four) and any of the other previously fixated stimuli. Based on these results, Bays and Husain (2008) stated that given the increased precision for the final stimulus can only survive a single eye-movement, suggesting that resources are reallocated when the focus of attention changes.

Extending this work, Bays et al. (2009) used a continuous report task wherein participants were presented with coloured squares and following a retention interval, were probed with the location of one of the squares. The goal of the task was to report the colour of the probed stimulus using a 360° colour wheel. Performance on this task was assessed across varying set sizes ($N = 1$, 2, 4, or 6) and stimulus display durations (100, 500, or 2,000ms). Critically, Bays et al. stated that while performance on the continuous report task had previously been associated with memory for colour (see e.g., Zhang & Luck, 2008), it also relied on memory for location as well. This led Bays et al. (2009) to extend the two-component mixture model to a *three-component* model which states that performance on continuous report is a probabilistic mixture of; 1) responding to the target value based on a noisy memory representation, 2) random guessing, and 3) responding to a non-target value (i.e., the location component). This model is given formally as:

$$p\left(\hat{\theta}\right) = \left(1 - p_u - p_n\right)\phi_\kappa\left(\hat{\theta} - \theta\right) + p_u\frac{1}{2\pi} + p_n\frac{1}{n}\sum_{i}^{n}\phi_\kappa\left(\hat{\theta} - \theta_i^*\right), \tag{9}$$

where θ, $\hat{\theta}$, p_u, ϕ, and κ are the same as in the two-component model, with p_n representing the probability of a non-target response and θ_i^* representing the non-target feature values (see e.g., Grange & Moore, 2022). Results from Bays et al. (2009) revealed that precision decreased as set size increased, even when set size increased from one to two, similar to the findings from Bays and Husain (2008) and again conflicting with previous research suggesting a fixed upper-limit to vSTM capacity. Crucially, the finding that responses were centred on non-target feature values at a rate greater than what would be predicted by chance and which increased as set size increased, led Bays et al. (2009) to highlight that memory for location was also essential for performance of the continuous report task; clearly, the results from this experiment show that on some trials, memory for location failed, resulting in participants reporting the feature value of one of the non-targets (i.e., a *swap error*).

4.1.5 The role of attention in resource allocation

While the research discussed previously provides some evidence for the role attention in vSTM resource allocation, additional research has also suggested that vSTM resources and attention are synonymous or at least that attention mediates the allocation of vSTM resources (see e.g., Bays & Husain, 2008; Dube et al., 2017; Emrich et al., 2017; Henderson et al., 2020; Huynh Cong & Kerzel, 2021; Kiyonaga & Egner, 2013; van den Berg et al., 2012). Such a view therefore suggests that the control of attention is essential for the appropriate allocation of vSTM resources, but on occasion, attentional control fails; therefore if attention and vSTM resource allocation are intertwined, such failures would lead to sub-optimal resource allocation and reduced task performance. Indeed, a study by Henderson et al. (2020) which investigated how deficits in attentional control as a result of healthy aging influence vSTM resource allocation provide support for such a view. Henderson et al. assessed performance on a continuous report task wherein stimuli consisted of coloured squares ($N = 4$; baseline condition $N = 1$), with probabilistic cues (lines) also presented alongside the stimuli; a total of four cuing conditions were used ($N = 4$, 100% valid; $N = 2$, 100% valid; $N = 1$, 100% or 50% valid). A further *no distractor* condition was implemented wherein a single coloured square was presented with no cue; as such, this meant that non-cued stimuli were only probed in the 50% valid, one cue condition, meaning non-cued stimuli could be ignored in other conditions. After presentation of the stimuli, a probe appeared indicating the location of one of the previously presented stimuli; responses were provided by manipulating a colourwheel to best match the colour of the probed stimulus. Performance was assessed between younger (M = 21.10 years) and older (M = 73.32 years) adults, with the authors hypothesising that response error would be poorer for older adults while also stating that older adults would have difficultly in allocating resources when attentional control demands peak.

Results of this study revealed evidence to show that both younger and older adults were able to utilise the cues to a similar level to flexibly allocate resources, with both groups reporting stimuli with the same validity (i.e., 50% valid, one cue and 100% valid, two cue conditions) with similar levels of error, however error was found to be greater overall for older adults. Additionally, comparison of the 100% and 50% valid one cue conditions showed greater error in the 50% valid condition, suggesting that participants used the cues to prioritise resource allocation, with older adults again having a greater response error than younger adults. In terms of distractors, comparison of performance on the no distractor and 100% valid one cue condition revealed that participants could effectively ignore distractors, with the older adults group again having greater response error. Taken together, the authors stated that these results showed that both groups were able to effectively allocate vSTM resources to relevant information through the use of probabilistic cues. However, fitting of the three-component model (see Bays et al., 2009) revealed that older adults were less precise across each of the conditions employed, with older adults also having a greater guess rate in general. Interestingly, in terms of non-target errors, little evidence was observed supporting a difference between younger and older adults, with the greatest evidence for a difference observed in the 50% valid one cue condition (BF_{10} = 4.03). Further analysis of this condition revealed that while target colour—but not cued colour—significantly predicted response colour in younger adults, both target *and* cued colour were significant predictors of response colour in older adults. This suggests that when probed with a non-cued stimulus, older adults often reported the cued stimulus instead; given that this is the only condition wherein a non-cued stimulus can be probed, this is the condition wherein attentional control demands will be greatest. As such, these results are indicative of attentional control deficits, with Henderson et al. (2020) suggesting that older adults may have found it difficult to inhibit the cued stimuli when a non-cued stimulus was probed.

4.1.6 Implementing switching within continuous report

While the use of probabilistic cuing and retroactive cuing has been used extensively to investigate the role of attentional control in vSTM resource allocation, a further, perhaps more simplistic method for investigation of this role of attentional control is through the use of task switching. Task switching involves either repeating performance of the same task or switching between performance of two or more tasks; a consistent finding is that switching tasks relative to repeating tasks results in a detriment to performance (typically assessed by error rate and response time) referred to as the *switch cost* (see e.g., Grange & Houghton, 2014; Kiesel et al., 2010; Vandierendonck et al., 2010). Critically, the difference between performance of a repetition trial and a switch trial lies in the need to update the *task set*. While definitions vary across researcher, a task set is simply a grouping of processes and representations required to perform a given task (see e.g., Grange & Houghton, 2014 for a discussion of the varying definitions). When repeating a task, the task set can remain the same however, when the task switches, new processes and representations are required to effectively carry out this task. This requirement to update task set forms the basis of many theories which attempt to account for switch costs, including two of the most prominent, the *interference* and *reconfiguration* accounts. The interference account is based on research by Allport et al. (1994) which suggested that when switching tasks, the previously relevant task set interferes with the newly relevant task set and thus disrupts performance (see also Allport & Wylie, 2000; Wylie & Allport, 2000), with such disruption leading to increased RTs and error rates. On the other hand, the reconfiguration account is based on research by Rogers and Monsell (1995) which stated that switch costs are a product of advance reconfiguration of the task set; increased RT on a switch trial reflects the time taken to update the task set while an error reflects a failure of the task set to be sufficiently reconfigured in order to perform the task.

Crucially to the present study however, task sets can be deconstructed into two constituent parts, the *attentional set* and the *intentional set*. The attentional set facilitates identification and selection of information relevant to the current task or goal, while the intentional set facilitates the assignment of stimulus-response mappings to provide appropriate responses for the current task or goal (see e.g., Kopp et al., 2006; Liefooghe & Verbruggen, 2019; Rushworth et al., 2005, 2002). Therefore, as implemented in previous chapters, switching between which feature dimension is relevant on a given trial while the task remains the same will require the updating of the attentional set only. Previous research by Meiran and Marciano (2002) implemented a dimension switch within a same or different judgement task; stimuli were made up of four features, shape, fill, size, and tilt, however shape and fill were the only relevant features, with participants allocated to one of three groups. In the dimension group, the task remained the same throughout but the relevant feature dimension could change at random. In the decision rule group, the relevant feature dimension remained the same but the task could change at random. Finally, in the response mapping group, both feature dimension and task remained the same throughout, but the response mapping could change randomly. The purpose of this study was to investigate the influence of advance reconfiguration through manipulation of the cue-stimulus-interval (CSI; either 170, 470, 1,470, or 2,470ms) on dimension switching. Importantly, results of Experiment 1 revealed that while increasing the duration of the CSI resulted in reductions to the switch costs observed for the decision rule and response mapping groups, this reduction was not found for the dimension group. Such a result not only follows with results obtained in Chapter 3 (Experiment 1a), but also fits well with a previous supposition by Rogers and Monsell (1995) based on the finding that residual switch costs were still observed even when preparation duration was long. The authors stated that while an endogenous control process was required to update the task set, a further, exogenous process, triggered upon stimulus presentation was required to complete reconfiguration, and which was unaffected by manipulations of preparation time.

4.1.7 The current study

The present study aimed to provide a more nuanced insight into the impact of dimension switching on vSTM resource allocation and the role of attentional control in such allocation. Previous chapters have examined the impact of dimension switching on change detection, finding that such switching impacts various measures of performance across a range of experimental manipulations. However, the cause of this disruption cannot be ascertained from change detection data, allowing only for speculative conclusions to be drawn. Through the use of mixture modelling, the parameters returned can be examined across sequencing condition, allowing identification of any changes that may occur between repetition and switch trials, which could provide some information relating to the underlying cause of the dimension switch costs observed in previous chapters. While some previous research has investigated the role of attentional control in vSTM resource allocation (see e.g., Dube et al., 2017; Emrich et al., 2017; Henderson et al., 2020), these studies have examined performance across a single feature dimension. This is clearly atypical of everyday life where stimuli are often multivalent; therefore, the mechanism by which vSTM resources are allocated to relevant information within a multivalent stimulus has not yet been fully investigated. As such, by embedding a continuous report task within a dimension switching design wherein the relevant feature dimension either repeats (e.g., colour–colour) or switches (e.g., colour–orientation), a more detailed picture of the impact of dimension switching on vSTM can be obtained. Through the application of mixture modelling (specifically the three-component mixture model, see Bays et al., 2009), one can determine the impact of dimension switching on precision, the probability of making target and non-target responses, and the probability of guessing.

4.2 Experiment 1: General method

4.2.1 Participants

A total of 50 participants completed Experiment 1. Data collection for this experiment began shortly after restrictions on laboratory-based testing due to the Coronavirus (COVID-19) pandemic were lifted. As such, protocols pertaining to controlling the spread of the virus in the laboratory were put in place, reducing the number of datasets that could be collected in a given time period. Additionally, previous guidance which had placed many teaching sessions online vastly reduced the available participant pool. Participants were aged between 18 and 60 years (inclusive) and self-reported normal or corrected-to-normal visual acuity and normal colour vision. Recruitment was conducted via the SONA Participant Recruitment System at Keele University. Participants were awarded course credit for taking part.

4.2.2 Stimuli and materials

The task used in Experiment 1 was created and run using PsychoPy (Peirce et al., 2019). Stimuli in Experiment 1 consisted of four coloured, oriented circular shapes (see Figure 13) and were presented in the centre of a 24.5 inch display (ASUS ROG Swift™ PG258Q) at a viewing distance of 52cm. Viewing distance was held constant through the use of a chin rest. Four colours were selected at random on each trial from a circular portion of the CIE L*a*b colour space (L = 90, a = 20, b = 38) with a radius of 60. Four orientations were also selected at random on each trial, with the restriction that each of the colours and orientations on a given stimulus display had at least 40° separation. The RGB coordinates of the background colour were 170, 171, 171, with cues and fixations crosses presented in black. The stimuli had a radius of 30 with the bar indicating orientation set to a height of 60. If colour was the relevant feature dimension, the word “colour” was presented as the cue, whereas if orientation was relevant, the word “orientation” was presented. On the probe screen, the locations of three of the stimulus locations were outlined with a black circle, with the location of the fourth, probed stimulus a solid black circle. When colour was relevant, a colour wheel of radius 250 and a width of 35 was presented surrounding the locations of the four stimuli on the probe screen; when orientation was relevant, this wheel was black. If a location was selected on either response wheel, the probed stimulus took on the value selected (e.g., if 180° was selected, the probe would display a stimulus oriented to 180°).

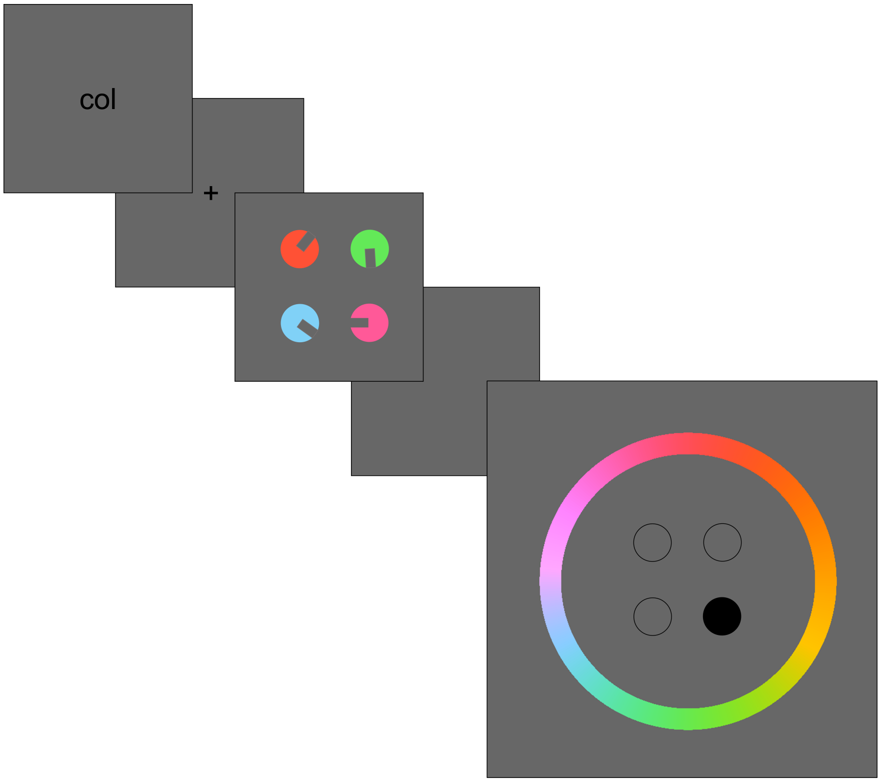

Figure 13
Figure showing example trial procedure in Experiment 1. Note that if orientation was relevant, the wheel presented on the probe display was black.

4.2.3 Procedure

In Experiment 1, participants were presented with four coloured, oriented circular shapes and then attempted to reproduce either the colour or orientation of a probed shape location. The relevant feature dimension could either repeat (e.g., colour–colour; repetition) or switch (e.g., colour–orientation; switch). After demonstrating three trials each for colour and orientation, participants performed 10 colour only trials, followed by 10 orientation only trials. The main experimental section consisted of a total of eight blocks of 50 trials for a total of 400 trials. Self-paced breaks were given every 50 trials. At the beginning of each trial, participants were presented with a fixation cross for 500ms followed by presentation of the cue for 500ms. If colour was relevant, the word "colour" was presented, whereas the word "orientation" was presented if orientation was relevant. After a 500ms cue-stimulus interval, the stimulus display was presented for 200ms. This was followed by a 1,000ms retention interval and subsequent presentation of the probe screen; the probe screen was presented until a response was provided. Participants responded by clicking on the location they believed best matched that of the stimulus in the probed location and pressing spacebar to confirm response; responses were not time-limited. A 500ms intertrial interval followed response.

4.2.4 Model fitting and statistical analysis

The three-component mixture model (see Bays et al., 2009) was fit to the data obtained from Experiment 1 using the R package *mixtur* (see Grange & Moore, 2022). Prior to model fitting, response time (RT) trimming was performed; RTs less than 1,000ms and 2.5 standard deviations above the mean were trimmed per participant, per condition using the R package *trimr* (see Grange, 2022). Following model fitting, participants with a probability of guessing (p_u) greater than 0.5 were removed from further analysis. This is based on previous simulations which showed that when p_u exceeds 0.5, precision as indicated by the parameter κ becomes unstable (see e.g., Grange & Moore, 2022). As such, seven participants were excluded from analysis post model fitting (final $N = 43$). Bayesian paired-samples t-tests were then used to investigate whether the parameters returned by the model fitting (κ, p_t, p_{nt}, p_u) differed across sequencing condition (repetition, switch); this was done using the R package *BayesFactor* (see Morey & Rouder, 2021), using package default priors, namely a Cauchy distribution centred on zero effect with scale parameter $r = 0.707$.

4.3 Experiment 1

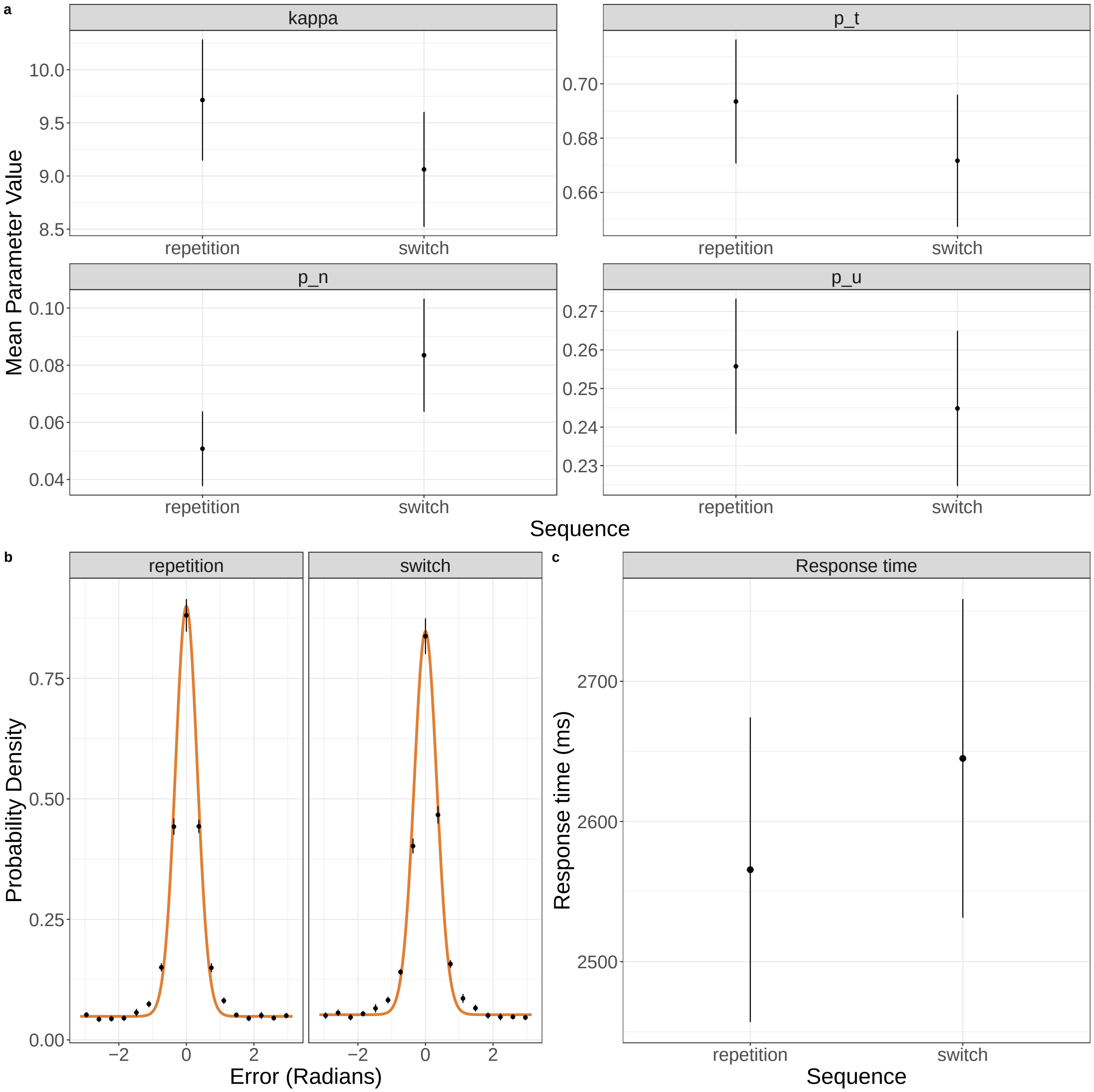

Figure 14

Plots displaying a) mean values for each of the parameters returned by the three-component model, b) participant-averaged response error (points) with overlaid model fit (solid line), and c) mean response time across sequencing condition in Experiment 1. Error bars represent the standard error of the mean.

4.3.1 Results

Means for each of the parameter values returned by the model fit, RTs, and model fit can be found in Figure 14. Analysis of κ parameter values between repetition (M = 9.71, SD = 3.72) and switch (M = 9.06, SD = 3.53) data revealed a BF_{10} = 0.28 indicating moderate evidence in support of the absence of a switch cost. Analysis of p_t parameter values between repetition (M = 0.69, SD = 0.15) and switch (M = 0.67, SD = 0.16) data revealed a BF_{10} = 0.89 indicating anecdotal evidence in support of the absence of a switch cost. Analysis of p_{nt} parameter values between repetition (M = 0.05, SD = 0.08) and switch (M = 0.08, SD = 0.13) data revealed a BF_{10} = 3.79 indicating moderate evidence for the presence of a switch cost; however, note the effect size is small (d = 0.26). Analysis of p_u parameter values between repetition (M = 0.26, SD = 0.11) and switch (M = 0.24, SD = 0.13) data revealed a BF_{10} = 0.20 indicating moderate evidence in support of the absence of a switch cost. Finally, analysis of RT data between repetition (M = 2,565.63, SD = 711.30) and switch (M = 2,644.93, SD = 744.32) revealed a BF_{10} = 151.34 indicating extreme evidence in support of the presence of a switch cost.

Additional analysis on behavioural measures returned from the *mixtur* package were also conducted. A Bayesian paired-samples t-test on precision between repetition (M = 0.75, SD = 0.33) and switch (M = 0.71, SD = 0.35) data revealed a BF_{10} = 0.34 indicating anecdotal evidence in support of the absence of a switch cost. Comparison of mean absolute error between repetition (M = 0.55, SD = 0.21) and switch (M = 0.59, SD = 0.25) data revealed a BF_{10} = 17.32 indicating strong evidence in support of a switch cost. Comparison of resultant vector length between repetition (0.65, SD = 0.15) and switch (M = 0.63, SD =

0.18) data revealed a $\mathrm{BF}_{10} = 1.67$ indicating anecdotal evidence in support of a switch cost. Finally, comparison of bias between repetition ($M = 0.00$, $SD = 0.11$) and switch ($M = 0.00$, $SD = 0.17$) data revealed a $\mathrm{BF}_{10} = 0.17$ indicating moderate evidence in support of the absence of a switch cost.

4.4 General discussion

4.4.1 Summary of results

The purpose of the present study was to gain a more detailed insight into the impact of dimension switching on vSTM resource allocation; while results from previous chapters revealed that dimension switching disrupts change detection performance across a range of measures and experimental manipulations, the underlying cause of such disruption cannot be inferred from binary change detection data. By utilising a continuous report task embedded with a dimension switching design, an element of attentional control was introduced to a task typically devoid of such a requirement, with mixture modelling then allowing for a more nuanced examination of the underlying cause of dimension switch costs. Specifically, the three-component mixture model (see Bays et al., 2009) was fit to the behavioural data, returning the parameters κ, an estimate of precision, p_t, the probability of responding to the target, p_{nt}, the probability of responding to a non-target, and p_u, the probability of random guessing. In general, comparison of the parameter values returned for repetition and switch data revealed little evidence in support of a switch cost; indeed, evidence for the absence of a switch cost was found for κ, p_t, and p_u, with p_{nt} showing weak evidence in support of a switch cost.

4.4.2 Speculation on the notion of resource misallocation

From this study, the most notable result was the finding that dimension switching impacted the probability of non-target responses, with responses to non-targets increasing when a dimension switch occurred relative to when the dimension repeated. While clearly the evidence for this cost is weak, given that it was the only parameter comparison which yielded any evidence in support of a cost, it is pertinent to offer some discussion of the implications; it may be the case that the effect does in fact exist however, the methodology employed in the present study may not have been sufficient to capture it. In previous chapters it was shown that switching between feature dimensions disrupted performance on a change detection task across a range of measures. It was suggested that these dimension switch costs may have been due to a misallocation of resources; when a switch in dimension occurred, attentional control is required in order to ensure that resources are allocated to the relevant feature dimension. However, on occasion, such control may mistakenly allocate resources to the irrelevant feature dimension if the representations (i.e., the attentional set) required to perform the task has not been sufficiently configured. Such misallocation cannot be ruled out at present; given the fact that the non-targets used within the modelling in the present study were taken from the relevant feature dimension, it cannot be determined whether a misallocation of resources to the irrelevant feature dimension occurred. In an ideal situation, the values of the irrelevant feature dimension would be used as the non-target values, allowing for investigation of resource misallocation. However, in the present experiment, this was not possible; while orientation may have influenced responses when colour was relevant (i.e., participants mistakenly report the orientation of a stimulus on the colour wheel), colour could not influence responses when orientation was relevant as responses were provided on a black wheel.

However, if in theory we take the finding of increased non-target responses in the present study to be a true effect, it is possible that dimension switching causes an increase in the misbinding of visual information (i.e., binding of features and location, see e.g., Siegel & Castel, 2018). In simple terms, misbinding causes the feature values from two different stimuli becoming swapped in memory; for example, if presented with multiple coloured stimuli, the colours of two of the stimuli may swap due to misbinding, resulting in an incorrect response if one of these stimuli are probed (i.e., a *swap error* reflected by the p_{nt} parameter in the three-component mixture model, see e.g., Bays et al., 2009). Much previous research has shown that the binding of visual features into an integrated object requires attention (see e.g., Rensink, 2000; Treisman, 1988; Treisman & Gelade, 1980; Treisman & Schmidt, 1982; Wheeler & Treisman, 2002). Feature Integration Theory (FIT, see e.g., Treisman, 1988; Treisman & Gelade, 1980) is perhaps one of the most well-known theories of attention, which states that features from different dimensions can be processed in parallel however, combining these features into an integrated object representation requires focused attention. Building on this, later work by Wheeler and Treisman (2002) led to the suggestion of a two-stage model of vSTM wherein the capacity of each feature dimension is independent and limited to a small number of items, but that binding between feature dimensions is only possible via attention. A number of later studies have also provided support for the view that attention is required for feature binding (see e.g., Hyun et al., 2009; Wan et al., 2020; Zokaei et al., 2014), however, it has also been suggested that feature binding is a passive process, requiring no additional attention (Allen et al., 2012; see e.g., Allen et al., 2006).

Based on the view that the binding in vSTM requires attention, it is therefore plausible that the disruption caused by dimension switching is due to both binding and dimension switching being attentionally demanding processes, resulting in misbinding. In a review of working memory and attention, Oberauer (2019) outlines the notion of attention as a limited resource, positing three differing classes based on what the attentional resource will be utilised for; *storage and processing*, *perception and maintenance*, and *controlling attention*. In the latter of these classes, Oberauer suggests that it is not attending to a stimulus which depletes the resource, but rather the top-down control of attention. For instance, attention may be automatically drawn to several stimuli during a visual search task however, the attentional resource is not required for such occurrences. Rather, controlling the action of allocating attention to stimuli of interest in service of task performance consumes the attentional resource, with such control limited to a single process at any given time. This therefore gives rise to two possibilities concerning the dimension switching design employed in the present study (and those in previous chapters): 1) the process of switching feature dimensions depleted the attentional resource to such a level that little was available for feature binding or 2) incomplete configuration of the attentional set meant that the action of feature binding could not be controlled by attention. In either case, such a view negates the requirement of a memory resource as the source of limitations in vSTM, as on this account the limiting factor concerns the ability to control a given action (e.g., allocation of perceptual attention) rather than a limitation based on the allocation of a memory resource.

4.4.3 On the general absence of dimension switch costs

To reiterate, the discussion of misbinding simply serves to provide some theoretical speculation based on the view that the increased probability of non-target response on switch trials observed in the present study is a true effect. Evidence for this effect in the current study is weak, therefore it is appropriate to reinforce that the previous discussion is purely speculative and not based on substantial evidence. While some studies in previous chapters found substantial costs to performance due to dimension switching in change detection tasks, this is not the case in the present study. Thus, it may be that methodological factors may determine the impact of dimension switching on vSTM performance. Perhaps the most notable difference between the change detection tasks used previously and the continuous report task used in the present study is the vast difference in the task to be performed. In change detection, the memory and test displays are compared with participants then tasked with deciding whether a change occurred. In the continuous report task however, the relevant feature value must be retained in memory and subsequently reproduced using a 360° response wheel. Indeed, it could be suggested that continuous report tasks are somewhat analogous to a retrocued visual search task (see e.g., D. Schneider et al., 2018 for an example of a retrocued visual search task) wherein search takes place following extinction of the stimulus display and requires reproduction of the target value rather than simple identification of the location of the target. As such, it is clear that continuous report tasks are more difficult to perform than change detection, which may have had an influence over the presentation of a switch cost.

It should also be highlighted that the observation of dimension switch costs observed in previous chapters were associated with change detection tasks wherein a constant change in the irrelevant feature dimension occurred. While there are clearly no changes in either relevant or irrelevant feature dimensions within continuous report tasks, the implementation of this irrelevant change ensured that the cue was essential for correct performance of the task. Experiments 1a and 2a in Chapter 2 utilised a change in the relevant feature dimension only, with results revealing little evidence in support of a dimension switch cost. It was suggested that this was due to the cue being non-essential for correct performance of the task; as the change only occurred within the relevant feature dimension, it was plausible that participants monitored the entire stimulus display for any change, removing the need to switch between feature dimensions and negating any potential switch costs that may have arisen. As such, it is possible that participants performing the continuous report task simply attended to the entire stimulus display, with performance reflecting the difficulty of the task rather than the influence of dimension switching.

In line with this rationale regarding the difficulty of the tasks, while responses in change detection are binary (change or no change), there are 360 different responses that can be provided in the continuous report task. Given the high number of possible responses and hence decisions on which value best reflected that presented in the probed location, it is possible that this may have influenced performance on the continuous report task, particularly if a whole-display monitoring technique was adopted. Responses in change detection are based on comparison of noisy representations of the memory and test displays, with increases in set size increasing the noise associated with the decision (see e.g., Palmer, 1990). In contrast, responses in continuous report are based on a noisy representation of the memory display and subsequent selection of one of 360 different response options. As such, the number of response options may contribute to the noise associated with decision making, thus adding to the notion that performance on the task used in the present study reflected task difficulty and not the influence of dimension switching.

In a similar manner, the method of response was also vastly different between the two tasks. In the change detection tasks, responses were provided by pressing one of two keys on a keyboard; first responses were final and providing a response progressed the experiment. In contrast, responses in the continuous report task involved clicking a location on a 360° wheel with a mouse-driven pointer and pressing spacebar to confirm response. Therefore, given the larger number of response options and the ability to select multiple options before confirming an answer in the continuous report task would result in more substantial RTs. Thus, it could be plausible that any impact of dimension switching is short-lived, perhaps only being observable up to a specific duration after stimulus extinction. Indeed, mean RT from the present study showed that participants took well over 2,000ms in each condition (repetition M = 2,565ms; switch M = 2,644ms), while mean RT within the change detection studies in Chapter 2 only ever exceeded 1,100ms once (Experiment 3a, switch).

4.4.4 Recommendations for future research

Based on the limitations discussed previously, future research using similar methods would benefit from developing a continuous report switching paradigm wherein responses on each of the tasks are provided in the same manner. Given that responses in the task used in the present study were either provided on a colour wheel or a solid black wheel, it was not possible to investigate resource misallocation to the irrelevant feature dimension on switch trials. By standardising the method of response across both tasks, the irrelevant feature dimension could therefore be used as non-target feature values for input to the three-component model and thus provide insight into the notion of resource misallocation; if the rate of non-target response increases on switch relative to repetition trials, this would indicate a misallocation of resources to the irrelevant feature dimension on switch trials. It may also be of benefit to investigate the influence of the retention interval on performance. As alluded to previously, it is possible that the influence of dimension switching is short-lived, with previous research also suggesting that the rate of misbinding increases with longer retention interval durations (see Pertzov et al., 2017). However, to dissociate between the influence of the retention interval and the influence of dimension switching on misbinding, this would require a task that includes blocks of pure repetition trials for both colour and orientation, with further blocks containing mixed trials (i.e., colour and orientation trials). Comparison of p_{nt} parameter values between these different sequencing conditions would enable dissociation of the impact of the retention interval and dimension switching on misbinding.

4.4.5 Conclusion

In general, the results of the present study suggest that dimension switching has little impact on precision, probability of target and non-target response, and guess rate as evidenced by comparisons between sequencing conditions for parameter values from the three-component mixture model (see Bays et al., 2009). Some speculative discussion based on the observation of weak evidence in support of a switch cost for non-target response rate suggests that dimension switching may cause an increase in misbinding. However, the evidence is by no means substantial enough to draw any robust conclusions. There are also a number of methodological differences between the present study and those which have previously found dimension switch costs which may indicate the unsuitability of the continuous report task for capturing such costs. As such, the final empirical chapter will adapt a change localisation task and modelling implemented by Shin and Ma (2017) in an attempt to identify the cause of dimension switch costs.

5 Chapter 5: Change localisation

This chapter outlines experiments which made use of a change localisation task embedded within a task switching paradigm to further investigate the impact of switching between different feature dimensions. Computational modelling was also used within this chapter with the aim to provide some form of explanatory account of dimension switch costs.

5.1 Introduction

5.1.1 Overview

A central aspect of the debate surrounding visual short-term memory (vSTM) capacity concerns whether items consisting of multiple features (e.g., colour, orientation) are stored in memory as integrated object representations or as groups of individual features (see e.g., Alvarez & Cavanagh, 2004; Luck & Vogel, 1997; Vogel et al., 2001; Wheeler & Treisman, 2002). An object-based storage view would suggest that when a task-relevant feature of an item is encoded then so too are other, task-irrelevant features on the same item. Evidence for such a view comes from research examining whether the addition of irrelevant features results in reduced performance tasks probing vSTM (see e.g., Luria & Vogel, 2011; Vogel et al., 2001), with findings from these studies revealing no impact on performance, leading the authors to suggest that irrelevant features are not encoded. However, it has also been highlighted that these findings can be accounted for by independent pools of memory resource for each feature dimension (see e.g., Hyun et al., 2009; Shen et al., 2013; Wheeler & Treisman, 2002). More recently, explanatory modelling of the processes of memory encoding and resource allocation by Shin and Ma (2017) using change localisation have enabled a greater insight into the notion of separate resource pools.

Previous chapters have explored the impact of dimension switching on vSTM performance across a range of experimental paradigms. Chapters 2 and 3 showed consistent and sometimes substantial disruptions to change detection performance when required to switch between feature dimension relative to repeating feature dimension, which was suggestive of a failure of attentional control to allocate resources to the relevant feature dimension. While the disruption caused by dimension switching was much more muted in Chapter 4 which used a *continuous report* task, there was weak evidence suggesting that participants responded with the feature of a non-probed stimulus at a greater rate when the feature dimension switched relative to when it repeated. To build upon these findings, the present chapter aims to develop an explanatory account of the impact of dimension switching on vSTM performance. The present study makes use of the methods and modelling (specifically Model 4) applied by Shin and Ma. A change localisation task will be used wherein participants determine the location of a change; crucially, the feature dimension in which the change occurs will either repeat (e.g., colour–colour) or switch (e.g., colour–orientation). By developing variations of the model used by Shin and Ma, it is hoped that this will help identify any differences in encoding and resource allocation that may occur between repeating and switching dimensions, having the potential to highlight the origin of disruption in dimension switching.

5.1.2 Object- vs. feature-based storage

Perhaps two of the most influential papers surrounding the notion of objects as the unit of storage in vSTM were those conducted by Luck and Vogel (1997; outlined in detail in Chapter 1) and Vogel et al. (2001). These studies employed similar methodologies, using change detection tasks with various stimulus types and experimental manipulations. In one such task (Experiment 11, Vogel et al., 2001), participants were presented with coloured and oriented bars (set size $N = 2$, 4, or 6) and instructed to remember the colour of the bars only (*colour condition*), the orientation of the bars only (*orientation condition*), or both the colour and orientation of the bars (*conjunction condition*). Therefore, at a given set size, the memory requirement on the conjunction condition was double that in either the colour or orientation conditions. Results from this experiment revealed that while set size impacted performance, with increasing set size reducing performance, there was no effect of condition or any interaction between set size and condition. This meant that participants were able retain the colour and orientation of a stimulus in memory as easily as retaining the colour or orientation only, which the authors stated provided evidence supporting the notion of an object-based vSTM.

However, the authors then stated that the lack of disruption in the conjunction condition could be explained based on an inability to selectively attend the relevant feature dimensions in the colour and orientation conditions. That is to say, if both features are encoded into memory irrespective of their relevance to the task, there is no need to invoke the notion of integrated object storage in vSTM. To test this, Experiment 12 of Vogel et al. removed information pertaining to the irrelevant feature dimension by making the stimuli all vertical in the colour condition and all black in the orientation condition; the conjunction condition remained the same as that used in Experiment 11. Results of this experiment revealed a similar picture to that found previously; despite removing variation in the irrelevant feature dimension, performance on the conjunction condition was on par with performance on the single-feature conditions. To ensure that participants did not obligatorily encode the irrelevant feature despite the lack of variation, a further experiment (Experiment 13) was conducted which eliminated irrelevant information entirely, with the coloured stimuli having no discernible orientation and the orientation stimuli having no discernible colour; again, the conjunction task remained the same as previous. Results from this experiment again revealed that performance between all three conditions was at a similar level; if performance in the previous experiments had been based on the number of features that were automatically encoded into memory, a notable increase in performance should have been observed between the colour and orientation conditions relative to the conjunction condition given the latter required encoding of double the number of features. Given that this was not the case, the authors stated that these results are in line with the view of integrated object storage in vSTM.

Extending these findings, in Experiment 14 Vogel et al. also tested whether this object-based view of vSTM held for stimuli defined by a greater number of features, namely colour, orientation, size, and the presence or absence of a gap. Thus, in the single-feature conditions, participants only needed to retain information pertaining to the relevant feature (e.g., colour), whereas in the conjunction condition, all four features present on each stimulus needed to be retained. Results from this experiment revealed no effect of condition or interaction between condition and set size; this meant that despite the vast number of features needed to be retained in the conjunction condition, participants were able to do this as well as retaining the four features needed in the single-feature conditions. While the authors stated that this provided further support for object-based storage, they also highlighted that the results could still be explained by independent memory stores for a given feature dimension. To investigate this possibility, a final experiment was conducted wherein stimuli were made up of an inner square of one colour and an outer square of a different colour (i.e., a colour–colour conjunction). Two single-feature conditions were used wherein either the inner squares or outer squares were presented; in the conjunction, both squares were presented. In the conjunction condition, a change occurred in the inner square on 25% of trials, on the outer square 25% of trials, and on the remaining 50% of trials there was no change. In keeping with the results of previous experiments, while set size did influence performance, there was no effect of condition or interaction between set size and condition. As such, the authors stated that participants could retain double the number of colours in the conjunction condition relative to the single-feature conditions, further supporting the view of integrated object storage in vSTM.

This object-based view of vSTM storage fits well with pre-established views on more general short-term memory and attention, with the *chunking* together of separate pieces of information (e.g., digits, letters) prevalent in verbal short-term memory for instance (see e.g., Miller, 1956). Additionally, previous research concerned with the distribution of visual attention (see e.g., Duncan, 1984; Egly et al., 1994; Vecera & Farah, 1994) has suggested that the grouping of features within an integrated object representation confers a benefit to performance. For instance, Egly et al. (1994) used a task wherein participants were cued to a location on one of two parallel rectangles. In the *valid condition*, the target square appeared in the same location as the cue; in the *invalid, same-object condition*, the target appeared in a different location within the same rectangle that had been cued on that trial;

and in the *invalid, different-object* condition, the target appeared in a location on the rectangle which had not been cued on that trial. In both invalid conditions, the target was equidistant from the cue meaning that there was no confound with increased spatial separation. Results from this experiment revealed that performance was faster in the valid condition relative to both invalid conditions, with performance on the invalid, same-object condition being faster than performance on the invalid, different-object condition. Given that the cues and targets were of equal distance in both invalid conditions, these results clearly show that a benefit to performance is conferred when cue and target remained within the same object (however, see also Chen et al., 2020).

However, despite the rather elegant links between the notion of object-based storage in vSTM and other research on short-term memory and attention, later failings to replicate findings from Luck and Vogel (1997; see e.g., Olson and Jiang, 2002; Wheeler and Treisman, 2002; Xu, 2002), as well as research which has highlighted that *stimulus complexity* also factors into vSTM storage (see e.g., Alvarez & Cavanagh, 2004), cast doubt on such a view. For example, Wheeler and Treisman (2002) attempted to replicate the finding from Luck and Vogel (1997) that two values from the same feature dimension (colour) could be retained in memory as easily as a single feature value (see also Vogel et al., 2001). Wheeler and Treisman used bicoloured stimuli with varying arrangements of the inner and outer squares (set size $N = 3$) and found significant differences between performance with each of the bicoloured stimulus types and a control condition wherein three single coloured stimuli were presented (Experiment 1). Wheeler and Treisman (2002) took these findings as evidence showing that performance on this task was limited by the total number of features present on the display, providing evidence against an object-based vSTM storage. However, it should be noted that these experiments differ from the colour–colour conjunction task used by Luck and Vogel (1997) wherein set size varied ($N = 2$, 4, or 6), arrangement of the inner square was not manipulated, and memory for simple features (i.e., squares with a single colour) was tested by presenting either the inner or outer squares separately. As such, Wheeler and Treisman (2002) conducted a further experiment (Experiment 2) which more closely replicated the task used by Luck and Vogel (1997). Results revealed a significant effect of condition, indicating that performance was worse with the single coloured squares relative to the bicoloured squares, providing further evidence against the notion of object-based storage, with Wheeler and Treisman (2002) stating that capacity for colour is limited by the number of colours presented and not the total number of objects.

In addition, work by Alvarez and Cavanagh (2004) highlighted that *visual information* (i.e., the features that make up a given object and that are encoded into memory) plays a role in the amount of information that can be stored in memory at a given time. While the stimuli used by Luck and Vogel (1997) and Vogel et al. (2001) were relatively simple coloured and/or oriented shapes, Alvarez and Cavanagh used coloured squares, letters, Chinese characters, polygons, shaded cubes, and Snodgrass images (simple drawings of objects, e.g., a pen; see Snodgrass and Vanderwart (1980)). The authors estimated information load using a visual search task, assuming that with increased information load the time taken to process these items will increase. Following this, a change detection task was then used to estimate capacity for each of the stimulus types using Pashler's (1988) *k measure*. In the visual search task, a target stimulus was presented in the centre of the display followed by a search array (set size $N = 4$, 8, or 12) containing stimuli from the same class; on 50% of trials, the target was present in the search array. In the change detection task, a memory display was presented (set size $N = 1$, 3, 5, 7, 9, 11, 13, or 15) with stimuli from the same class, followed by a brief blank and then the test display. A change occurred on 50% of the trials. Results revealed that coloured squares had the largest capacity ($N = 4.4$), followed by letters (capacity $N = 3.7$), Snodgrass images (capacity $N = 2.6$), Chinese characters (capacity $N = 2.8$), polygons (capacity $N = 2$), and finally shaded cubes (capacity $N = 1.6$). Alvarez and Cavanagh stated that these results clearly show that the total number of stimuli that can be retained in memory is dependent on the complexity of the stimulus, and not solely on the total number of stimuli present. This, the authors argued, conflicts with theories of vSTM which postulate that capacity is based soley on the number of objects. In addition, the authors also highlighted that it is apparent given the inverse relationship between information load and capacity found in this study that additional capacity must be directed to more complex stimuli, a view which on face value appears more compatible with *resource models* of vSTM (see e.g., Bays et al., 2009; Bays & Husain, 2008; Wilken & Ma, 2004) than *discrete capacity* models (see e.g., Luck & Vogel, 1997; Vogel et al., 2001; Zhang & Luck, 2008).

5.1.3 Resource models of visual short-term memory

A consequence of research which postulates that vSTM stores integrated objects (see e.g., Luck & Vogel, 1997; Vogel et al., 2001) is that the capacity limitation of vSTM was held to be dependent on the total number of items that needed to be retained in memory at a given time. Perhaps the most influential of these papers is that by Luck and Vogel (1997), wherein the authors estimated memory capacity using a change detection task and Pashler's (1988) *k measure*. With varying set size, the authors observed that performance up to set sizes of four did not differ to any great extent however, once set size exceeded four, they found systematic reductions in performance; therefore, Luck and Vogel took this as evidence to suggest that vSTM is limited in capacity to approximately four items. However, a later study by Wilken and Ma (2004) challenged this view, employing both change detection tasks (Experiments 1–6) and continuous report tasks (Experiments 7–9). In the latter of these tasks, participants were instructed to retain the feature values for presented stimuli in memory (set size $N = 2$, 4, 6, or 8 for the colour and spatial frequency tasks; 2, 3, 4, or 5 for the orientation task); the location of a single stimulus was then probed with participants required to report the feature value they believed best matched the stimulus in that location. This was achieved by clicking on a 360° colour wheel in the colour task or by manipulating a probe Gabor patch using the left and right arrow keys in the spatial frequency and orientation tasks. By calculating the angular deviation between participant response and the true value of the target, an estimate of how precisely the probed feature was held in memory could be obtained. Findings from this experiment revealed that as set size increased, the precision of participant response decreased (i.e., responses became further from the target value). Wilken and Ma took this as evidence against a capacity limitation of vSTM defined by the total number of items that could be retained simultaneously, arguing that if such a view was correct, precision on the continuous report task should have been excellent until memory capacity was full, at which point a substantial reduction in precision should be observed. Rather, the authors stated that the supposed upper-limit to vSTM capacity espoused by previous research (see e.g., Luck & Vogel, 1997; Vogel et al., 2001), may simply be artefactual, reflecting the increase in noise associated with encoding and storage as set size increases.

However, later work by Zhang and Luck (2008) revivified the notion of a fixed-capacity vSTM through a series of experiments using continuous report with coloured squares and varying set sizes (Experiment 1 set size $N = 3$ or 6; Experiment 2 set size $N =$ 1, 2, 3, or 6; Experiment 3 set size $N = 4$; Experiment 4 set size $N = 3$). From the data obtained from these experiments, Zhang and Luck were able to extract two parameters using standard estimation, P_m, representing the probability that an item was held in memory and $s.d.$, representing the precision at which the probed item was represented in memory. Results from Experiment 1 revealed that when set size doubled from three to six, P_m was found to have halved whereas no notable difference was observed in $s.d.$. The authors took this as initial evidence in support of a model of vSTM wherein only a small number of high-resolution item representations can be stored at a given time, suggesting that vSTM contained a fixed number of slots, each able to retain a single item representation at a high resolution (i.e., the *slots model*). Subsequent experiments (Experiments 2 & 3) aimed to examine predictions of two distinct versions of the *slots model.* The models in question were the *slots-plus-resource* and *slots-plus-averaging* models. On the slots-plus-resources model, both slots and resources are held as having an influence over performance, with resources allocated to an item within a slot to generate a robust representation. Conversely, the slots-plus-averaging model states that when set size does not exceed the capacity limit, items can be stored across multiple slots, thus increasing the representational strength of that item; responses are based on an average of the representations stored across multiple slots. Zhang and Luck stated that crucially, on both models, while set size remains below the capacity limit, precision will be improved relative to when set size exceeds capacity. Indeed, results from Experiment 2 track with this, revealing that precision was reduced as set size increased from 1–3 and then plateaued at set sizes of six, whereas P_m was found to decline as set size increased from 1–3, with a marked reduction at set sizes of six.

While both models were able to account for these results, Experiment 3 utilised probabilistic cuing in an attempt to further examine their predictions; on valid trials (70% of trials), a cue was presented simultaneously with the stimulus display indicating a single stimulus that would then be probed. On invalid trials (10%), one of the non-cued stimuli were probed, with neutral trials making up the remainder wherein all four stimulus locations were cued. Thus, the slots-plus-resources model would predict that the majority of resources will be allocated to the cued item which will result in a notable difference in precision between valid, invalid, and neutral trials, whereas P_m will be mostly unaffected. In contrast, the slots-plus-averaging model would predict that the majority of slots will be allocated to the cued item resulting in a notable difference in P_m on valid and invalid trials, with a difference in $s.d.$ only observable between valid and neutral trials due to averaging. Results from this experiment revealed a larger P_m on valid relative to invalid trials, with $s.d.$ reduced on valid relative to neutral trials, in line with predictions made by the slots-plus-averaging account. Zhang and Luck took this as evidence against a resource model of vSTM, stating that these results appear to show that low-fidelity representations cannot be generated by allocating a minimal amount of memory resource. Extending these results, Zhang and Luck further investigated the possibility that low resolution representations can be generated by employing a visual mask, arguing that if representations become more stable over time, an early mask (110ms) will reveal the presence of low resolution representations (i.e., reduced *s.d.*); conversely, if the generation of robust representations requires an "all-or-none" process, masking would only impact P_m. Presentation of the early mask was found to influence P_m only (however, see van Moorselaar et al., 2015), leading the authors to suggest that low resolution representations are not sufficient to survive further visual input, providing further evidence against resource models of vSTM.

Despite this work by Zhang and Luck (2008) seemingly providing substantial evidence ruling out resource models of vSTM, subsequent research by Bays and colleagues (see e.g., Bays et al., 2009; Bays & Husain, 2008) highlighted a number of issues surrounding the way in which these conclusions were reached. Notably, Bays and Husain (2008) highlighted that Zhang and Luck (2008) did not control for eye movements which, based on the results from Bays and Husain (2008), play a crucial role in biasing precision in favour of the targets of a fixation. In Experiment 2, Bays and Husain presented participants with five stimuli (coloured squares in the location task; coloured, oriented arrows in the orientation task), requiring each to be fixated in turn; successful fixation was indicated by an auditory click. Once a saccade to the final target was detected, the screen was blanked and a subsequent display probing a random stimulus from the initial display, either displaced horizontally (location task) or rotated (orientation task), with participants required to indicate the direction of the displacement or rotation. Findings from this experiment revealed that when the final stimulus in the display was probed (the non-fixated stimulus), performance was highest in both the location and orientation tasks. There were also no significant differences found between performance on any of the previously fixated stimuli, nor did fixation duration have any influence on performance. Bays and Husain took this as evidence to suggest that the high degree of precision observed for a saccade target can only survive a single eye-movement. This not only suggests that covert shifts of attention are sufficient to generate high resolution item representations, but also that resources appear to be recovered from a previously fixated item and redistributed to a new target.

Extending this work further, Bays et al. (2009) used a continuous report task almost identical to that used by Zhang and Luck (2008) with the exception of differing set sizes ($N = 1, 2, 4,$ or 6) and varying display durations (100, 500, or 2,000ms); Bays et al. also monitored eye movements during performance of the task. The purpose of this experiment was to investigate precisely what is required when performing the continuous report task; while the continuous (colour) report task had previously been thought of as simply a test of memory for colour, Bays et al. (2009) highlighted that memory for location was also a crucial requirement of the task. A mixture model developed by Zhang and Luck (2008; the *two-component model*, given formally below) on the basis of their results stated that responses on the continuous report task were a probabilistic mixture of responding to the target based on a noisy memory representation and random guessing:

$$p\left(\hat{\theta}\right) = \left(1 - p_u\right)\phi_\kappa\left(\hat{\theta} - \theta\right) + p_u\frac{1}{2\pi}, \tag{10}$$

where θ represents the target value, $\hat{\theta}$, participant response (in radians), p_u, the probability of random guessing, and ϕ, the probability density of the von Mises distribution which is the circular analogue of the normal distribution with a mean of zero and standard deviation κ (see e.g., Grange & Moore, 2022). However, such a model fails to account for the influence of memory for the location of a given stimulus, instead collating all failures of memory for location as random guessing. Indeed, results from the experiment conducted by Bays et al. (2009) revealed that as set size increased, precision was reduced, even as set size increased from one to two. More crucially however, responses were found to be centred on non-target feature values more often than would be expected based on chance performance alone, with this also increased with set size increases. Thus, not only do these results track with that found previously by Wilken and Ma (2004) in terms of reductions in precision as set size increases even below the supposed capacity threshold, but also that memory for location is a crucial component for successful performance of the continuous report task. As such, Bays et al. (2009) extended the two-component model developed by Zhang and Luck (2008) to also account for location (the *three-component model*):

$$p\left(\hat{\theta}\right) = \left(1 - p_u - p_n\right)\phi_\kappa\left(\hat{\theta} - \theta\right) + p_u\frac{1}{2\pi} + p_n\frac{1}{n}\sum_{i}^{n}\phi_\kappa\left(\hat{\theta} - \theta_i^*\right), \tag{11}$$

where θ, $\hat{\theta}$, ϕ, and κ the same as in the two-component model, with p_n representing the probability of a non-target response and θ_i^* $(\theta_1^*, ..., \theta_N^*)$ representing the non-target values.

5.1.4 Modelling resource allocation and encoding

While the development of the two- and three-component models has led to significant advances in our understanding of vSTM as well as the development of further models (see e.g., Bays, 2014; Schurgin et al., 2020; van den Berg et al., 2012), these mixture models only provide a descriptive account of performance. While descriptive models provide a description of a given behaviour or function, they do little in terms of *explaining* the underlying causes of the behaviour or function. Explanatory models on the other hand offer an explanation of the way in which a given behaviour or function occurred, allowing for identification of processes of interest (e.g., resource allocation) or, more critically for the present research, processes disrupted by a given variable (e.g., dimension switching). More recently, Shin and Ma (2017) developed several explanatory models which allow for a closer examination of the processes of resource allocation and memory encoding in an attempt to determine the nature of vSTM storage (i.e., features or objects). In total, Shin and Ma developed six models, three wherein memory resources were shared between feature dimensions (the *shared resource* models; Models 1–3) and three wherein each feature dimension had an independent pool of resources (the *independent resource* models; Models 4–6). Additionally, the irrelevant feature dimension was encoded in two of the models from each class, with this dimension either being ignored or influencing decisions, with a third model from each class not encoding the irrelevant feature dimension at all. Each of these models were fit to data obtained from *change localisation* tasks wherein participants were presented with coloured and/or oriented stimuli (set size $N = 4$ or 8) and instructed to respond with the location they believed a change occurred. In *Condition A*, stimuli were either coloured or oriented, in *Condition B*, stimuli were coloured and oriented with both feature dimensions being relevant (i.e., a change could occur in either colour or orientation), in *Condition C*, stimuli were coloured and oriented, with one relevant and one irrelevant feature dimension, with *Condition D* identical to Condition C with the exception that a change also occurred in the irrelevant feature dimension. Prior to discussing the results of Shin and Ma (2017), it would first be pertinent to outline the steps involved with the model, encoding, inference, and response probability. In the interest of clarity, the discussion will concerned with stimuli defined by a single feature dimension.

Encoding. On each memory display, feature values of stimuli are represented by a vector of values θ $(\theta_1,\ldots,\theta_N)$ with N representing set size. In Shin and Ma, each θ_i was drawn from a uniform distribution. Between the memory and test displays a single stimulus chosen at random changes, denoted L (location), with the magnitude of that change drawn from a uniform distribution and denoted Δ. Stimuli on each test display are represented by φ $(\varphi_1,\ldots,\varphi_N)$, where $\varphi_i = \theta_i$ with the exception of the changing stimulus which is $\varphi_L = \theta_L + \Delta$. On both memory and test displays, memory of each stimulus will be noisy, with such noise independent across stimulus displays and stimuli; on the memory display, this is represented by the measurement vector x_i $(x_1,\ldots,x_N)$, while on the test display this is represented by y_i $(y_1,\ldots,y_N)$. Measurement vectors for both stimulus displays are drawn from a von Mises distribution such that the measurement vector for the memory display is given by:

$$p(x_i|\theta_i) = \frac{1}{2\pi I_0\left(\kappa_{x,i}\right)} e^{\kappa_{x,i} cos(x_i - \theta_i)}, \tag{12}$$

while the measurement vector for the test display is given by:

$$p(y_i|\varphi_i) = \frac{1}{2\pi I_0\left(\kappa_{y,i}\right)} e^{\kappa_{y,i} cos(y_i - \varphi_i)}, \tag{13}$$

where most crucially, $\kappa_{x,i}$ and $\kappa_{y,i}$ are concentration parameters (see e.g., Grange & Moore, 2022), with I_0 representing the modified Bessel function of the first kind of order zero.

The precision at which a feature on a given stimulus was encoded was modelled based on the amount of memory resource allocated to that stimulus and feature. For a given stimulus location (i) in a display (J'_{array}), the memory resource allocated to this stimulus was free to vary across stimulus displays (memory and test), locations, and trials. Therefore, $J'_{array,i}$ was drawn from a gamma distribution with mean $\bar{J}'$ and scale parameter τ independently across stimulus displays, stimuli, and trials such that:

$$\left(\bar{J}', \tau\right) = J_{array,i} \tag{14}$$

Following this[8], the precision values are then converted into concentration parameters ($\kappa_{array,i}$) via:

$$J_{array,i} = \kappa_{array,i} \frac{I_1\left(\kappa_{array,i}\right)}{I_0\left(\kappa_{array,i}\right)}, \tag{15}$$

where I_1 represents the modified Bessel function of the first kind of order one.

Inference. During the inference stage, the vector representing noisy memory for the initial memory display is compared with the vector representing noisy memory for the test display. The model then determines the location in which a change occurred by selecting the

---

[8] Prior to this step, Shin and Ma (2017) multiply $J'_{array,i}$ by the bottom-up factor α such that $J_{array,i} = \alpha J'_{array,i}$. However, given that Shin and Ma provide little information regarding the α parameter and in the interest of simplicity, it was held constant at one in the modelling within the current study, negating any influence it may have had.

location with the greatest difference between the memory and test displays. Given the noisy measurement vectors x and y which represent the memory and test display respectively, the posterior probability of a change in a given stimulus $p(L|x, y)$ independent of other locations is computed such that:

$$d_L = \frac{I_0\left(\kappa_{x,L}\right) I_0\left(\kappa_{y,L}\right)}{I_0\left(\sqrt{\kappa_{x,L}^2 + \kappa_{y,L}^2 + 2\kappa_{x,L}\kappa_{y,L} cos\left(x_L - y_L\right)}\right)}, \tag{16}$$

where d_L represents the likelihood ratio of the decision variable, with the location (i.e., stimulus) in which d_L is highest being chosen.

Response probabilities. Monte Carlo simulations were used to calculate the probability of a correct response for combinations of parameters (ω) and a given change magnitude ($p(correct|\Delta, \omega)$). 1,280 samples of $J_{x,i}$ and $J_{y,i}$ were generated for the memory and test displays using the same procedure as that detailed in the encoding section. Concentration parameters ($\kappa_{x,i}$ and $\kappa_{y,i}$) were then computed from these measurement vectors using Equation 15, with measurement vectors x and y then drawn from von Mises distributions using the given concentration parameters. Means for each of the von Mises distributions was held constant at zero with the exception of one measurement from the y vector which had a mean of Δ. The decision rule was then assessed, with correct responses providing an estimate of the probability of a correct response for a given combination of parameters and change magnitudes.

In Experiments 1 and 2, Shin and Ma ruled out all three models wherein resources are shared across feature dimensions. Experiment 1 tested Models 1 (irrelevant dimension encoded and influences decision) and 2 (irrelevant dimension encoded but ignored) in Conditions A and C. Both models predicted lower performance on Condition C (bivalent stimuli, one relevant feature, one un-changing irrelevant feature) relative to Condition A (univalent stimuli) given that they share resources between feature dimensions. Analysis of behavioural data revealed no effect of condition, going against the predictions of the models. Formal model comparison between Models 1 and 2 and Models 5 and 6 (which would also predict lower performance in Condition C relative to Condition A) using Akaike's Information Criterion (AIC) revealed higher values for Models 1 and 2 (lower AIC value determines best model), leading to their rejection. Subsequently, Experiment 2 ruled out the last remaining shared resource model, Model 3; this experiment compared performance on Condition B (two relevant features) and Conditon C. Given resources are shared on this model and irrelevant features are not encoded, Model 3 would predict that performance on Condition C would be markedly better than performance on Condition B given there is no need for resources to be shared to the irrelevant feature dimension; given both features are relevant in Condition B, the resources would be shared across features. Again, analysis of the behavioural data revealed no effect of condition, rejecting the predictions made by Model 3; subsequent formal model comparison revealed the AIC for Model 3 was greater than that for both Models 5 and 6.

Previous work by Shin and Ma (2016) had already eliminated the possibility of Model 6 wherein each feature dimension had a separate pool of resources but where the irrelevant feature dimension was not encoded. They used change localisation tasks (Experiments 1 and

2) wherein participants were presented with coloured, oriented stimuli (set size $N = 4$) and tasked with indicating the location of a change (both colour and orientation changed on each trial). The relevant feature dimension remained constant throughout the first 30 trials, followed by a trial wherein the change in the irrelevant feature dimension was also probed following a probe of the relevant feature change (e.g., in Experiment 1, the colour change was also probed on the 31^{st} as the orientation change was probed on the initial 30). A total of 31 trials were used with 600 participants providing analysable data. Results from Experiments 1 and 2 revealed that performance was poorer on the online version (Mechanical Turk) of the task relative to the lab-based study and that performance on the irrelevant trial was not significantly different from chance performance. The authors stated that the chance performance observed in these experiments could be the result of a number of factors, such as the requirement to first respond to the relevant feature before responding to the irrelevant feature. This would therefore result in a delay, potentially causing degradation of any representation of the irrelevant feature held in memory. Shin and Ma also highlight a study by Marshall and Bays (2013) which stated that while irrelevant features are automatically encoded when attending a relevant feature on a multivalent stimulus, there is no requirement to maintain this irrelevant feature in memory. Therefore, Shin and Ma (2016) argued that any representation of the irrelevant feature may have been discarded upon trial completion. Furthermore, the authors also suggest that irrelevant features may only be encoded when set size is lower however, what appears to be the most likely explanation (and indeed, the least difficult to verify) is that participants may not have been aware that they were required to respond differently. Shin and Ma found that on trial 31 in Experiments 1 and 2, 36% and 41% of participants respectively, responded to the irrelevant

feature change with the same response given for the relevant feature change.

To combat these issues, Shin and Ma (2016) conducted a further two experiments making use of the continuous report task; participants were briefly (100ms) presented with a single coloured, oriented stimulus and tasked with reporting the colour or orientation of the stimulus after a retention interval (1,000ms). As in Experiments 1 and 2, the relevant feature dimension remained constant throughout the first 30 trials; on trial 31, the irrelevant feature dimension was instead probed of the relevant feature dimension (e.g., in Experiment 3, colour was probed instead of orientation as orientation had been probed on the initial 30 trials). Therefore, a total of 31 trials were used and a further 600 participants naive to the irrelevant manipulation were recruited. Given that Model 6 (Shin & Ma, 2017) predicts the irrelevant feature dimension is not encoded, response distributions on trial 31 would be uniform as participants would simply be guessing. Results from Experiments 3 and 4 of Shin and Ma (2016) revealed that this was not the case; while error rate certainly increased for the irrelevant feature dimension relative to the relevant feature dimension (although the difference in trial numbers should also be taken into account when comparing error distributions), the response distributions were significantly different from uniform[9].

In Experiment 3 of Shin and Ma (2017), the remaining two models were tested, Models 4 and 5; both models assume independent resource pools for each feature dimension and while both also assume the irrelevant feature dimension is encoded, the irrelevant feature dimension is taken into account during the decision process in Model 4 and ignored

---

[9] Shin and Ma (2016) note that there is a substantial difference between performance for colour and orientation (particularly noticeable from the response distributions on p. 6) reinforcing the notion that colour is a more salient or perhaps, as stated by Shin and Ma, more ecologically necessary.

in Model 5. These models were tested in Condition C (bivalent stimuli, one relevant feature, one unchanging irrelevant feature) and Condition D (bivalent stimuli, one relevant feature, one irrelevant feature that could change). Six total sessions were conducted in Experiment 3, three wherein colour was relevant and three wherein orientation was relevant: two sessions of Condition C, two sessions of Condition D wherein one change occurred, and two sessions of Condition D wherein two changes occurred. Results from this study failed to provide evidence supporting one model over the alternative; while Model 5 would predict similar performance across all three conditions as the irrelevant feature dimension is ignored during decision making (the only difference between Models 4 and 5), Model 4 predicts that performance would be highest on Condition C, followed by one-change Condition D, and finally, two-change Condition D. However, analysis of behavioural data did not reveal any support for Model 4 but based on the experiment reported in Shin and Ma (2016), the authors highlighted the potential for the irrelevant feature to be stored in memory at a low resolution. Critically, calculation of AIC for Models 4 and 5 revealed that Model 4 was preferred however, Shin and Ma (2017) highlight that they were unable to determine with any certainty that the irrelevant feature dimension is considered during the decision process, going on to state that if it is considered, it has little influence on decisions as it is encoded at such a low resolution.

5.1.5 The current study

The primary purpose of the present study is to investigate how dimension switching impacts the resource allocation and memory encoding processes by adapting the methods employed by Shin and Ma (2017). Previous chapters made use of tasks which probe vSTM (e.g., change detection, continuous report) embedded within dimension switching designs to investigate the role of attentional control in the allocation of vSTM resources. In these experiments, participants performed the same task (e.g., change detection), while the relevant feature dimension either repeated (e.g., colour–colour) or switched (e.g., colour–orientation). These tasks were based on more traditional *task switching* experiments wherein participants switch between two (or more) tasks; a consistent finding from the task switching literature shows that when there is a requirement to switch between tasks, performance is reduced relative to when the task repeats, a detriment referred to as the *task switch cost* (for comprehensive reviews on task switching, see Grange & Houghton, 2014; Kiesel et al., 2010; Vandierendonck et al., 2010). While findings from the initial change detection experiments (see Chapters 2 & 3) showed a notable impact of dimension switching—which was not attenuated by manipulations of trial timings—use of the continuous report task and mixture modelling in Chapter 4 revealed little evidence to suggest dimension switching impacted performance. The present work serves to further these findings by introducing a dimension switching element into the change localisation task; thus, on a given trial, the relevant feature dimension may repeat (e.g., colour–colour) or switch (e.g., colour–orientation). Crucially, Model 4 of Shin and Ma (2017) was adapted to develop six models; while in some models, parameters were fixed across sequencing condition (repetition, switch), in others they were free to vary across sequencing condition. Therefore, should differences in the model parameters between repetition and switch trials be observed, this would indicate a source of disruption as a result of dimension switching. For instance, differences in $\bar{J}$ and/or τ parameter values between sequencing conditions would indicate that dimension switching impacts the precision at which the relevant features are encoded into memory, while differences in p would indicate an impact of dimension switching on resource allocation, specifically if higher p values are observed on switch trials, this would suggest a misallocation of resources to the irrelevant feature dimension.

5.2 Experiment 1: General method

5.2.1 Participants

A total of 40 participants took part in Experiment 1a and a further 40 participants took part in Experiment 1b. Participants were aged between 18 and 60 years (inclusive) and self-reported normal or corrected-to-normal visual acuity and normal colour vision. Participants who took part in Experiment 1a were prevented from participating in Experiment 1b. Participation was again limited to those within the United Kingdom and United States of America. Recruitment was conducted via the use of Prolific (prolific.co) and participants were paid a small fee for taking part.

5.2.2 Stimuli and materials

Both experiments were created and run using Gorilla Experiment Builder (Anwyl-Irvine et al., 2020) with participation restricted to those using desktop or laptop computers only. The make-up of the tasks used in Experiment 1a and 1b were almost identical. Stimuli in both experiments consisted of four coloured, oriented circular shapes (see Figure 15). Colours were selected at random on each trial from a circular portion of the CIE L*a*b colour space (L = 70, a = 20, b = 38) with a radius of 60. Four orientations were also selected at random on each trial. In both experiments, a change occurred on every trial, with the magnitude of a change in Experiment 1a being either 5, 15, 30, or 45°, while in Experiment 1b the magnitude of a change was either 5, 30, 60, or 90°; both the changing stimulus and change magnitude was selected at random on each trial with equal probability. The RGB coordinates of the background colour were 127, 128, 128, with cues and fixation crosses presented in white. If colour was the relevant feature dimension, the word "col" was presented as a cue, whereas if orientation was the relevant feature dimension, the word "ori" was presented.

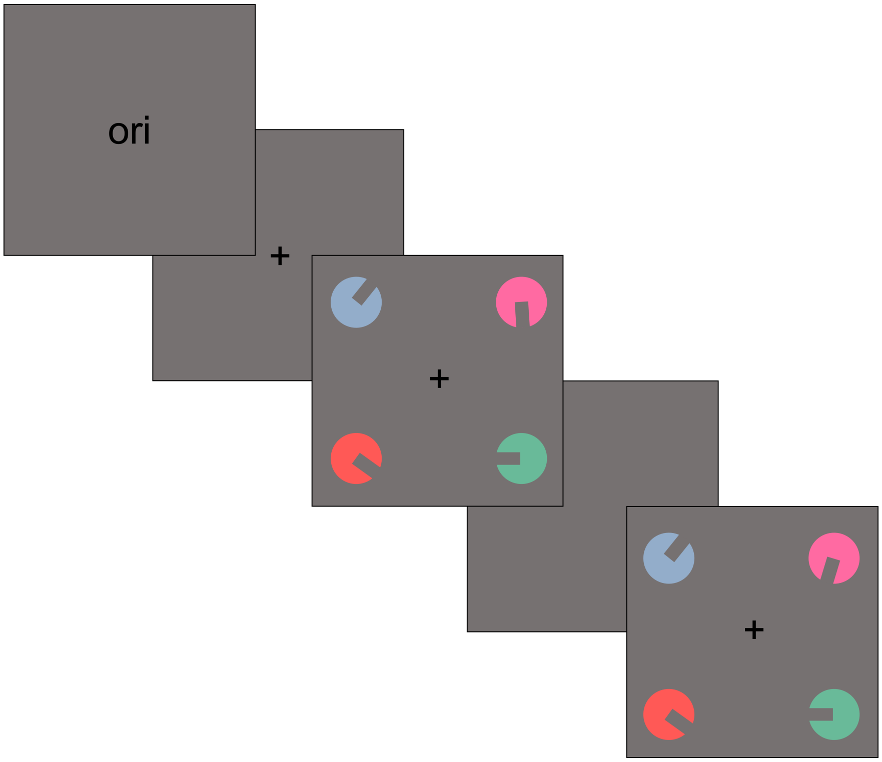

Figure 15
Figure showing example trial procedure in Experiment 1. Note that the same stimuli were used in both Experiment 1a and 1b with the only difference between experiments being the change magnitudes used. The change in this example occurred in the top-right stimulus.

5.2.3 Procedure

In both experiments, participants were presented with four coloured, oriented circular shapes with the goal of the task to determine in which of the four shapes a change occurred. Crucially, the change could either occur in colour or orientation on a given trial meaning that the relevant feature dimension could repeat (e.g., colour–colour) or switch (e.g., colour–orientation). A change occurred on every trial, however the magnitude of the change varied across trials. The practice section of the experiments consisted of 10 colour only trials, 10 orientation only trials, and 20 mixed trials, with the main section of each experiment consisting of a total of eight blocks of 50 trials for a total of 400 trials. Self-paced breaks were given every 50 trials. At the beginning of each trial, a fixation cross was presented for 250ms followed by presentation of the cue for 500ms; if colour was relevant, the word "col" was presented, while if orientation was relevant, the word "ori" was presented. This was followed by a 250ms cue-stimulus-interval and then subsequent presentation of the memory display for 200ms. A retention interval of 1,000ms was then followed by presentation of the test display for 200ms. Upon extinction of the test display, participants were able to provide a response. Response was provided by pressing one of four keys: if the change was believed to have occurred in the top-left, response was provided with the "A" key; top-right, response was provided with the "K" key; bottom-left, response was provided with the "Z" key; bottom-right, response was provided with the "M" key. Responses were not time-limited and were followed by a 500ms intertrial interval.

5.2.4 Model fitting and statistical analysis

Six variants of Model 4 as implemented in Shin and Ma (2017) were used in the model fitting for both experiments; these are named the *full model*, *fixed model*, *fully free model*, *free j-bar model*, *free tau model*, and the *free p model.* All models had some variation of the parameters $\bar{J}$, τ, and p, with variation based on whether the parameters were free to vary across sequencing condition (repetition and switch) and feature dimension (colour and orientation). The p parameter relates to the amount of resource allocated to the irrelevant feature dimension; as such, it is not included in Shin and Ma's single feature dimension model discussed in the Introduction. The p parameter is initially drawn from a random uniform distribution and in simple terms, constricts the precision at which the irrelevant feature values represented, with the model fit routine subsequently finding the optimal p values to implement. Thus, if colour is relevant for example, precision would be determined by:

$$\left(\bar{J}_{col}, \tau_{col}\right) = J_{col}, \tag{17}$$

whereas orientation would be determined by:

$$\left(p\bar{J}_{ori}, \tau_{ori}\right) = J_{ori} \tag{18}$$

The parameter composition for each of the models is as follows: *full model*, $\bar{J}$, τ, and p could vary across sequence, but only $\bar{J}$ and τ could vary across dimension; *fixed model*, no parameter could vary across sequence and only $\bar{J}$ and τ could vary across dimension; *fully free model*, all parameters could vary across sequence and dimension; *free j-bar model*, $\bar{J}$ could vary across both sequence and dimension, τ could vary across dimension only, and p was fixed across both sequence and dimension; *free tau model*, $\bar{J}$ could vary across dimension only, τ could vary across both sequence and dimension, and p was fixed across both sequence and dimension; *free-p model*, $\bar{J}$ and τ could vary across dimension only, while p could vary across sequence only.

For each model, a genetic algorithm was used to fit the model to the experimental data of each individual participant separately. Genetic algorithms are conceptually based on evolutionary genetics, wherein a population evolves over generations, with reproduction based on the fitness of a mate (see e.g., Lewandowsky & Farrell, 2011, pp. 90–95). In the present study, 256 parameter sets (i.e., the initial population) were generated for a given model. Each of these parameter sets were then used as input to simulate the given model, with 5,000 trials simulated per parameter sets. For each of the 256 parameter sets, log-likelihood was calculated indicating the quality of the model fit. Following this, noise was added to the initial parameter values to generate a new generation of parameters, which were then used as input to simulate the model in question; once complete, log-likelihood was calculated for each of these parameter sets. The initial population and the new generation were then bound and the best fitting parameter values (determined by log-likelihood) were retained, with population size reducing with each generation. Thus, the updated population size (α) would be given as $\alpha = \sigma\beta$, where β represents the previous population size, and σ the reduction rate which was set to 0.98 in the present study. Prior to noise being added to the remaining parameter values, the level of noise was manipulated, reducing with each generation to narrow in on the optimal parameter values. This process repeated for a total of 128 generations with the remaining parameter sets being those providing the best fit based on the model fit routine in question. Akaike's Information Criterion (AIC)—a method for estimation of the quality of a model based on prediction error which is penalised by the total number of parameters—was calculated for each of the models and is given by:

$$AIC = -2\widehat{LL} + 2p, \tag{19}$$

where LL represents the log-likelihood of the model fit and p the number of parameters in the model. The AIC for each model was calculated for each participant with the overall model AIC calculated by summing participant-level AIC, with the preferred model being that with the lowest AIC value overall.

Bayesian ANOVAs were used to examine the impact of sequencing condition (repetition, switch) and change magnitude on accuracy and RT. Further Bayesian ANOVAs were used to investigate the impact of sequencing condition and feature dimension on parameter values within the full, fully-free, free j-bar, free-tau, and free-p models. In instances where a parameter only varied across either sequencing condition or feature dimension, Bayesian paired-samples t-tests were used. All Bayesian analysis was conducted using the R package *BayesFactor* (see Morey & Rouder, 2021) and used package default priors (a Cauchy distribution centred on zero effect with scale parameter $r = 0.707$). Prior to response time (RT) analysis, RTs slower than 150ms and 2.5 standard deviations above the mean were trimmed using the *trimr* package (see Grange, 2022). Given the expectation that performance would be poorer at smaller change magnitudes (i.e., 5°, 15°), no trials were removed on the basis of errors, nor were any participants excluded from analysis based on not performing significantly different than chance.

5.3 Experiment 1a

The aim of Experiment 1a was to provide an initial look at how dimension switching may impact resource allocation and memory encoding processes. This was achieved through modifying the change localisation task used by Shin and Ma (2017) to also include trials wherein the dimension repeated (e.g., colour–colour) or switched (e.g., colour–orientation). In addition, six models were fit to the behavioural data obtained from the change localisation task.

5.3.1 Results

Behavioural analysis. See Table 3 for BFs for all models in the accuracy and RT analysis in Experiment 1a. A Bayesian ANOVA examining the effects of sequencing condition (repetition, switch) and change magnitude (5, 15, 30, 45) on accuracy data revealed the model containing the main effect of change magnitude only as the best model (BF_{10} = 1.42e+36), with the data being 4.43 times more likely under this model than the next best model (sequence + change magnitude; BF_{10} = 3.22e+35), and 75.30 times more likely than the third best model (sequence + change magnitude + interaction; BF_{10} = 1.89e+34). The model containing sequence only showed evidence in support of the null, thus no comparisons were made. To determine where the effect of change magnitude lay, three Bayesian paired-samples t-tests were conducted. The first compared performance between change magnitudes of 5° ($M = 0.25$, $SD = 0.07$) and 15° ($M = 0.30$, $SD = 0.07$) revealing a BF_{10} = 242.02 indicating extreme evidence in support of a difference in performance. The second compared performance between change magnitudes of 15° and 30° ($M = 0.39$, $SD = 0.11$) revealing a BF_{10} = 19,722,904.14 indicating extreme evidence in support of a difference

Table 3

Table displaying Bayes Factors (BF_{10}) for each model across accuracy and response time analyses in Experiment 1a. The BF_{10} for each model is in comparison to the null model.

Measure	Model	BF	Error
Accuracy	Sequence	0.171	0.038
	Change magnitude	1.42e+36	0.009
	Sequence + Change magnitude	3.22e+35	0.012
	Sequence + Change magnitude + Interaction	1.89e+34	0.015
Response time	Sequence	0.2	0.013
	Change magnitude	0.044	0.005
	Sequence + Change magnitude	0.009	0.03
	Sequence + Change magnitude + Interaction	0.001	0.017

in performance. The final t-test compared performance between change magnitudes of 30° and 45° (M = 0.46, SD = 0.14) revealing a BF_{10} = 1,006.91 indicating extreme evidence in support of a difference in performance. A further Bayesian ANOVA examining the effects of sequencing condition and change magnitude on RT data revealed that all models showed evidence in support of the null.

Formal model comparison. Log-likelihoods and AICs for each of the models in Experiment 1a can be found in Table 4. AIC was calculated for each of the six models, revealing the fixed model as the best (AIC = 19764), with this model being the best for 77.50% of participants. The second best model was the free tau model (AIC = 19841; 10% of participants), followed by the free j-bar model (AIC = 19845; 12.50% of participants), the free p model (AIC = 19882), the full model (AIC = 20028), and finally the fully free model (AIC = 20184). Despite the fixed model being the preferred model on the basis of the model comparison, analysis on parameter values for each model was still undertaken and will be reported in turn, preceded by results of the behavioural analysis.

Table 4

Table displaying variability of parameters across sequencing condition and feature dimension, log-likelihood, and AIC for each model in Experiment 1a. A ⊕ symbol indicates that the parameter was free to vary across sequencing condition, while a ⊖ symbol indicates that the parameter was not free to vary across feature dimension.

Model	$\bar{J}$	τ	p	*LL*	*AIC*
Fully free	⊕ ⊕	⊕ ⊕	⊕ ⊕	-9,611.86	20,183.72
Full	⊕ ⊕	⊕ ⊕	⊕ ⊖	-9,613.84	20,027.68
Free j-bar	⊕ ⊕	⊖ ⊕	⊖ ⊖	-9,642.72	19,845.44
Free tau	⊖ ⊕	⊕ ⊕	⊖ ⊖	-9,640.45	19,840.90
Free p	⊖ ⊕	⊖ ⊕	⊕ ⊖	-9,661.00	19,882.00
Fixed	⊖ ⊕	⊖ ⊕	⊖ ⊖	-9,681.77	19,763.54

Note. Here, the resources dedicated to colour and orientation are held to vary, hence why $\bar{J}$ and τ are free to differ across feature dimension in all models.

Full model. Within the full model, the $\bar{J}$ and τ parameters could vary across both feature dimension and sequencing condition, while the p parameter could only vary across sequencing condition. A Bayesian ANOVA conducted on $\bar{J}$ parameter values revealed evidence in support of the null for all models. Similarly, a Bayesian ANOVA conducted on τ parameter values revealed evidence in support of the null for all models. As the p parameter could only vary across sequencing condition, comparison of p values for repetition ($M = 0.64$, $SD = 0.34$) and switch ($M = 0.64$, $SD = 0.36$) revealed a $\mathrm{BF}_{10} = 0.17$ indicating moderate evidence in support of the absence of a switch cost for the p parameter.

Fixed model. The $\bar{J}$ and τ parameters could only vary across feature dimension in the fixed model, while he p parameter was not allowed to vary across either feature dimension or sequencing condition. Comparison of $\bar{J}$ parameter values between colour ($M = 5.65$, $SD =$ 4.26) and orientation ($M = 5.93$, $SD = 5.32$) revealed a $\mathrm{BF}_{10} = 0.19$ indicating moderate evidence in support of no difference between $\bar{J}$ parameter values between feature dimension. Comparison of τ parameter values between colour (M = 22.78, SD = 15.18) and orientation (M = 24.21, SD = 14.91) revealed a $\mathrm{BF}_{10} = 0.20$ indicating moderate evidence in support of no difference in τ parameter values between feature dimension.

Fully-free model. Within the fully-free model, all parameters could vary across feature dimension and sequencing condition. A Bayesian ANOVA conducted on $\bar{J}$ parameter values revealed evidence in support of the null for all models. A further Bayesian ANOVA conducted on τ parameter values also revealed evidence in support of the null for all models. Finally, a Bayesian ANOVA conducted on p parameter values revealed the model containing dimension only as the best model ($\mathrm{BF}_{10} = 9.62$), being 5.18 times more likely than the second best model (sequence + dimension; $\mathrm{BF}_{10} = 1.86$). All other models showed evidence in support of the null.

Free j-bar model. In the free j-bar model, the $\bar{J}$ parameter could vary across feature dimension and sequencing condition, while the τ parameter could only vary across feature dimension. The p parameter could not vary across either feature dimension or sequencing condition. A Bayesian ANOVA conducted on $\bar{J}$ parameter values revealed evidence in support of the null for all models. Comparison of τ parameter values between colour (M = 27.43, SD = 10.48) and orientation (M = 22.99, SD = 14.04) revealed a $\mathrm{BF}_{10} = 0.83$ indicating anecdotal evidence in support of no difference between feature dimensions.

Free tau model. In the free tau model, the $\bar{J}$ parameter could vary across feature dimension only, while the τ parameter could vary across both feature dimension and sequencing condition. The p parameter could not vary across either feature dimension or sequencing condition. Comparison of $\bar{J}$ parameter values between colour (M = 4.66, SD = 3.51) and orientation (M = 5.10, SD = 4.23) revealed a $\text{BF}_{10} = 0.23$ indicating moderate evidence in support of no difference between feature dimensions. A Bayesian ANOVA conducted on τ parameter values revealed the model containing dimension only as the best model ($\text{BF}_{10} = 2.50$); all other models revealed evidence in support of the null hence no comparisons were made.

Free p model. In the free p model, both the $\bar{J}$ and τ parameters could vary across feature dimension only, while the p parameter could vary across sequencing condition only. Comparison of $\bar{J}$ parameter values between colour (M = 5.67, SD = 4.12) and orientation (M = 5.03, SD = 4.67) revealed a $\text{BF}_{10} = 0.29$ indicating moderate evidence in support of no difference between feature dimensions. Comparison of τ parameter values between colour (M = 22.88, SD = 14.05) and orientation (M = 18.41, SD = 13.52) revealed a $\text{BF}_{10} = 0.73$ indicating anecdotal evidence in support of no difference between feature dimensions. Finally, comparison of p parameter values between repetition (M = 0.49, SD = 0.38) and switch (M = 0.68, SD = 0.33) revealed a $\text{BF}_{10} = 2.00$ indicating anecdotal evidence in support of a switch cost.

5.3.2 Discussion

While results from the behavioural analysis in Experiment 1a revealed a substantial effect of change magnitude on accuracy, there was a distinct lack of evidence to suggest an effect of sequencing condition or an interaction between sequencing condition and feature dimension. Furthermore, results from comparisons between model parameters revealed that values did not differ to any great extent between sequencing condition in models wherein parameters were free to vary. Support for these findings is also given through formal model comparison, which found the fixed model to be the best. On this model, the $\bar{J}$ and τ parameters could vary only across feature dimension, while the p parameter was fixed across all conditions. This suggests that the data obtained here are best explained by a model where switching does not influence the precision of representations in memory. However, it should be noted that performance on this task was somewhat poor; this may have been due to the relatively small change magnitudes employed in this experiment (5, 15, 30, and 45). As such, Experiment 1b employed the use of more substantial change magnitudes.

5.4 Experiment 1b

As stated, Experiment 1b used larger change magnitudes (5, 30, 60, 90) than those used in Experiment 1a. This was based on the possibility that performance on Experiment 1a suffered due to the use of smaller change magnitudes, which may have impacted results and model fits.

5.4.1 Results

Behavioural analysis. See Table 5 for BFs for all models in the accuracy and response time analysis in Experiment 1b. A Bayesian ANOVA examining the effect of sequencing condition (repetition, switch) and change magnitude (5, 30, 60, 90) on accuracy data revealed the model containing the main effect of change magnitude only as the best model (BF_{10} = 3.6e+78) with the data being 6.76 times more likely under this model than the next best model (sequence + change magnitude; BF_{10} = 5.33e+77), and 199 times more likely than the third best model (sequence + change magnitude + interaction; BF_{10} = 1.81e+76).

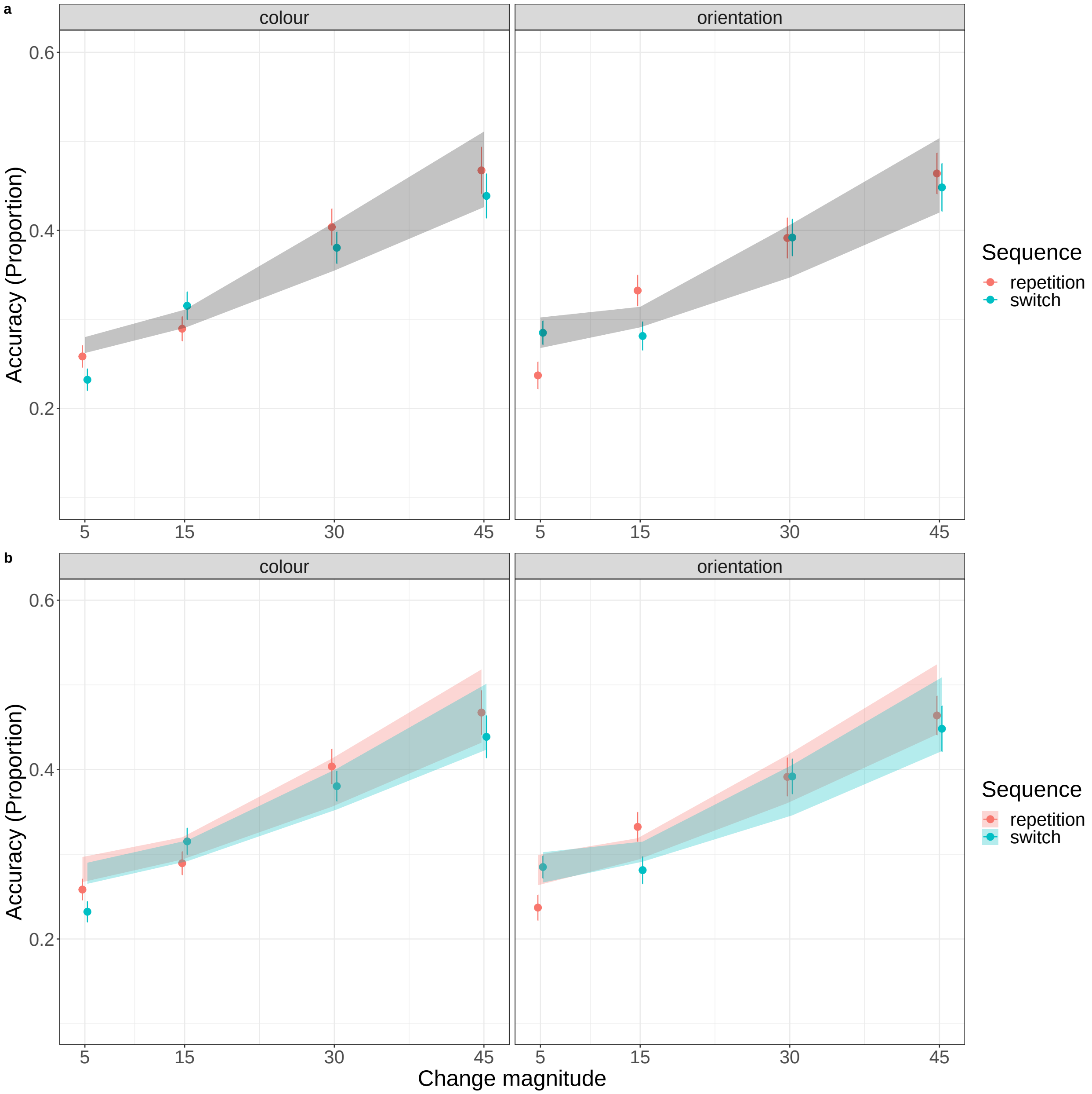

Figure 16

Plots displaying a) fixed model fit and b) free tau model fit in Experiment 1a. Note that the fixed model provided the best fit to the data, followed by the free tau model.

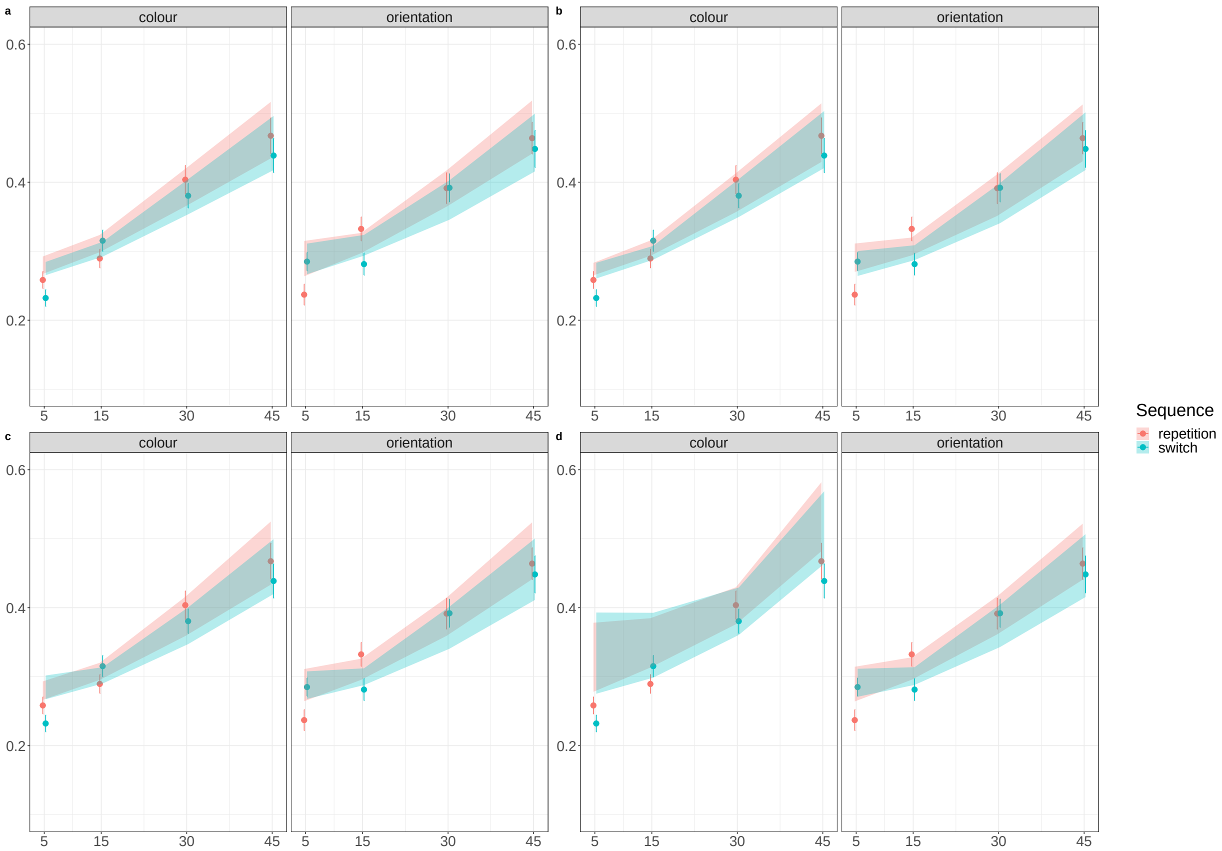

Figure 17

Plots displaying the remaining model fits in Experiment 1a; a) free j-bar model, b) free p model, c) full model, d) fully-free model.

The model containing the main effect of sequence only showed evidence in support of the null, hence no comparisons were made. To determine where the effect of change magnitude lay, three Bayesian paired-samples t-tests were conducted. The first compared performance between change magnitudes of 5° ($M = 0.24$, $SD = 0.05$) and 30° ($M = 0.41$, $SD = 0.09$) revealing a $\mathrm{BF}_{10} = 418198763271607\text{28}$ indicating extreme evidence in support of a difference in performance. The second compared performance between change magnitudes of 30° and 60° ($M = 0.56$, $SD = 0.13$) revealing a $\mathrm{BF}_{10} = 3{,}159{,}000{,}436{,}247{,}178.00$ indicating extreme evidence in support of a difference in performance. The final comparison was between change magnitudes of 60° and 90° ($M = 0.60$, $SD = 0.13$) revealing a $\mathrm{BF}_{10} = 32.80$

Table 5

Table displaying Bayes Factors (BF_{10}) for each model across accuracy and response time analyses in Experiment 1b. The BF_{10} for each model is in comparison to the null model.

Measure	Model	BF	Error
Accuracy	Sequence	0.13	0.039
	Change magnitude	3.6e+78	0.009
	Sequence + Change magnitude	5.33e+77	0.012
	Sequence + Change magntidue + Interaction	1.81e+76	0.015
Response time	Sequence	0.12	0.014
	Change magnitude	6617215980.609	0.004
	Sequence + Change magnitude	846210059.834	0.031
	Sequence + Change magnitude + Interaction	60438529.44	0.017

indicating strong evidence in support of a difference in performance.

A further Bayesian ANOVA examining the effects of sequencing condition and change magnitude on RT data also revealed the model containing the main effect of change magnitude only as the best model (BF_{10} = 6,617,215,980.61), with the data being 7.82 times more likely under this model than the next best model (sequence + change magnitude; BF_{10} = 846,210,059.83), and 109.49 times more likely compared to the third best model (sequence + change magnitude + interaction; BF_{10} = 60,438,529.44). The model containing the main effect of sequence only showed evidence in support of the null, hence no comparisons were made. To determine where the effect of change magnitude lay, three Bayesian paired-samples t-tests were conducted. The first examined performance between change magnitudes of 5° (M = 847.09, SD = 248.00) and 30° (M = 817.26, SD = 264.07) revealing a BF_{10} = 0.38 indicating anecdotal evidence in support of the absence of a difference in performance. The second compared performance between change magnitudes of 30° and 60° (M = 739.25, SD = 186.78) revealing a BF_{10} = 20.00 indicating strong evidence in support of a difference in

Table 6

Table displaying variability of parameters across sequencing condition and feature dimension, log-likelihood, and AIC for each model in Experiment 1b. A ⊕ symbol indicates that the parameter was free to vary across sequencing condition, while a ⊖ symbol indicates that the parameter was not free to vary across feature dimension.

Model	$\bar{J}$	τ	p	*LL*	*AIC*
Fully free	⊕ ⊕	⊕ ⊕	⊕ ⊕	-9,724.06	20,408.12
Full	⊕ ⊕	⊕ ⊕	⊕ ⊖	-9,723.93	20,247.86
Free j-bar	⊕ ⊕	⊖ ⊕	⊖ ⊖	-9,773.38	20,106.76
Free tau	⊖ ⊕	⊕ ⊕	⊖ ⊖	-9,758.05	20,076.10
Free p	⊖ ⊕	⊖ ⊕	⊕ ⊖	-9,777.20	20,114.40
Fixed	⊖ ⊕	⊖ ⊕	⊖ ⊖	-9,791.50	19,983.00

Note. The models used in Experiment 1b were identical to those used in Experiment 1a.

performance. The final comparison was between change magnitudes of 60° and 90° (M = 705.10, SD = 171.88) indicating a $\text{BF}_{10} = 1.24$ indicating anecdotal evidence in support of a difference in performance.

Formal model comparison. Log-likelihoods and AICs for each of the models in Experiment 1b can be found in Table 6. AIC was calculated for each of the six models, revealing the fixed model as the best (AIC = 19983), with this model being the best for 87.50% of participants. The second best model was the free tau model (AIC = 20076; 7.50% of participants), followed by the free j-bar model (AIC = 20107; 5% of participants), the free p model (20114), the full model (AIC = 20248), and finally the fully free model (AIC = 20408).

Full model. Within the full model, the $\bar{J}$ and τ parameters could vary across both feature dimension and sequencing condition, while the p parameter could only vary across sequencing condition. A Bayesian ANOVA conducted on $\bar{J}$ parameter values revealed that all models showed evidence in support of the null. A further Bayesian ANOVA conducted on τ parameter values revealed the model containing feature only as the best model (BF_{10} = 6.74); the data was 5.72 times more likely under this model than the next best model (sequence + dimension; BF_{10} = 1.18). All other models showed evidence in support of the null. A Bayesian paired-samples t-test between repetition (M = 0.65, SD = 0.34) and switch (M = 0.69, SD = 0.30) values for the p parameter revealed a BF_{10} = 0.20 indicating moderate evidence in support of the absence of a switch cost.

Fixed model. The $\bar{J}$ and τ parameters could only vary across feature dimension in the fixed model, while he p parameter was not allowed to vary across either feature dimension or sequencing condition. Comparison of $\bar{J}$ parameter values between colour (M = 5.42, SD = 4.05) and orientation (M = 6.77, SD = 4.33) revealed a BF_{10} = 0.57 indicating anecdotal evidence in support of no difference between values across feature dimension. Comparison of τ parameter values between colour (M = 15.63, SD = 14.27) and orientation (M = 24.92, SD = 12.14) revealed a BF_{10} = 14.11 indicating strong evidence for a difference between values across feature dimension.

Fully-free model. Within the fully-free model, all parameters could vary across feature dimension and sequencing condition. A Bayesian ANOVA conducted on $\bar{J}$ parameter values revealed that all models showed evidence in support of the null. Analysis of τ parameter values revealed the model containing feature only as the best model (BF_{10} = 2.80); all other models showed evidence in support of the null, hence no comparisons were made. Finally, analysis of p parameter values also revealed the model containing feature only as the best model ($\mathrm{BF}_{10} = 4.38$); again, all other models showed evidence in support of the null.

Free j-bar model. In the free j-bar model, the $\bar{J}$ parameter could vary across feature dimension and sequencing condition, while the τ parameter could only vary across feature dimension. The p parameter could not vary across either feature dimension or sequencing condition. A Bayesian ANOVA conducted on $\bar{J}$ parameter values revealed that all models showed evidence in support of the null. Comparison of τ parameter values between colour ($M = 24.80$, $SD = 12.39$) and orientation ($M = 26.86$, $SD = 10.90$) revealed a $\mathrm{BF}_{10} = 0.23$ indicating moderate evidence in support of no difference between values across feature dimension.

Free tau model. In the free tau model, the $\bar{J}$ parameter could vary across feature dimension only, while the τ parameter could vary across both feature dimension and sequencing condition. The p parameter could not vary across either feature dimension or sequencing condition. Comparison of $\bar{J}$ parameter values between colour ($M = 5.63$, $SD = 3.93$) and orientation ($M = 6.31$, $SD = 3.75$) revealed a $\mathrm{BF}_{10} = 0.28$ indicating moderate evidence in support of no difference between values across feature dimension. A Bayesian ANOVA conducted on τ parameter values revealed the model containing feature only as the best model ($\mathrm{BF}_{10} = 2{,}488.40$); the data is 5.16 times more likely under this model than the next best model (feature + sequence; $\mathrm{BF}_{10} = 481.81$) and 20.49 times more likely under this model than the third best model (sequence + change magnitude + interaction; $\mathrm{BF}_{10} = 121.47$). The model containing sequence only showed evidence in support of the null.

Free p model. In the free p model, both the $\bar{J}$ and τ parameters could vary across feature dimension only, while the p parameter could vary across sequencing condition only. Comparison of $\bar{J}$ parameter values between colour ($M = 5.78$, $SD = 4.35$) and orientation ($M = 6.69$, $SD = 4.36$) revealed a $\mathrm{BF}_{10} = 0.28$ indicating moderate evidence in support of no difference between values across feature dimension. Comparison of τ parameter values between colour ($M = 17.47$, $SD = 15.15$) and orientation ($M = 25.66$, $SD = 12.88$) revealed a $\mathrm{BF}_{10} = 4.12$ indicating moderate evidence in support of a difference between values across feature dimension. Finally, comparison of p parameter values between repetition ($M = 0.61$, $SD = 0.31$) and switch ($M = 0.59$, $SD = 0.34$) revealed a $\mathrm{BF}_{10} = 0.18$ indicating moderate evidence in support of the absence of a switch cost.

5.4.2 Discussion

Overall, the results of Experiment 1b paint a similar picture to those from Experiment 1a. Behavioural analysis revealed a substantial impact of change magnitude on performance however, a lack of evidence for an effect of sequencing condition. Perhaps the most notable deviation in results relative to Experiment 1a was found in comparison of parameters within the free tau model; while not being of critical importance to the present study, a substantial effect of feature dimension was observed for the τ parameter. Interestingly, the model containing both main effects of sequencing condition and feature dimension, as well as the model containing the interaction, also showed substantial evidence for the presence of an effect, yet the model containing only the main effect of sequencing condition showed evidence in favour of the null. Comparison of parameter values within the remaining models revealed a distinct lack of evidence for any differences across both sequencing condition and feature dimension. Formal model comparison again revealed the fixed model as being the best model, with the free tau model came in at second-best, as it did in Experiment 1a. Taken together, the results of Experiments 1a and 1b suggest that—at least in terms of the paradigm used in this study—dimension switching has little impact on memory encoding and resource allocation.

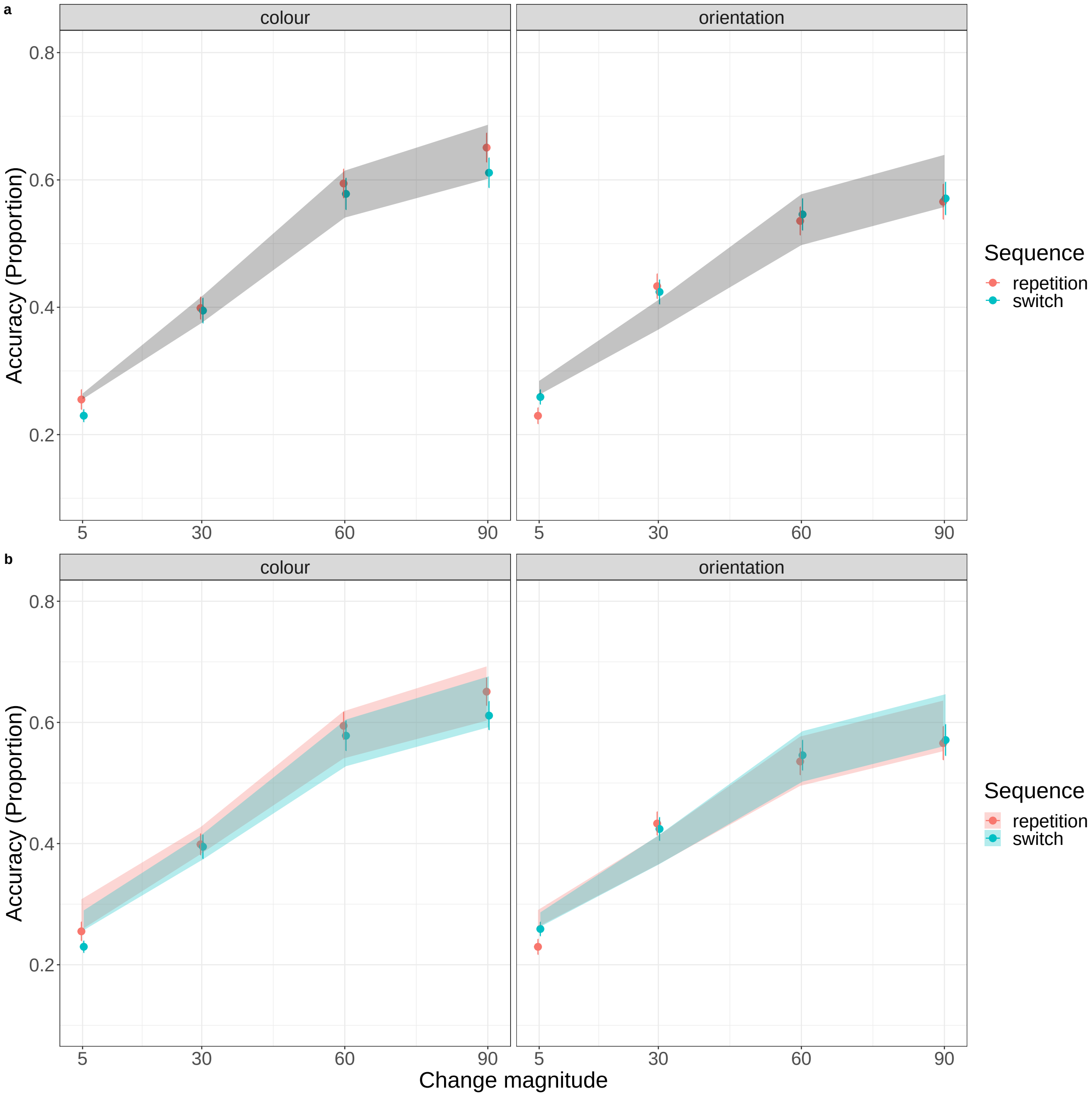

Figure 18

Plots displaying a) fixed model fit and b) free-j model fit in Experiment 1b. Note that the fixed model provided the best fit to the data, followed by the free-j model.

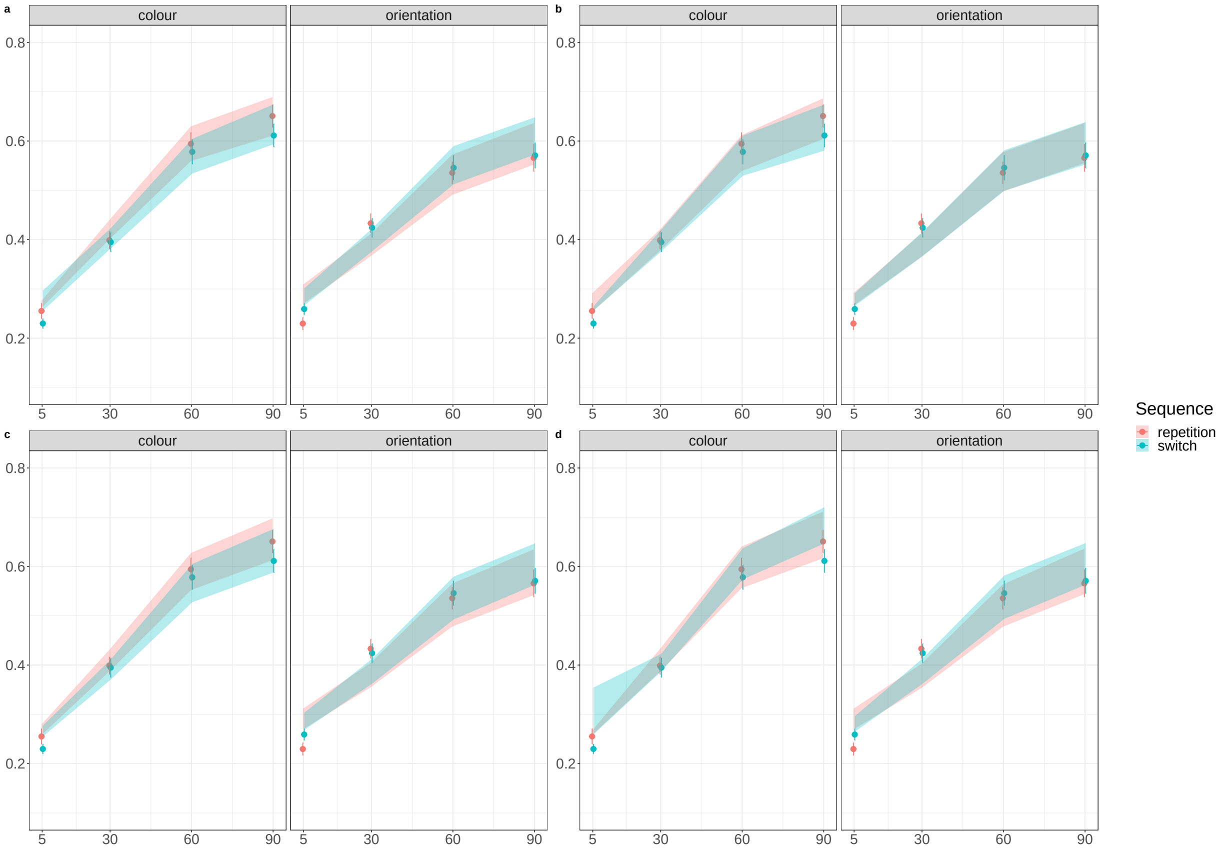

Figure 19

Plots displaying the remaining model fits in Experiment 1b; a) free-p model, b) free-tau model, c) full model, d) fully-free model.

5.5 General discussion

5.5.1 Summary of results

The aim of the present study was to investigate the impact of dimension switching on the processes of resource allocation and memory encoding by adapting methods and models from Shin and Ma (2017). This involved two change localisation experiments wherein changes varied in magnitude and the goal was to determine where the change occurred. Six models were then fit to the data obtained from these experiments, with all models containing the parameters $\bar{J}$, τ, and p. The $\bar{J}$ and τ parameters determine the precision at which a feature is encoded into memory, with p reflecting the amount of resource allocated to the irrelevant feature dimension. Thus, variations in these parameter values across sequencing condition (repetition, switch) would indicate issues within the processes these parameters reflect. Whether these parameters were free to vary across sequencing condition and feature dimension was variable across models. Analysis of behavioural data across Experiments 1a and 1b revealed no evidence to suggest any main effect of dimension switching, with evidence pointing to a substantial impact of change magnitude. Formal model comparison across Experiments 1a and 1b also found that a model wherein parameters could not vary across sequencing condition was also favoured over all other models. Taken together, these results suggest that dimension switching has little influence over the processes of resource allocation and memory encoding, contrasting results from initial change detection experiments (Chapters 2 and 3), but supporting those found in the continuous report experiment (Chapter 4).

5.5.2 On the lack of evidence for a dimension switch cost

While some theoretical speculation was offered based on similar results in Chapter 4, it is more difficult to offer the same kind of speculation here given the overwhelming lack of evidence in support of any impact of dimension switching. As such, the majority of this discussion will be dedicated to potential reasons why a lack of evidence may have been observed in the present study, highlighting directions for future versions of this task. The results of the present study seemingly indicate that dimension switching does not have any impact on resource allocation and memory encoding, as well as having little influence on performance in the change localisation task, contrasting findings from previous chapters (Chapters 2 and 3). Perhaps most crucially, the results obtained in the present study provide evidence against the notion that dimension switching causes a misallocation of resources to the irrelevant feature dimension. The p parameter in all of the model fit routines represents the amount of resource allocated to the irrelevant feature dimension. Results from the formal model comparison revealed that the fixed model was favoured over all other models, indicating that the amount of resource allocated to the irrelevant feature dimension was sufficiently captured by a single value regardless of sequencing condition (or feature dimension). Had dimension switching caused a misallocation of resources to the irrelevant feature dimension, this would have been reflected by poor fits for models wherein p could not vary across sequencing condition and differences between p values across sequencing condition in models wherein p was free to vary.

The present results also contradict speculation from Chapter 4 concerning misbinding. While weak evidence was observed for an increased rate of non-target responses on dimension switch trials relative to dimension repetition trials, a theoretical account was given based on this being a true effect. Previous research has stated that attention is required for the binding of relevant visual information (Hyun et al., 2009; Rensink, 2000; Treisman, 1988; Treisman & Gelade, 1980; Treisman & Schmidt, 1982; Wan et al., 2020; Wheeler & Treisman, 2002; Zokaei et al., 2014), with attention also required for switching between feature dimensions. Thus, if attention is considered as a finite resource responsible for controlled action (see e.g., Oberauer, 2019), it was suggested that the process of configuring the relevant attentional set either depletes this attentional resource to such an extent that little or no resource is available for feature binding or that a failure to sufficiently configure attentional set prevents binding from occurring, as attention can only be directed to a single process at any given time.

The absence of dimension switch costs within the behavioural analysis in the present study would suggest that this notion of misbinding due to depletion of attention is incorrect. If such depletion occurred and was responsible for misbinding leading to poorer performance on switch relative to repetition trials, this would be reflected by a switch cost within the behavioural measures, primarily accuracy. In terms of the model fitting, while it is difficult to state with any certainty which parameters would be affected by misbinding, it is likely that there would have been some difference in the $\bar{J}$ and/or τ parameters across sequencing condition had misbinding occurred. As shown by previous research (see e.g., Bays et al., 2009), misbinding leads to an increase in errors; thus, had misbinding been observed in the present study, it could be suggested that this may have been reflected within the model fits in terms of less precise memory encoding. However, results clearly indicate that the data is best explained in terms of a model wherein switching has no influence on the precision of representations in memory (i.e., the fixed model). Within the fixed model the $\bar{J}$ and τ parameters were free to vary across feature dimension, but no parameters (including p) were able to vary across sequencing condition. This suggests that the process of memory encoding across sequencing condition was sufficiently captured by a single value, without the need to invoke separate values for each sequencing condition.

It is also worth noting that less detailed speculation within the discussion in Chapter 4 suggested that the effect of dimension switching may have been short-lived. This was based on the differences in the change detection tasks wherein dimension switch costs were observed and the continuous report task wherein little evidence for such a cost was found. In the change detection tasks, responses were binary (change or no change), with response selection causing the task to progress. In the continuous report task, participants had 360° different response options to choose from and could select multiple responses using a mouse-driven pointer before confirming a response by pressing spacebar. As such, the number of responses and method of response is vastly different between the tasks, with overall RTs substantially larger on continuous report tasks relative to change detection tasks. This led to the suggestion that the absence of any major evidence in support of dimension switch costs in Chapter 4 may have been due to the effect only lasting for a brief period after stimulus presentation. However, based on results from the present set of experiments, this notion appears to be incorrect. While change detection and change localisation tasks are somewhat different, with change localisation having an increased number of possible responses and variation in change magnitudes, responses should typically be provided at faster rates than those observed in continuous report. Indeed, RTs from the present study were found to be no greater than 1,110ms (for a given change magnitude) relative to RTs from the continuous report task in Chapter 4 which exceeded 2,000ms. Based on this and the results from the present study, it therefore appears that the notion of the effects of dimension switching being short-lived can be ruled out.

However, it is important to highlight what may be a crucial methodological difference between previous change detection experiments which found the presence of a dimension switch cost (i.e., most experiments within Chapters 2 and 3) and the change localisation experiments within the present study. Throughout the present thesis, robust evidence for dimension switch costs were only observed in experiments which employed an additional, constant change in the irrelevant feature dimension. In Experiments 1a and 2a in Chapter 2 however, a change was only implemented in the relevant feature dimension, with results showing little impact of dimension switching on performance. It was suggested that this was due to the cue being non-essential for correct performance of the task; as a change only occurred in the relevant dimension, it was possible for participants to simply monitor the entire display and report any change, eliminating the need to switch between dimensions and negating any potential switch costs that may have presented. As such, it is possible that a similar strategy was adopted by participants in the present study; as a change only occurred within the relevant feature dimension, it could have been the case that participants simply ignored the cue and monitored the entire stimulus display for any change, again negating any potential switch costs that may have presented. While initial evidence for such a possibility was only offered by behavioural data from the change detection tasks, the finding that the fixed model provided the best fit to the data for both experiments within the present study may provide further support for the notion that participants did not switch between feature dimensions, but simply attended to the entire stimulus display.

While an additional change in the irrelevant feature dimension could be implemented in this change localisation task, the difficulty of the task led to the decision not to include such a change. In contrast to change detection, there are four possible responses in a change localisation task, increasing the number of decisions that need to be made which could impact performance and cloud the dimension switch cost. In addition, the use of smaller change magnitudes means that on some trials, the change is barely noticeable, making localisation of the change much more difficult. Critically, these factors may not only have influenced behavioural performance on the task, but may also have confounded the results from model fitting and formal model comparison and as such, it was deemed appropriate to not include a constant irrelevant change. Future research implementing similar methods may benefit from including an additional change in the irrelevant feature dimension to ensure that the cue is essential for correct performance of the task, either as a irrelevant change with a constant change magnitude or an irrelevant change with a change magnitude matching that of the relevant change.

In relation to methodological limitations, there are a number of further issues which may have contributed to the results obtained in the present study. Perhaps most importantly was the trial number used; while Shin and Ma (2017) used between 2,400 and 3,600 trials, only 400 trials were used per participant in each of experiments detailed previously. As such, it would be pertinent to perform a conceptual replication of this study using a higher number of trials which may also allow for sample size to be reduced. Furthermore, while the online experimentation used in the present study did provide decent data, it would be more appropriate to conduct this experiment in a laboratory setting. This allows for greater control over experimental parameters which may have impacted performance in the current study. For instance, differences in screen sizes used by participants when performing the task may have had some impact; it may, for example, have been more difficult to notice changes of smaller magnitudes with a smaller screen size or different brightness settings.

5.5.3 Conclusion

The present study aimed to determine the influence of dimension switching on resource allocation and memory encoding processes through the application of explanatory models to behavioural data obtained from change localisation tasks. Findings show an overwhelming absence of evidence in support of any influence of dimension switching on these processes, both in terms of analysis of behavioural data and formal model comparison. While these results may suggest that dimension switching has little impact on resource allocation, methodological issues may be the cause for the lack of evidence for such an impact. Perhaps the most crucial issue is that pertaining to the use of a constant change in the irrelevant feature dimension which has been suggested to ensure the cue is used for correct task performance. Thus, it would be appropriate for future research to investigate the influence of an additional irrelevant change on this task however, there are obvious concerns regarding the influence such a change may have on model fitting.

6 Chapter 6: General discussion

This chapter provides an overview of the results obtained in Chapters 2-5 in relation to the overarching research question concering the role of attentional control in the allocation of visual short-term memory resources. In addition to the theoretical implications of the findings from previous chapters, some theoretical speculation will also be discussed.

6.1 The research question

Stimuli within our everyday environment are often multivalent, containing any number of feature dimensions (e.g., colour, shape) or indeed, multiple values within the same dimension. Therefore, given the limited nature of visual short-term memory (vSTM), the cognitive system must ensure that the limited resources available are allocated to the most relevant information. As such, the aim of the research within this thesis was to investigate the role of attentional control in the allocation of vSTM resources to relevant information. This involved embedding a range of tasks which probe vSTM (e.g., change detection) into task switching paradigms. Stimulus displays consisted of either bivalent, coloured and oriented shapes or univalent shapes displaying either colour or orientation, but critically, each stimulus display always contained both colour and orientation. While in traditional task switching participants either repeat or switch between two (or more) tasks, here participants performed the same task throughout with the critical switching element relating to which feature dimension was relevant on a given trial. Therefore, on some trials the relevant feature dimension could repeat (e.g., colour–colour) or switch (e.g., colour–orientation). Prior to discussion of the theoretical contributions and limitations of this research, an overview of the results obtained within Chapters 2–5 is given.

6.2 Overview of results

6.2.1 Chapter 2

Chapter 2 consisted of a total of three change detection experiments wherein participants were presented with either bivalent stimuli containing colour and orientation or univalent stimuli containing either colour or orientation. On each trial, the relevant feature dimension (colour or orientation) could either repeat (e.g., colour–colour) or switch (e.g., colour–orientation); the task always remained the same throughout (determine if a change occurred in the relevant dimension). This introduced an element of attentional control to a task which otherwise did not require such control. Experiments 1 and 2 aimed to provide initial insight into the role of attentional control in the allocation of vSTM resources using bivalent and univalent stimuli respectively. Results from Experiment 1a revealed that dimension switching appeared to have no impact on change detection performance across a range of measures. While this may have suggested that dimension switching did not influence vSTM performance, a possible methodological oversight may have led to these results; given that a change only occurred in the relevant feature dimension, participants would be able to ignore the cue and monitor the entire display for any change. As such, this would remove the requirement to switch between dimensions and negate any cost to performance that may be observed. Therefore, in Experiment 1b, an additional, constant change in the irrelevant feature dimension was introduced, making the cue essential for correct performance of the task. Results from Experiment 1b revealed a substantial impact of dimension switching on performance across accuracy, RT, and sensitivity, suggesting that dimension switching did impact change detection performance, but only when the cue was essential for correct task performance. Results from Experiment 2a were much in line with those from Experiment 1a, with the finding that dimension switching did not impact change detection performance to any great extent. However, results from Experiment 2b revealed a somewhat differing picture from those in Experiment 1b; the impact to performance appeared to be most prevalent between pure (wherein the feature dimension always repeated) and mixed blocks of trials. The general lack of evidence in support of a dimension switch cost here could therefore be attributed to the fact that stimuli were univalent, meaning only two stimuli were relevant on a given trial.

Experiment 3 aimed to both confirm findings from Experiments 1 and 2 and to extend these findings. Experiment 3a and 3b were conceptual replications of Experiments 1b and 2b, with Experiment 3a finding substantial evidence for the presence of a switch cost for accuracy, RT, and sensitivity, again suggesting that dimension switching impacted performance on the change detection task. While also providing support for the presence of a dimension switch cost, the extent of evidence in Experiment 3b was much more modest, again highlighting the possibility that stimulus valency moderates the extent of the dimension switch cost. In Experiment 3c, univalent stimuli were again used however, set size was doubled from four to eight. This meant that while a greater number of stimuli were presented, the total number of features presented was equivalent to that of Experiment 3a, as were the total number of relevant stimuli on a given trial. Results from Experiment 3c revealed that dimension switching impacted performance, with the highest evidence found for RT, sensitivity, and response bias measures. Finally, Experiment 3d used spatially overlapping stimuli (coloured circles overlaid onto white isosceles triangles). Interestingly, results from Experiment 3d were notably similar to those from Experiment 3c, with both accuracy and response bias for sequencing condition being almost identical. Taken together, the results from Chapter 2 provided support for the notion that dimension switching impacts change detection performance, with the suggestion that such disruption was due to a misallocation of resources to the irrelevant feature dimension on switch trials.

6.2.2 Chapter 3

Chapter 3 consisted of two experiments which sought to investigate the impact of trial timing manipulations and cue switching on the extent of the dimension switch cost in the same change detection paradigm used in Chapter 2. Experiment 1a manipulated trial timings by having either a long or short cue-stimulus interval (CSI), while Experiment 1b had either a long or short response cue-interval (RCI). The CSI is the duration between presentation of the cue and stimulus on trial *N* and is associated with the process of advance reconfiguration (see e.g., Rogers & Monsell, 1995), whereas the RCI is the duration between response on trial *N-1* and cue presentation on trial *N* and is associated with the effect of proactive interference (see e.g., Allport et al., 1994). Crucially, extending the duration of either of these intervals had previously been shown to attenuate the task switch cost (for reviews on task switching, see e.g., Grange & Houghton, 2014; Kiesel et al., 2010; Vandierendonck et al., 2010), with Experiment 1 aiming to investigate whether a similar effect was found for dimension switch costs. Results from both Experiments 1a and 1b revealed that manipulation of the CSI or RCI respectively did not have any impact on the extent of the dimension switch cost, suggesting that the origin of such costs lies with processes not associated with those concerning interference from previous task sets or configuration of new task sets.

Experiment 2 employed a two cues per task (2:1 mapping) paradigm wherein participants performed the same change detection switching task outlined in Experiment 1b; in this experiment however, each feature dimension could be probed by one of two cues, giving rise to three different sequencing conditions that those seen previously. On some trials, both the cue and relevant feature dimension could repeat, referred to as a *cue repetition*; on some trials the cue could switch while the relevant feature dimension repeated, referred to as a *cue switch*; and finally, on some trials, both the cue and relevant feature dimension could switch, referred to as a *dimension switch*. Differences in performance between cue repetition and cue switch trials reflects the impact of a cue switch independent of a dimension switch (Logan & Bundesen, 2004; for the impact of a cue switch on measures of the task switch cost, see e.g., Logan & Bundesen, 2003; Mayr & Kliegl, 2003), while differences in performance between a cue switch and dimension switch reflects the true impact of a dimension switch cost. In Experiment 2a, all three sequencing conditions were used with results showing that cue switching did not contribute to the overall dimension switch cost. Results from Experiment 2b—which removed cue repetition trials and increased cue switch and dimension switch trial numbers—also found that cue switching did not contribute to the overall dimension switch cost.

6.2.3 Chapter 4

Chapter 4 aimed to provide a more nuanced insight into the cause of dimension switch costs by implementing the same dimension switching design within a continuous report task. Here, participants were presented with four bivalent coloured, oriented stimuli and then probed with the location of one of the stimuli; if colour was relevant, participants used a

360° colour wheel to provide the colour of the stimulus in the probed location, whereas if orientation was relevant, participants used a 360° black wheel to provide the orientation of the stimulus in the probed location. The three-component mixture model (see Bays et al., 2009) was then fit to the data obtained from this experiment; this model returns parameters estimating the precision at which the probed items were held in memory (κ), the probability of making a target (p_t) and non-target (p_{nt}) response, and the probability of guessing (p_u). Results from the model fitting revealed support for no dimension switch cost in most parameters, with only weak evidence for a switch cost in the p_{nt} parameter, suggesting that dimension switching does not impact resource allocation on the continuous report task. However, if a true effect, the weak evidence in support of the presence of a switch cost in the p_{nt} parameter may reflect misbinding of relevant information. As dimension switching and binding are attentionally demanding processes, it is possible that on switch trials attention was depleted to such an extent that improper binding of relevant information upon stimulus presentation occurred, leading to a higher rate of non-target responses.

6.2.4 Chapter 5

Chapter 5 attempted to provide deeper insight into the cause of dimension switch costs by adapting methods and models used by Shin and Ma (2017). Two experiments involving change localisation tasks were used. On each trial, a change occurred in either colour or orientation, meaning the relevant feature could either repeat (e.g., colour–colour) or switch (e.g., colour–orientation). The magnitude of the change on each trial varied, with participants required to report the location of where the change occurred. Six different models were generated based on Model 4 from Shin and Ma (2017), with these models containing the parameters $\bar{J}$, τ, and p. Each model differed in relation to whether the parameters were free to vary across feature dimension (colour, orientation) and sequencing condition (repetition, switch). Results from Experiment 1a revealed evidence to suggest that dimension switching had no impact on performance in this task, with substantial effects of change magnitude observed. Perhaps more crucially, formal model comparison revealed that the model wherein no parameter was free to vary across sequencing condition (the fixed model) was favoured. This suggested that the data obtained in this experiment was best explained by a model wherein switching has no influence on the allocation of resources or the precision of representations in memory, a finding which stood for the results of Experiment 1b. Experiment 1b was a conceptual replication of Experiment 1a with larger change magnitudes. Results again revealed no impact of dimension switching on behavioural measures of performance and formal model comparison again revealed the fixed model as the best, suggesting that dimension switching has little influence on resource allocation and memory encoding. As such, these results suggest that the previous theory surrounding a misallocation of resources to the irrelevant feature dimension on switch trials is incorrect as such misallocation would have been reflected by switch costs within the behavioural measures, as well as formal model comparison favouring a model wherein the p parameter was free to vary across sequencing condition.

6.3 Attentional control and visual short-term memory resource allocation

6.3.1 Support for attentional control as the mechanism for resource allocation

As stated previously, the aim of the present research was to investigate the role of attentional control in the allocation of vSTM resources to relevant information. Overall, results from the studies conducted within this thesis seemingly show that attentional control plays some part in how vSTM resources are allocated; however, this appears to be based on the methodology employed. Initial investigations using change detection tasks embedded within the dimension switching design showed a relatively consistent cost to performance as the result of switching feature dimensions; it was postulated that these costs may have arisen as the result of a misallocation of resources to the irrelevant feature dimension. This was based on previous task switching research which stated that to perform a given task, a *task set* must be activated; task sets can simply be thought of as a grouping of processes and representations necessary for performance of a task (see e.g., Grange & Houghton, 2014; Logan & Gordon, 2001; Logan & Schneider, 2010; Rogers & Monsell, 1995). More crucially to the present research however, task sets can be further deconstructed into *attentional sets* and *intentional sets*. While intentional sets are associated with the assignment of stimulus-response (S-R) mappings, the attentional set is responsible for the identification and selection of relevant information (see e.g., Kopp et al., 2006; Liefooghe & Verbruggen, 2019; Meiran, 2000; Rushworth et al., 2005, 2002; Yeung & Monsell, 2003b). Given that the task always remained the same in the present research (e.g., detect a change in the relevant feature dimension), the intentional set could remain the same while the attentional set required updating on dimension switch trials. Thus, a failure to sufficiently update the attentional set on a dimension switch trial would likely lead to vSTM resources being allocated to the irrelevant feature dimension, leading to reduced performance on switch trials.

Such a view appears to be supported from the results of CSI and RCI manipulations within Chapter 3. Previous research had shown that increasing the response-stimulus interval (RSI)—the duration between response on trial *N-1* and stimulus presentation on trial *N*—attenuated the extent of the task switch cost (see e.g., Allport et al., 1994; Rogers & Monsell, 1995). While Allport et al. (1994) attributed this reduction of the switch cost to allowing more time for the previously relevant task set to dissipate (i.e., reduce proactive interference), Rogers and Monsell (1995) stated that the reduction was due to more time being allowed for advance reconfiguration of the task set. Given that both the process of task set dissipation and advance reconfiguration occurred within the RSI, there was no way to dissociate between these processes using the paradigms employed by both studies. However, later work by Meiran (1996) facilitated such a dissociation; by cuing participants to which task should be performed on a given trial, this separated the RSI into the RCI—the duration between response on trial *N-1* and cue presentation on trial *N*—and the CSI—the duration between cue and stimulus presentation on trial *N*. During the RCI, the upcoming task is not yet known as the cue has not been presented, meaning this only allows for dissipation of the previously relevant task set. Alternatively, during the CSI the cue is known, which allows for reconfiguration of the task set; theoretically, task set dissipation can still occur during this time however, this interval is primarily associated with task set reconfiguration. Critically, CSI and RCI manipulations within the change detection switching task in Chapter 3 were used to determine whether proactive interference or reconfiguration impacted the extent of the dimension switch cost. Results revealed that the effect of interval duration showed evidence in favour of the null, indicating that interval duration did not influence the extent of the dimension switch cost.

Previously Rogers and Monsell (1995) hypothesised that increasing the RSI allows completion of an *endogenous control process* which is responsible for the updating of task sets. However, the finding that a task switch cost still presented when the RSI was large led the authors to hypothesise that an additional, *exogenous component* was required for the completion of task set reconfiguration. This exogenous component is only initiated upon presentation of the stimuli and therefore is unaffected by manipulations of preparation time. It could therefore be suggested that performance on the change detection switching task reflects this exogenous component only which is unaffected by manipulations of CSI and RCI; indeed, it would be plausible to assume that the process of resource allocation is only initiated upon stimulus presentation, with the dimension switch cost reflecting a disruption of this process due to insufficient configuration of the attentional set. Indeed, a more refined view of such disruption may be gained through the *executive control theory of visual attention* (ECTVA, see e.g., Logan & Gordon, 2001; see also Bundesen, 1990; Bundesen et al., 2005; Nosofsky & Palmeri, 1997). In simple terms, this theory proposes that executive control processes are used to manipulate parameters associated with subordinate processes, configuring them in such a way that allows for performance of an upcoming task. Two such parameters are β and π, representing a measure of response bias and attentional weighting of stimuli respectively; broadly speaking however, these parameters could also be taken as representations of intentional and attentional set. Therefore, given that the intentional set remained the same throughout, it could be suggested that the dimension switch cost originates from processes related to stimulus processing only (e.g., resource allocation, encoding), with this reflected by updating of the π parameter only in terms of the ECTVA model. Implementation of this model within the dimension switching design used in the present research may provide some additional insight into which aspects associated with task sets contribute to the dimension switch cost.

Experiment 2 within Chapter 3 investigated the contributions of cue switches to the dimension switch cost by implementing a 2:1 cues to task mapping. In typical versions of the *explicit cuing procedure*—wherein participants are presented with a cue at the beginning of each trial indicating which task to perform—a single cue is used per task. As such, when the task repeats, so too does the cue, with a switch in task also requiring a switch in cue. As such, some researchers suggested that the cue switch may in fact contribute to the overall task switch cost (see e.g., Logan & Bundesen, 2003; Mayr & Kliegl, 2003). To determine the contributions of a cue switch to the task switch cost, a 2:1 cues to task mapping was employed wherein two cues were used per task (e.g., for colour, the cues could be "colour" and "shade"). This gave rise to three sequencing conditions, cue repetitions, wherein both cue and task repeat, cue switches, wherein the cue switches and the task repeats, and task switches, wherein both cue and task switch. Here, the true impact of a task switch is reflected in the difference between a cue switch and task switch, whereas the influence of a cue switch is reflected in the difference between cue repetition and cue switch trials. Both Logan and Bundesen (2003) and Mayr and Kliegl (2003) found that cue switching does indeed contribute to the overall task switch cost however, neither of the theories surrounding proactive interference and task set reconfiguration offered an explanation for cue switch costs. Therefore, Logan and Bundesen (2003; see also Arrington and Logan, 2004a; Logan and Bundesen, 2004; D. W. Schneider and Logan, 2005) stated that participants encode both the cue and stimulus, with this *compound cue* used to retrieve the correct response from long-term memory (LTM). Critically, participants encode both the cue and stimulus on each trial and respond accordingly, meaning that there is no requirement to update task set as the task requirement would always be encode both the cue and stimulus. Thus, on this account, cue repetition trials receive a benefit through priming; as the cue had been encoded on the previous trial, this speeds up the process of cue encoding on the subsequent trial. Mayr and Kliegl (2003) also invoked a process of retrieval from LTM, suggesting that advance preparation requires retrieval of the appropriate S-R mapping from LTM, followed by implementation of the task when stimuli are presented. As such, a benefit is conferred on cue repetition trials as the retrieval path in LTM remains the same as the previous trial.

In general, the results from Experiment 2 revealed that cue switching does not contribute to the overall dimension switch cost. This is somewhat in line with the notion that the dimension switch cost reflects processes associated with changes in the attentional set only. As the cue in the present study simply indicated which feature dimension was relevant on a given trial and the task remained the same throughout (determine whether a change occurred in the relevant feature dimension), no further rules or parameters for task completion needed to be retrieved from LTM. Quite simply, the only aspect which required updating was which information was relevant on a given trial. As such, it could be posited that the cue switch cost reflects changes in intentional set. Given both theories of cue switch costs described above state that performance is highest when the cue repeats, it therefore follows that the poorer performance on trials wherein the cue switches is entirely due to the requirement to encode the cue in order to retrieve task rules from LTM; for instance, the compound cue encoding account states that performance on cue switches and task switches should be similar however, on a cue switch the task remains the same whereas on a task switch both cue and task change. This therefore suggests that despite the same task being performed, a change in cue requires a complete process of retrieval of task rules in the same manner as when both the cue and task switch. In the context of the present task, no such retrieval of task rules or S-R mappings was required as the cue did not indicate which task to perform, but rather indicated which feature dimension was relevant, therefore only necessitating the updating of attentional set.

6.3.2 Evidence against the requirement of attentional control for resource allocation

Overall, Chapters 2 and 3 appear to provide evidence in support of the view that attentional control is required for the appropriate allocation of vSTM resources to relevant information. However, in an attempt to extend these results and investigate the specific cause of the disruption by dimension switching, later studies employing continuous report (Chapter 4) and change localisation (Chapter 5) tasks revealed a distinct lack of evidence in support of such a view. The use of a continuous report task in Chapter 4 was to determine the cause of the dimension switch cost by fitting the *three-component mixture model* (see Bays et al., 2009) to the data. On this model, responses are a probabilistic mixture of 1) responding to the true target value based on a noisy memory representation, 2) responding to a non-target value, and 3) random guessing. When fit to data, the three-component model returns the parameters κ, an estimate of precision, p_t, the probability of responding to the target, p_{nt}, the probability of responding to the non-target, and p_u, the probability of guessing. Crucially, this mixture model enables identification of the source of error on continuous report tasks. Therefore, by including a dimension switching element within a continuous report task, differences in model parameters across sequencing condition (repetition, switch) would highlight the source of disruption as a result of dimension switching.

Results from this experiment revealed little evidence to suggest that dimension switching impacts continuous report performance. Indeed, the most notable findings were weak evidence in support of a switch cost in the p_{nt} parameter and stronger evidence in support of a switch cost in mean absolute error, with the latter suggesting that dimension switching does have some impact on performance, the three-component model is unable to establish precisely what has caused this. While overall there is little evidence to suggest dimension switching impacts continuous report performance, it is perhaps of interest that p_{nt} was the only parameter within the three-component model which showed evidence supporting a switch cost. If a true effect, this would suggest that dimension switching impacts the ability to bind relevant information. While previously suggested that the dimension switch cost was the result of a misallocation of resources to the irrelevant feature dimension, if we hypothetically assume that the switch cost in the p_{nt} parameter is a true effect, this would provide evidence against such a view and rather support the notion that dimension switching impacts the ability to bind information. As participants were cued to the relevant feature dimension on each trial, the non-target feature values used in the model fit routine were those belonging to the relevant feature dimension only. This therefore suggests that on switch trials, participants were less able to bind the relevant feature values with the spatial location of the stimuli, leading to what Bays et al. (2009) referred to as a swap error (e.g., the colours of two stimuli may be swapped in memory).

Previous research has shown that binding requires attention (see e.g., Hyun et al., 2009; Rensink, 2000; Treisman, 1988; Treisman & Gelade, 1980; Wan et al., 2020; Wheeler & Treisman, 2002; Zokaei et al., 2014); given that both the process of switching dimension and binding are attentionally demanding processes, it could be suggested that on switch trials, attention is depleted to such a point that binding becomes extremely difficult. Such a theoretical stance is perhaps more tenable when attention is considered as a resource. For instance, Oberauer (2019) provides a review of working memory and attention, discussing the notion of attention as a finite resource, highlighting three distinct classes; in one class, the attentional resource is responsible for attentional control of a given action. While attention may be drawn automatically to a certain stimulus when performing a task (e.g., when a stimulus is more salient than others), this does not require any attentional resource. Rather, it is the control of a given action (e.g., selectively attending a stimulus) which requires this attentional resource; crucially, this control can only be allocated to a single action at a given time. As such, it could be suggested that dimension switch costs observed within the present research are caused by a depletion of this attentional resource; dimension switching is an attentionally demanding process which could consume most of the attentional resource, leaving little for the process of binding upon stimulus presentation.

Alternatively, incomplete configuration of the attentional set during a switch in dimension could mean that the resource is prevented from controlling the action of binding relevant information upon stimulus presentation.

This view of misbinding would also account for the dimension switch costs observed within the studies implementing change detection tasks. Given that two stimulus displays are presented within these tasks, there is a possibility that misbinding could occur on both the memory and test displays however, misbinding on the memory display would likely have a greater impact on performance. For example, if on a switch trial wherein colour is relevant, if the colours of two stimuli are swapped, this would lead to the perception of a change in colour when the test display is subsequently presented. In fact, there is a possibility that this perception of a change in colour may be accentuated due to the appearance of a colour change in two of the stimuli. This would invariably lead to a change response, which is indeed reflected in the response bias measure across many of the change detection experiments within this thesis. While the level of evidence for a switch cost in response bias varied across experiments, it was consistently found that on switch trials response bias became more liberal relative to repetition trials, which could be explained by misbinding. In the case of both a depletion of attentional resources or prevention of the attentional resource from controlling binding, the notion of such a resource being responsible for the processes necessary for attending to and encoding stimuli into memory removes the need for a vSTM resource to explain dimension switch costs; on such an account, the limitation of vSTM would lie solely with the control of attention rather than the allocation of specific memory resources which are responsible for the precision of internal representations.

However, results from Chapter 5 of the present thesis provide some evidence against such a view and more critically, evidence against the notion of a role for attentional control in vSTM resource allocation. Behavioural analysis within the change localisation studies revealed that dimension switching had no impact on performance; while this may suggest that dimension switching did not impact vSTM resource allocation on this task, it is also possible that methodological limitations are responsible for the null results. For instance, the experiments may have been underpowered or, perhaps more critically, the data may have been too noisy given that it was collected using online means. In addition to the behavioural analysis, six models were developed to determine the influence of dimension switching on memory encoding and resource allocation. These models contained the parameters $\bar{J}$, τ, and p; the $\bar{J}$ and τ parameters reflect the precision of memory encoding whereas the p parameter reflects the amount of resource allocated to the irrelevant feature dimension. Thus, an increase in the p parameter value on switch trials would be indicative of a misallocation of resources to the irrelevant feature dimension, which had previously been suggested to be the cause of the dimension switch cost (see Chapter 2). While each model contained these parameters, whether these could vary across sequencing condition and feature dimension varied depending on the model in question. Accuracy data was used as input to these models and it was revealed that the fixed model was favoured over all other models in both experiments within Chapter 5.

In the fixed model, while the $\bar{J}$ and τ parameters were free to vary across feature dimension, all parameters including p were unable to vary across sequencing condition. That this model was favoured over other models—including those wherein different values of the p

parameter for repetition and switch trials—provides support for the results of the behavioural analysis, suggesting that a model wherein a single value can capture the process of resource allocation across both repetition and switch trials was sufficient to capture the data from these experiments. This therefore provides weight to the notion that attentional control is not required for vSTM resource allocation. In terms of previous supposition that dimension switch costs are caused by misbinding, from the results of Chapter 5, it is difficult to state with any certainty whether this is the case. Given that no dimension switch costs were observed, it could be suggested that misbinding is not responsible for the dimension switch costs observed in previous experiments or that some methodological issues may have contributed to the null effect. Additionally, none of the parameters within the models used in this study directly reflect the process of binding relevant information. Indeed, $\bar{J}$ and τ relate to the precision at which features are represented in memory, similar to the κ parameter returned by the three-component model (see e.g., Bays et al., 2009). Given that κ parameter values were found to be similar across sequencing condition within the continuous report task in Chapter 4, suggesting that the fidelity of feature representations was not impacted by dimension switching, it is unlikely that misbinding would have resulted in differences between the $\bar{J}$ and τ parameters in Chapter 5.

6.3.3 Summary of findings

The aim of the present thesis was to investigate the role of attentional control in the allocation of vSTM resources to relevant visual information contained on multivalent stimuli. This was investigated through the use task which probe vSTM wherein the relevant feature dimension could either repeat or switch; on switch trials, this introduced an element of attentional control required to perform the task, with poorer performance on switch relative to repetition trials indicative of insufficient attentional control. Taken together, the results of the present research are somewhat inconclusive. While dimension switch costs were observed across a range of measures for change detection tasks with different experimental manipulations (e.g., set size, interval duration), suggesting attentional control plays some role in vSTM resource allocation, other than relatively strong evidence in support of a switch cost in mean absolute error in the continuous report task, there was little evidence for such a cost within other behavioural measures in the continuous report and little evidence overall within the change localisation tasks. Furthermore, data from the continuous report and change localisation tasks were used in modelling which also revealed little evidence for disruption by dimension switching. Weak evidence was observed for a dimension switch cost in non-target response rate for modelling of continuous report data, with formal model comparison of models using change localisation data as input revealing that a model wherein parameters for encoding precision and resource allocation are fixed is sufficient to capture the data. Together, the results of the behavioural analysis and modelling provide weight to the notion that attentional control is not required for vSTM resource allocation.

However, it is also possible that the lack of dimension switch costs within chapters employing continuous report (Chapter 4) and change localisation (Chapter 5) tasks may be due to methodological limitations. These have been highlighted previously and are also discussed later in this chapter however, there is one limitation which may be of greatest importance, specifically that concerning whether the cue is essential for correct performance of the task. In Chapter 2, little evidence for a dimension switch cost was observed when a change occurred in the relevant feature dimension only; it was suggested that this may have led to participants ignoring the cue and monitoring the entire display, rather than encoding the cue and attending to the relevant feature dimension. As such, an additional, constant change in the irrelevant feature dimension was used to ensure that the cue was essential for correct task performance, resulting in dimension switch costs across a range of measures and experimental manipulations. Critically, within the continuous report and change localisation tasks, no mechanism was implemented to ensure that the cue was essential for correct performance of these tasks. Thus, it is possible that the lack of dimension switch costs observed in these studies is simply due to participants ignoring the cue and attending to the entire stimulus display, removing the need to switch between feature dimensions and negating any impact of dimension switching. To combat this issue, it would be beneficial to rerun these experiments with the addition of a mechanism to ensure that the cue is essential for correct performance of the task. While this is somewhat straightforward within change localisation tasks (i.e., adding a constant change in the irrelevant feature dimension like in the change detection tasks), such a mechanism is perhaps more difficult to implement within the continuous report task. However, given that modelling of change localisation data offers a greater insight into the cause of the dimension switch cost (e.g., the ability to investigate potential resource misallocation), it may be that change localisation may be the more preferred task in any case.

6.4 Limitations and future research

6.4.1 Methodological limitations

While alluded to in the previous section, it is necessary to discuss the limitations relating to the use of online testing within the present research. While use of online experimentation has increased within the last few years (even prior to the COVID-19 pandemic, see e.g., Dance, 2015), there are still some issues with the quality of data obtained from these sources. The most notable issue with online data collection concerning the research detailed within this thesis arose with the continuous report task. Given that at response, this task presents a wheel with 360 different options, it is perhaps understandable that the quality of data obtained from two online experiments was poor. This could be due to a range of different reasons; when colour was relevant for example, screen size, screen quality, or brightness may have impacted the ability to discern between similar colour values when responding. Furthermore, experiment timings are heavily reliant on secure internet connections; indeed, in a few instances within the online version of the continuous report task, it was apparent that the internet connection became unstable, resulting in a response being provided multiple times and with all responses subsequently appended to the data.

A further limitation of the experiments used within the present research, particularly within the change localisation study, is the number of trials used. For example, within Shin and Ma (2017), 1,200 trials per session were used in Experiment 1 compared to the 400 total trials used in each of the change localisation experiments within Chapter 5. Given that the experiments within Chapter 5 were conducted online, it would therefore be beneficial to run an in-person version of this task with an increased number of trials per condition in order to gain a more reliable picture of performance and indeed, provide more reliable results pertaining to model fitting. The duration of the retention interval may have also been a methodological limitation of the current research given that retention intervals have been shown to influence both precision and memory strength (see e.g., Bays et al., 2009; Brady et al., 2013; Nosofsky & Gold, 2016). As the retention interval did not fall below 1,000ms in any of the present studies, it could therefore be suggested that the performance observed was based on noisier or weaker memory representations, which may have clouded the true influence of dimension switching. Thus, implementing a range of retention intervals would allow a more accurate picture of performance within these paradigms.

Finally, there are also methodological limitations associated with participant screening; participants in the present study all self-reported normal or corrected to normal visual acuity and normal colour vision. Given the lack of access to validated online and in-person versions of the Ishihara Test, this was deemed to be the most appropriate option however, more rigorous methods should be used in any future research to ensure that participant screening is 100% accurate.

6.4.2 The impact of the Coronavirus pandemic

Given that the majority of the research discussed within this thesis was conducted during the Coronavirus (COVID-19) pandemic, it would be pertinent to discuss subsequent issues which arose as a result. Perhaps the most notable issue was the suspension of in-person testing from March 2020 to October 2021 at Keele University which disrupted plans to conduct all research in-person and required a switch to online testing. Initially, the research plan was to utilise the continuous report task for most of the experimental work, implementing a number of experimental manipulations (e.g., set size); however, given the need to transition research online, this posed a challenge for programming of the continuous report experiment. While the in-person version of the experiment was programmed in Python using PsychoPy (see Peirce et al., 2019), the School of Psychology had a subscription for online testing with Gorilla Experiment Builder (see Anwyl-Irvine et al., 2020) which uses JavaScript. This therefore required a complete rewrite of the experiment in another programming language, which required assistance from the team at Gorilla. Thus, it seemed appropriate to develop a similar set of experiments using change detection tasks which were much simpler to program while awaiting development of the online version of the continuous report task. Unfortunately, once completed, the data obtained from the online version of the continuous report task was extremely noisy; given that the task measures the precision of a participant's response, it is likely that extraneous variables influenced performance (e.g., screen size, brightness, the environment the task was performed in). As such, the data obtained from the online version of the continuous report task was essentially unusable, particularly when exclusion procedures were implemented which resulted in the removal of a large number of participants. Therefore, given that in-person testing was still unavailable at this stage, the online change localisation studies were developed. Once restrictions on in-person testing were relaxed, some data from the original version of the continuous report task was collected which provided much more usable data. However, given that previous restrictions on in-person teaching had resulted in many students deciding to attend university remotely, this greatly reduced the size of the participant pool for recruitment, meaning only 50 datasets were collected.

6.4.3 Future research

Given discussion concerning the limitations of the present research, perhaps the most obvious first step for future research would be to run each of the experiments within this thesis in an laboratory setting As alluded to previously, there is the potential for the data collected using online methods to be noisier than that collected in an in-person setting, which may have contributed to results within the present thesis. In-person testing not only offers a more controlled environment (e.g., calibrated monitors, noise, lighting), which leads to greater reliability, but may also highlight areas of discrepancy between online and in-person data, adding to the growing body of knowledge surrounding differences between the two methods of data collection. While the continuous report study detailed within this thesis was conducted in-person, it would also be of interest to repeat this study with a greater sample size or increased number of trials. While previous research (see e.g., Grange & Moore, 2022) has shown that 200 trials per cell of the experiment design are required to obtain good recovery of parameter values within the three-component model, increasing the number of trials would lead to more accurate model predictions which could offer greater insight into whether the weak evidence of a switch cost in the p_{nt} parameter (reflecting non-target response rate) is a true effect.

In terms of extending the theoretical contributions of this research, perhaps the most logical first step would be to examine the impact of dimension switching with varying set sizes. Much previous research has highlighted the importance of set size in various tasks probing vSTM performance (see e.g., Bays & Husain, 2008; Luck & Vogel, 1997; Vogel et al., 2001; Wilken & Ma, 2004; Zhang & Luck, 2008), therefore it would be of interest to determine whether the extent of the dimension switch cost observed in change detection switching tasks is influenced by set size, or if set size has any impact on results from the modelling outlined in Chapters 4 and 5. Based on previous research, it would be expected that set size would have a detrimental impact on performance, primarily accuracy. However, it is difficult to state with any large degree of certainty whether set size would impact the extent of the dimension switch cost; hypothetically, if misbinding is responsible for the dimension switch cost, it could be assumed that switch costs may be greater at larger set sizes given that there is an increase in the number of stimuli and thus, greater potential for misbinding.

It would also be beneficial to determine the impact of the duration of the retention interval on performance in these tasks. For instance, shorter retention intervals—which have been suggested to result in the generation of stronger memory representations (see e.g., Brady et al., 2013)—may in fact influence the extent of the dimension switch cost. In a similar manner, investigation of the possibility for an asymmetrical switch cost could provide greater insight into the role of attentional control in vSTM resource allocation. When switching between tasks of unequal difficulty, the asymmetrical switch cost reflects the heightened switch cost when switching from the difficult task to the easy task relative to switching from the easy to difficult task (see e.g., Meuter & Allport, 1999; Wylie & Allport, 2000; Yeung & Monsell, 2003a). The easier task is dominant therefore requiring greater top-down control when this task is not relevant (i.e., during performance of the more difficult, non-dominant task). Thus, when switching from the difficult task, more substantial switch costs (relative to switching from the easy task) present as a result of activating the task set associated with the easy task which had previously been well-inhibited (see e.g., Wu et al., 2015). It would therefore follow that greater attentional control would be required to inhibit a feature dimension that is somewhat automatically identified (e.g., colour) relative to a feature dimension that is less automatically identified (e.g., size).

In line with limitations discussed previously, implementation of a mechanism ensuring that the cue is essential for correct task performance in both continuous report and change localisation tasks would be essential for advancement of any theory. Given that the majority of dimension switch costs were observed within change detection tasks employing a constant change in the irrelevant feature dimension—making the cue essential for correct task performance—it is possible that the absence of such a mechanism within the continuous report and change localisation studies in the present thesis contributed to the general absence of dimension switch costs within these tasks. While such a mechanism is easily implemented in change localisation tasks by simply adding a constant change (either of a set or variable magnitude) in the irrelevant feature dimension as in the change detection tasks reported here, implementation of such a mechanism within a continuous report task is more problematic. Perhaps one method of identifying whether participants are attending to the entire stimulus display or the cued feature is to compare performance between the explicitly cued continuous report task and a continuous report task using retrocues. While in the former, participants could either attend to the cued feature or the entire stimulus display, the latter requires attendance to the entire stimulus display, with comparison of performance between methodologies perhaps indicating what participants are attending to upon stimulus presentation.

Perhaps most crucially to theoretical advancement however, is the ability to investigate the possibility of resource misallocation. While this is possible through the modelling of change localisation data, additional evidence obtained from mixture modelling would strengthen this theory. Within the present thesis, such misallocation could not be investigated using the continuous report task or mixture modelling. In the task used here, orientation could have influenced responses on colour trials (i.e., participants selected an orientation on the colourwheel rather than a colour) however, colour could not influence responses on the orientation task as responses to the orientation task were provided on a black wheel. Thus, in order to investigate the misallocation of resources using continuous report and mixture modelling, a task wherein responses are given using the same format (i.e., the probe screen for both features is identical) is required.

6.4.4 Conclusion

In conclusion, the present research aimed to investigate the role of attentional control in the allocation of vSTM resources to relevant information. This was achieved through a combination of tasks which probe vSTM performance and adaptation of task switching paradigms to require participants to either repeat or switch between relevant feature dimensions while performing the same task. Evidence from a number of studies using change detection tasks suggests that attentional control does play some role in vSTM resource allocation, with costs to performance observed when required to switch between feature dimensions relative to when the feature dimension repeated. Furthermore, it was found that manipulations of the CSI and RCI—which have been shown to influence the extent of the task switch cost—have no influence on the dimension switch cost and that cue switching does not contribute to the overall dimension switch cost. Together, results from these initial studies not only indicate a role for attentional control in vSTM resource allocation, but also that the origin of the dimension switch cost lies outside that associated with the widely reported task switch cost. However, later studies using continuous report and change localisation tasks are in direct opposition, in general finding little evidence for the presence of dimension switch costs, suggesting that attentional control is not required for resource allocation. Furthermore, modelling of data collected within these continuous report and change localisation studies provide support for this view, with perhaps most crucially, change localisation modelling revealing greatest support for a model wherein resource allocation is sufficiently captured by a single parameter rather than separate parameters for repetition and switch trials. While these results may be indicative of the lack of need for attentional control to allocate vSTM resources, weak evidence observed for a switch cost in non-target response rate within the continuous report task was found. It is speculated that this may reflect misbinding as the result of a depletion of attentional resources however, such a view will require much further research, not only due to the weak evidence, but also some methodological limitations of the present work. Thus, at present it appears that attentional control is required for vSTM resource allocation in some contexts as evidenced by dimension switch costs, but not others; however, it is possible that limitations of the research contained within this thesis meant that the methods used were insufficient to capture dimension switch costs within some experimental paradigms.

References

Ach, N. (1910–2006). *On Volition (Über den Willen)* (T. Herz, Trans.). Verlag von Quelle und Meyer.

Agaoglu, S., Agaoglu, M. N., Breitmeyer, B., & Ogmen, H. (2015). A statistical perspective to visual masking. *Vision Research*, *115*, 23–39. https://doi.org/10.1016/j.visres.2015.07.003

Allen, R. J., Baddeley, A. D., & Hitch, G. J. (2006). Is the binding of visual features in working memory resource-demanding? *Journal of Experimental Psychology: General*, *135*(2), 298–313. https://doi.org/10.1037/0096-3445.135.2.298

Allen, R. J., Hitch, G. J., Mate, J., & Baddeley, A. D. (2012). Feature binding and attention in working memory: A resolution of previous contradictory findings. *Quarterly Journal of Experimental Psychology*, *65*(12), 2369–2383. https://doi.org/10.1080/17470218.2012.687384

Allport, A. D., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance series. Attention and performance XV: Conscious and nonconscious information processing* (pp. 421–452). The MIT Press.

Allport, A. D., & Wylie, G. (2000). 'Task switching', stimulus-response bindings and negative priming. In S. Monsell & J. Driver (Eds.), *Control of Cognitive Processes: Attention and Performance XVIII* (pp. 35–70). The MIT Press.

Altmann, E. M., & Gray, W. D. (2008). An integrated model of cognitive control in task switching. *Psychological Review*, *115*(3), 602–639. https://doi.org/10.1037/0033-295X.115.3.602

Alvarez, G. A., & Cavanagh, P. (2004). The Capacity of Visual Short-Term Memory is Set Both by Visual Information Load and by Number of Objects. *Psychological Science*, *15*(2), 106–111. https://doi.org/10.1111/j.0963-7214.2004.01502006.x

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x

Arrington, C. M., & Logan, G. D. (2004a). Episodic and semantic components of the compound-stimulus strategy in the explicit task-cuing procedure. *Memory & Cognition*, *32*(6), 965–978. https://doi.org/10.3758/BF03196874

Arrington, C. M., & Logan, G. D. (2004b). The Cost of a Voluntary Task Switch. *Psychological Science*, *15*(9), 610–615. https://doi.org/10.1111/j.0956-7976.2004.00728.x

Arrington, C. M., Logan, L., Gordon D., & Schneider, D. W. (2007). Separating cue encoding from target processing in the explicit task-cuing procedure: Are there "true" task switch effects? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(3), 484–502. https://doi.org/10.1037/0278-7393.33.3.484

Arrington, C. M., & Weaver, S. M. (2015). Rethinking volitional control over task choice in multitask environments: Use of a stimulus set selection strategy in voluntary task switching. *Quarterly Journal of Experimental Psychology*, *68*(4), 664–679. https://doi.org/10.1080/17470218.2014.961935

Aust, F., & Barth, M. (2020). *papaja: Prepare reproducible APA journal articles with R Markdown*. https://github.com/crsh/papaja

Awh, E., Barton, B., & Vogel, E. K. (2007). Visual Working Memory Represents a Fixed Number of Items Regardless of Complexity. *Psychological Science*, *18*(7), 622–628. https://doi.org/10.1111/j.1467-9280.2007.01949.x

Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends in Cognitive Sciences*, *5*(3), 119–126. https://doi.org/10.1016/S1364-6613(00)01593-X

Baddeley, A. D., & Hitch, G. (1974). Working Memory. In *Psychology of Learning and Motivation* (Vol. 8, pp. 47–89). Elsevier. https://doi.org/10.1016/S0079-7421(08)60452-1

Barth, M. (2022). *tinylabels: Lightweight variable labels*. https://cran.r-project.org/package=tinylabels

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bates, D., Maechler, M., & Jagan, M. (2022). *Matrix: Sparse and dense matrix classes and methods*. https://CRAN.R-project.org/package=Matrix

Bays, P. M. (2014). Noise in Neural Populations Accounts for Errors in Working Memory.

Journal of Neuroscience, *34*(10), 3632–3645. https://doi.org/10.1523/JNEUROSCI.3204-13.2014

Bays, P. M. (2016). Evaluating and excluding swap errors in analogue tests of working memory. *Scientific Reports*, *6*(1), 19203. https://doi.org/10.1038/srep19203

Bays, P. M., Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*, *9*(10), 7–7. https://doi.org/10.1167/9.10.7

Bays, P. M., & Husain, M. (2008). Dynamic Shifts of Limited Working Memory Resources in Human Vision. *Science*, *321*(5890), 851–854. https://doi.org/10.1126/science.1158023

Berman, R., & Colby, C. (2009). Attention and active vision. *Vision Research*, *49*(10), 1233–1248. https://doi.org/10.1016/j.visres.2008.06.017

Bocincova, A., van Lamsweerde, A. E., & Johnson, J. S. (2017). The role of top-down suppression in mitigating the disruptive effects of task-irrelevant feature changes in visual working memory. *Memory & Cognition*, *45*(8), 1411–1422. https://doi.org/10.3758/s13421-017-0738-8

Brady, T. F., Konkle, T., Gill, J., Oliva, A., & Alvarez, G. A. (2013). Visual Long-Term Memory Has the Same Limit on Fidelity as Visual Working Memory. *Psychological Science*, *24*(6), 981–990. https://doi.org/10.1177/0956797612465439

Bryck, R. L., & Mayr, U. (2008). Task selection cost asymmetry without task switching. *Psychonomic Bulletin & Review*, *15*(1), 128–134. https://doi.org/10.3758/PBR.15.1.128

Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, *97*(4), 523–547. https://doi.org/10.1037/0033-295X.97.4.523

Bundesen, C., Habekost, T., & Kyllingsbæk, S. (2005). A Neural Theory of Visual Attention: Bridging Cognition and Neurophysiology. *Psychological Review*, *112*(2), 291–328. https://doi.org/10.1037/0033-295X.112.2.291

Burmester, A., & Wallis, G. (2012). Contrasting Predictions of Low- and High-Threshold Models for the Detection of Changing Visual Features. *Perception*, *41*(5), 505–516. https://doi.org/10.1068/p7176

Carlisle, N. B., Arita, J. T., Pardo, D., & Woodman, G. F. (2011). Attentional Templates in Visual Working Memory. *Journal of Neuroscience*, *31*(25), 9315–9322.

https://doi.org/10.1523/JNEUROSCI.1097-11.2011

Chen, Z., Cave, K. R., Basu, D., Suresh, S., & Wiltshire, J. (2020). A region complexity effect masquerading as object-based attention. *Journal of Vision*, *20*(24), 1–15. https://doi.org/10.1167/jov.20.7.24

Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A Taxonomy of External and Internal Attention. *Annual Review of Psychology*, *62*(1), 73–101. https://doi.org/10.1146/annurev.psych.093008.100427

Coltheart, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics*, *27*(3), 183–228. https://doi.org/10.3758/BF03204258

Cooper, R. P., Schwartz, M. F., Yule, P., & Shallice, T. (2005). The simulation of action disorganisation in complex activities of daily living. *Cognitive Neuropsychology*, *22*(8), 959–1004. https://doi.org/10.1080/02643290442000419

Cowan, N., Elliott, E. M., Scott Saults, J., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, *51*(1), 42–100. https://doi.org/10.1016/j.cogpsych.2004.12.001

Dance, A. (2015). How online studies are transforming psychology research. *Proceedings of the National Academy of Sciences*, *112*(47), 14399–14401. https://doi.org/10.1073/pnas.1520769112

Desimone, R., & Duncan, J. (1995). Neural Mechanisms of Selective Visual Attention. *Annual Review of Neuroscience*, *18*(1), 193–222. https://doi.org/10.1146/annurev.ne.18.030195.001205

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, *64*(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Dreisbach, G., Goschke, T., & Haider, H. (2007). The role of task rules and stimulus–response mappings in the task switching paradigm. *Psychological Research*, *71*(4), 383–392. https://doi.org/10.1007/s00426-005-0041-3

Dube, B., Emrich, S., & Al-Aidroos, N. (2017). More than a filter: Feature-based attention regulates the distribution of visual working memory resources. *Journal of Vision*, *17*(10), 206. https://doi.org/10.1167/17.10.206

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, *113*(4), 501–517. https://doi.org/10.1037/0096-3445.113.4.501

Egly, R., Driver, J., & Rafal, R. D. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General*, *123*(2), 161–177. https://doi.org/10.1037/0096-3445.123.2.161

Emrich, S. M., Lockhart, H. A., & Al-Aidroos, N. (2017). Attention mediates the flexible allocation of visual working memory resources. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(7), 1454–1465. https://doi.org/10.1037/xhp0000398

Eng, H. Y., Chen, D., & Jiang, Y. (2005). Visual working memory for simple and complex visual stimuli. *Psychonomic Bulletin & Review*, *12*(6), 1127–1133. https://doi.org/10.3758/BF03206454

Fougnie, D., & Alvarez, G. A. (2011). Object features fail independently in visual working memory: Evidence for a probabilistic feature-store model. *Journal of Vision*, *11*(12), 3–3. https://doi.org/10.1167/11.12.3

Frick, R. W. (1988). Issues of representation and limited capacity in the visuospatial sketchpad. *British Journal of Psychology*, *79*(3), 289–308. https://doi.org/10.1111/j.2044-8295.1988.tb02289.x

Friedman, N. P., & Miyake, A. (2004). The Relations Among Inhibition and Interference Control Functions: A Latent-Variable Analysis. *Journal of Experimental Psychology: General*, *133*(1), 101–135. https://doi.org/10.1037/0096-3445.133.1.101

García-Ogueta, M. I. (1993). Internal attentional switching: Effects of predictability, complexity and practice. *Acta Psychologica*, *83*(1), 13–32. https://doi.org/10.1016/0001-6918(93)90032-M

Goschke, T. (2000). Intentional reconfiguration and involuntary persistence in task-set switching. In S. Monsell & J. Driver (Eds.), *Control of Cognitive Processes: Attention and Performance XVIII* (pp. 331–355). The MIT Press.

Grange, J. A. (2022). *Trimr: An implementation of common response time trimming methods*. https://CRAN.R-project.org/package=trimr

Grange, J. A., & Houghton, G. (Eds.). (2014). *Task Switching and Cognitive Control*. Oxford University Press. https://doi.org/10.1093/acprof:osobl/9780199921959.001.0001

Grange, J. A., & Houghton, G. (2010). Cue-switch costs in task-switching: Cue priming or control processes? *Psychological Research*, *74*(5), 481–490. https://doi.org/10.1007/s00426-009-0270-y

Grange, J. A., & Moore, S. B. (2022). Mixtur: An R package for designing, analysing, and modelling continuous report visual short-term memory studies. *Behavior Research Methods*. https://doi.org/10.3758/s13428-021-01688-1

Grange, J. A., & Moore, S. B. (2021). *Mixtur: Modelling continuous report visual short-term memory studies*. https://CRAN.R-project.org/package=mixtur

Hautus, M. J. (1995). Corrections for extreme proportions and their biasing effects on estimated values of d$'$. *Behavior Research Methods, Instruments, & Computers*, *27*(1), 46–51. https://doi.org/10.3758/BF03203619

Henderson, S. E., Lockhart, H. A., Davis, E. E., Emrich, S. M., & Campbell, K. L. (2020). Reduced Attentional Control in Older Adults Leads to Deficits in Flexible Prioritization of Visual Working Memory. *Brain Sciences*, *10*(8), 542. https://doi.org/10.3390/brainsci10080542

Henry, L., & Wickham, H. (2020). *Purrr: Functional programming tools*. https://CRAN.R-project.org/package=purrr

Hollingworth, A., Richard, A. M., & Luck, S. J. (2008). Understanding the function of visual short-term memory: Transsaccadic memory, object correspondence, and gaze correction. *Journal of Experimental Psychology: General*, *137*(1), 163–181. https://doi.org/10.1037/0096-3445.137.1.163

Horoufchin, H., Philipp, A. M., & Koch, I. (2011a). Temporal distinctiveness and repetition benefits in task switching: Disentangling stimulus-related and response-related contributions. *Quarterly Journal of Experimental Psychology*, *64*(3), 434–446. https://doi.org/10.1080/17470218.2010.496857

Horoufchin, H., Philipp, A. M., & Koch, I. (2011b). The dissipating task-repetition benefit in cued task switching: Task-set decay or temporal distinctiveness. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(2), 455–472.

https://doi.org/10.1037/a0020557

Horowitz, T. S., & Wolfe, J. M. (1998). Visual search has no memory. *Nature*, *394*(6693), 575–577. https://doi.org/10.1038/29068

Houghton, G., & Tipper, S. P. (1994). A model of inhibitory mechanisms in selective attention. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 53–123). Academic Press.

Hsieh, S. (2012). Two decades of research on task switching: What more can we ask? *Chinese Journal of Psychology*, *54*(1), 67–93.

Huynh Cong, S., & Kerzel, D. (2021). Allocation of resources in working memory: Theoretical and empirical implications for visual search. *Psychonomic Bulletin & Review*, *28*(4), 1093–1111. https://doi.org/10.3758/s13423-021-01881-5

Hyun, J., Woodman, G. F., Vogel, E. K., Hollingworth, A., & Luck, S. J. (2009). The comparison of visual working memory representations with perceptual inputs. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(4), 1140–1160. https://doi.org/10.1037/a0015019

Irwin, D. E. (1991). Information integration across saccadic eye movements. *Cognitive Psychology*, *23*(3), 420–456. https://doi.org/10.1016/0010-0285(91)90015-G

Irwin, D. E. (1992). Visual Memory Within and Across Fixations. In K. Rayner (Ed.), *Eye Movements and Visual Cognition* (pp. 146–165). Springer New York. https://doi.org/10.1007/978-1-4612-2852-3_9

Irwin, D. E., & Andrews, R. V. (1996). Integration and accumulation of information across saccadic eye movements. In T. Inui & J. L. McClelland (Eds.), *Attention and performance 16: Information integration in perception and communication* (pp. 125–155). The MIT Press.

Jaswal, S., & Logie, R. H. (2011). Configural encoding in visual feature binding. *Journal of Cognitive Psychology*, *23*(5), 586–603. https://doi.org/10.1080/20445911.2011.570256

Jersild, A. T. (1927). Mental set and shift. *Archives of Psychology*, *14, 89*, 81.

Kassambara, A. (2020). *Ggpubr: 'ggplot2' based publication ready plots.* https://CRAN.R-project.org/package=ggpubr

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I.

(2010). Control and interference in task switching—A review. *Psychological Bulletin*, *136*(5), 849–874. https://doi.org/10.1037/a0019842

Kiyonaga, A., & Egner, T. (2013). Working memory as internal attention: Toward an integrative account of internal and external selection processes. *Psychonomic Bulletin & Review*, *20*(2), 228–242. https://doi.org/10.3758/s13423-012-0359-y

Kopp, B., Tabeling, S., Moschner, C., & Wessel, K. (2006). Fractionating the Neural Mechanisms of Cognitive Control. *Journal of Cognitive Neuroscience*, *18*(6), 949–965. https://doi.org/10.1162/jocn.2006.18.6.949

Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian Cognitive Modeling: A Practical Course*. Cambridge University Press. https://doi.org/10.1017/CBO9781139087759

Lewandowsky, S., & Farrell, S. (2011). *Computational Modeling in Cognition: Principles and Practice*. SAGE Publications, Inc. https://doi.org/10.4135/9781483349428

Lhermitte, F. (1983). 'Utilization behaviour' and its relation to lesions of the frontal lobe. *Brain*, *106*(2), 237–255. https://doi.org/10.1093/brain/106.2.237

Liefooghe, B., & Verbruggen, F. (2019). On the Assimilation of Instructions: Stimulus-response Associations are Implemented but not Stimulus-task Associations. *Journal of Cognition*, *2*(1), 20. https://doi.org/10.5334/joc.78

Logan, G. D. (1980). Attention and automaticity in Stroop and priming tasks: Theory and data. *Cognitive Psychology*, *12*(4), 523–553. https://doi.org/10.1016/0010-0285(80)90019-5

Logan, G. D., & Bundesen, C. (2004). Very clever homunculus: Compound stimulus strategies for the explicit task-cuing procedure. *Psychonomic Bulletin & Review*, *11*(5), 832–840. https://doi.org/10.3758/BF03196709

Logan, G. D., & Bundesen, C. (2003). Clever homunculus: Is there an endogenous act of control in the explicit task-cuing procedure? *Journal of Experimental Psychology: Human Perception and Performance*, *29*(3), 575–599. https://doi.org/10.1037/0096-1523.29.3.575

Logan, G. D., & Gordon, R. D. (2001). Executive control of visual attention in dual-task situations. *Psychological Review*, *108*(2), 393–434. https://doi.org/10.1037/0033-295X.108.2.393

Logan, G. D., & Schneider, D. W. (2010). Distinguishing Reconfiguration and

Compound-cue Retrieval in Task Switching. *Psychologica Belgica*, *50*(3-4), 413. https://doi.org/10.5334/pb-50-3-4-413

Logan, G. D., & Schneider, D. W. (2006). Interpreting instructional cues in task switching procedures: The role of mediator retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(2), 347–363. https://doi.org/10.1037/0278-7393.32.3.347

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–281. https://doi.org/10.1038/36846

Luria, R., & Vogel, E. K. (2011). Shape and color conjunction stimuli are represented as bound objects in visual working memory. *Neuropsychologia*, *49*(6), 1632–1639. https://doi.org/10.1016/j.neuropsychologia.2010.11.031

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience*, *17*(3), 347–356. https://doi.org/10.1038/nn.3655

Machizawa, M. G., Goh, C. C. W., & Driver, J. (2012). Human Visual Short-Term Memory Precision Can Be Varied at Will When the Number of Retained Items Is Low. *Psychological Science*, *23*(6), 554–559. https://doi.org/10.1177/0956797611431988

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*(2), 163–203. https://doi.org/10.1037/0033-2909.109.2.163

Macmillan, N. A., & Creelman, C. D. (2004). *Detection Theory: A User's Guide* (0th ed.). Psychology Press. https://doi.org/10.4324/9781410611147

Markov, Y. A., Tiurina, N. A., & Utochkin, I. S. (2019). Different features are stored independently in visual working memory but mediated by object-based representations. *Acta Psychologica*, *197*, 52–63. https://doi.org/10.1016/j.actpsy.2019.05.003

Marshall, L., & Bays, P. M. (2013). Obligatory encoding of task-irrelevant features depletes working memory resources. *Journal of Vision*, *13*(2), 21–21. https://doi.org/10.1167/13.2.21

Mathôt, S. (2018). Pupillometry: Psychology, Physiology, and Function. *Journal of Cognition*, *1*(1), 16. https://doi.org/10.5334/joc.18

Mayr, U., & Kliegl, R. (2000). Task-set switching and long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(5), 1124–1140.

https://doi.org/10.1037/0278-7393.26.5.1124

Mayr, U., & Kliegl, R. (2003). Differential effects of cue changes and task changes on task-set selection costs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(3), 362–372. https://doi.org/10.1037/0278-7393.29.3.362

Meiran, N. (2010). Task Switching: Mechanisms Underlying Rigid vs. Flexible Self-Control. In R. Hassin, K. Ochsner, & Y. Trope (Eds.), *Self Control in Society, Mind, and Brain* (pp. 202–220). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195391381.003.0011

Meiran, N. (2000). Modeling cognitive control in task-switching. *Psychological Research*, *63*(3-4), 234–249. https://doi.org/10.1007/s004269900004

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(6), 1423–1442. https://doi.org/10.1037/0278-7393.22.6.1423

Meiran, N., Chorev, Z., & Sapir, A. (2000). Component Processes in Task Switching. *Cognitive Psychology*, *41*(3), 211–253. https://doi.org/10.1006/cogp.2000.0736

Meiran, N., & Marciano, H. (2002). Limitations in advance task preparation: Switching the relevant stimulus dimension in speeded same—different comparisons. *Memory & Cognition*, *30*(4), 540–550. https://doi.org/10.3758/BF03194955

Meuter, R. F. I., & Allport, A. (1999). Bilingual Language Switching in Naming: Asymmetrical Costs of Language Selection. *Journal of Memory and Language*, *40*(1), 25–40. https://doi.org/10.1006/jmla.1998.2602

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81–97. https://doi.org/10.1037/h0043158

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex “Frontal Lobe” Tasks: A Latent Variable Analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, *7*(3), 134–140. https://doi.org/10.1016/S1364-6613(03)00028-7

Monsell, S., & Mizon, G. A. (2006). Can the task-cuing paradigm measure an endogenous task-set reconfiguration process? *Journal of Experimental Psychology: Human Perception and Performance*, *32*(3), 493–516. https://doi.org/10.1037/0096-1523.32.3.493

Monsell, S., Yeung, N., & Azuma, R. (2000). Reconfiguration of task-set: Is it easier to switch to the weaker task? *Psychological Research*, *63*(3-4), 250–264. https://doi.org/10.1007/s004269900005

Moore, S. B. (2022). *Sdtt: Simple detection theory tool*. https://github.com/s-b-moore/sdtt

Morey, R. D., & Rouder, J. N. (2021). *BayesFactor: Computation of bayes factors for common designs*. https://CRAN.R-project.org/package=BayesFactor

Müller, K., & Wickham, H. (2022). *Tibble: Simple data frames*. https://CRAN.R-project.org/package=tibble

Nosofsky, R. M., & Gold, J. (2016). Memory strength versus memory variability in visual change detection. *Attention, Perception, & Psychophysics*, *78*(1), 78–93. https://doi.org/10.3758/s13414-015-0992-4

Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review*, *104*(2), 266–300. https://doi.org/10.1037/0033-295X.104.2.266

Oberauer, K. (2019). Working Memory and Attention – A Conceptual Analysis and Review. *Journal of Cognition*, *2*(1), 36. https://doi.org/10.5334/joc.58

Olson, I. R., & Jiang, Y. (2002). Is visual short-term memory object based? Rejection of the "strong-object" hypothesis. *Perception & Psychophysics*, *64*(7), 1055–1067. https://doi.org/10.3758/BF03194756

Ooms, J. (2021). *Magick: Advanced graphics and image-processing in r*. https://CRAN.R-project.org/package=magick

Palmer, J. (1990). Attentional limits on the perception and memory of visual information. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(2), 332–350. https://doi.org/10.1037/0096-1523.16.2.332

Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics*, *44*(4), 369–378. https://doi.org/10.3758/BF03210419

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E.,

& Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Persuh, M., Genzer, B., & Melara, R. D. (2012). Iconic memory requires attention. *Frontiers in Human Neuroscience*, *6*. https://doi.org/10.3389/fnhum.2012.00126

Pertzov, Y., Manohar, S., & Husain, M. (2017). Rapid forgetting results from competition over time between items in visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(4), 528–536. https://doi.org/10.1037/xlm0000328

Philipp, A. M., Gade, M., & Koch, I. (2007). Inhibitory processes in language switching: Evidence from switching language-defined response sets. *European Journal of Cognitive Psychology*, *19*(3), 395–416. https://doi.org/10.1080/09541440600758812

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics*, *16*(2), 283–290. https://doi.org/10.3758/BF03203943

Plummer, M., Best, N., Cowles, K., & Vines, K. (2006). CODA: Convergence diagnosis and output analysis for MCMC. *R News*, *6*(1), 7–11. https://journal.r-project.org/archive/

Prinzmetal, W., Amiri, H., Allen, K., & Edwards, T. (1998). Phenomenology of attention: I. Color, location, orientation, and spatial frequency. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(1), 261–282. https://doi.org/10.1037/0096-1523.24.1.261

Purdy, J. E., Eimann, D. G., & Cross, H. A. (1980). Persistence of a briefly presented visual stimulus in sensory memory. *Bulletin of the Psychonomic Society*, *16*(5), 374–376. https://doi.org/10.3758/BF03329569

R Core Team. (2022). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.R-project.org/

Rensink, R. A. (2000). The Dynamic Representation of Scenes. *Visual Cognition*, *7*(1-3), 17–42. https://doi.org/10.1080/135062800394667

Rensink, R. A. (2014). Limits to the usability of iconic memory. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00971

Rensink, R. A. (2001). Change Blindness: Implications for the Nature of Visual Attention. In M. Jenkin & L. Harris (Eds.), *Vision and Attention* (pp. 169–188). Springer New

York. https://doi.org/10.1007/978-0-387-21591-4_9

Reuter, B., Philipp, A. M., Koch, I., & Kathmann, N. (2006). Effects of switching between leftward and rightward pro- and antisaccades. *Biological Psychology*, *72*(1), 88–95. https://doi.org/10.1016/j.biopsycho.2005.08.005

Rogers, R. D., & Monsell, S. (1995). Costs of a predictible switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, *124*(2), 207–231. https://doi.org/10.1037/0096-3445.124.2.207

Rouder, J. N., Morey, R. D., Morey, C. C., & Cowan, N. (2011). How to measure working memory capacity in the change detection paradigm. *Psychonomic Bulletin & Review*, *18*(2), 324–330. https://doi.org/10.3758/s13423-011-0055-3

Roussy, M., Mendoza-Halliday, D., & Martinez-Trujillo, J. C. (2021). Neural Substrates of Visual Perception and Working Memory: Two Sides of the Same Coin or Two Different Coins? *Frontiers in Neural Circuits*, *15*, 764177. https://doi.org/10.3389/fncir.2021.764177

Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(4), 763–797. https://doi.org/10.1037/0096-1523.27.4.763

Rushworth, M. F. S., Passingham, R. E., & Nobre, A. C. (2005). Components of Attentional Set-switching. *Experimental Psychology*, *52*(2), 83–98. https://doi.org/10.1027/1618-3169.52.2.83

Rushworth, M. F. S., Passingham, R. E., & Nobre, A. C. (2002). Components of Switching Intentional Set. *Journal of Cognitive Neuroscience*, *14*(8), 1139–1150. https://doi.org/10.1162/089892902760807159

Santee, J. L., & Egeth, H. E. (1980). Selective attention in the speeded classification and comparison of multidimensional stimuli. *Perception & Psychophysics*, *28*(3), 191–204. https://doi.org/10.3758/BF03204375

Schneider, D. W. (2015). Attentional control of response selection in task switching. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(5), 1315–1324. https://doi.org/10.1037/xhp0000091

Schneider, D. W., & Logan, G. D. (2014). Modelling Response Selection in Task Switching:

Testing the Contingent Encoding Assumption. *Quarterly Journal of Experimental Psychology*, *67*(6), 1074–1095. https://doi.org/10.1080/17470218.2013.843009

Schneider, D. W., & Logan, G. D. (2005). Modeling Task Switching Without Switching Tasks: A Short-Term Priming Account of Explicitly Cued Performance. *Journal of Experimental Psychology: General*, *134*(3), 343–367. https://doi.org/10.1037/0096-3445.134.3.343

Schneider, D. W., & Logan, G. D. (2007). Task switching versus cue switching: Using transition cuing to disentangle sequential effects in task-switching performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(2), 370–378. https://doi.org/10.1037/0278-7393.33.2.370

Schneider, D. W., & Logan, G. D. (2011). Task-switching performance with 1:1 and 2:1 cue–task mappings: Not so different after all. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(2), 405–415. https://doi.org/10.1037/a0021967

Schneider, D., Bonmassar, C., & Hickey, C. (2018). Motivation and short-term memory in visual search: Attention's accelerator revisited. *Cortex*, *102*, 45–56. https://doi.org/10.1016/j.cortex.2017.06.022

Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition*, *80*(1-2), 1–46. https://doi.org/10.1016/S0010-0277(00)00152-9

Schönbrodt, F. D., & Wagenmakers, E.-J. (2018). Bayes factor design analysis: Planning for compelling evidence. *Psychonomic Bulletin & Review*, *25*(1), 128–142. https://doi.org/10.3758/s13423-017-1230-y

Schurgin, M. W., Wixted, J. T., & Brady, T. F. (2020). Psychophysical scaling reveals a unified theory of visual memory strength. *Nature Human Behaviour*, *4*, 1156–1172. https://doi.org/10.1038/s41562-020-00938-0

Shen, M., Tang, N., Wu, F., Shui, R., & Gao, Z. (2013). Robust object-based encoding in visual working memory. *Journal of Vision*, *13*(2), 1–11. https://doi.org/10.1167/13.2.1

Shin, H., & Ma, W. J. (2016). Crowdsourced single-trial probes of visual working memory for irrelevant features. *Journal of Vision*, *16*(5), 10. https://doi.org/10.1167/16.5.10

Shin, H., & Ma, W. J. (2017). Visual short-term memory for oriented, colored objects. *Journal of Vision*, *17*(9), 12. https://doi.org/10.1167/17.9.12

Siegel, A. L. M., & Castel, A. D. (2018). The role of attention in remembering important item-location associations. *Memory & Cognition*, *46*(8), 1248–1262. https://doi.org/10.3758/s13421-018-0834-4

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2022). *Afex: Analysis of factorial experiments*. https://CRAN.R-project.org/package=afex

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(2), 174–215. https://doi.org/10.1037/0278-7393.6.2.174

Spector, A., & Biederman, I. (1976). Mental Set and Mental Shift Revisited. *The American Journal of Psychology*, *89*(4), 669. https://doi.org/10.2307/1421465

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, *74*(11), 1–29. https://doi.org/10.1037/h0093759

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. https://doi.org/10.1037/h0054651

Sudevan, P., & Taylor, D. A. (1987). The cuing and priming of cognitive operations. *Journal of Experimental Psychology: Human Perception and Performance*, *13*(1), 89–103. https://doi.org/10.1037/0096-1523.13.1.89

Torchiano, M. (2020). *Effsize: Efficient effect size computation*. https://doi.org/10.5281/zenodo.1480624

Treisman, A. (1988). Features and Objects: The Fourteenth Bartlett Memorial Lecture. *The Quarterly Journal of Experimental Psychology Section A*, *40*(2), 201–237. https://doi.org/10.1080/02724988843000104

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, *14*(1), 107–141. https://doi.org/10.1016/0010-0285(82)90006-8

Treisman, A., & Zhang, W. (2006). Location and binding in visual working memory. *Memory & Cognition*, *34*(8), 1704–1719. https://doi.org/10.3758/BF03195932

van den Berg, R., Shin, H., Chou, W.-C., George, R., & Ma, W. J. (2012). Variability in encoding precision accounts for visual short-term memory limitations. *Proceedings of the National Academy of Sciences*, *109*(22), 8780–8785. https://doi.org/10.1073/pnas.1117465109

van Moorselaar, D., Gunseli, E., Theeuwes, J., & Olivers, C., N. L. (2015). The time course of protecting a visual memory representation from perceptual interference. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.01053

Vandierendonck, A., Liefooghe, B., & Verbruggen, F. (2010). Task switching: Interplay of reconfiguration and interference control. *Psychological Bulletin*, *136*(4), 601–626. https://doi.org/10.1037/a0019791

Vecera, S. P., & Farah, M. J. (1994). Does visual attention select objects or locations? *Journal of Experimental Psychology: General*, *123*(2), 146–160. https://doi.org/10.1037/0096-3445.123.2.146

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(1), 92–114. https://doi.org/10.1037/0096-1523.27.1.92

Wan, X., Zhou, Y., Wu, F., He, K., Shen, M., & Gao, Z. (2020). The role of attention in retaining the binding of integral features in working memory. *Journal of Vision*, *20*(7), 16. https://doi.org/10.1167/jov.20.7.16

Waszak, F., Hommel, B., & Allport, A. (2003). Task-switching and long-term priming: Role of episodic stimulus–task bindings in task-shift costs. *Cognitive Psychology*, *46*(4), 361–413. https://doi.org/10.1016/S0010-0285(02)00520-0

Waszak, F., Hommel, B., & Allport, A. (2005). Interaction of task readiness and automatic retrieval in task switching: Negative priming and competitor priming. *Memory & Cognition*, *33*(4), 595–610. https://doi.org/10.3758/BF03195327

Waszak, F., Hommel, B., & Allport, A. (2004). Semantic generalization of stimulus-task bindings. *Psychonomic Bulletin & Review*, *11*(6), 1027–1033. https://doi.org/10.3758/BF03196732

Wheeler, M. E., & Treisman, A. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, *131*(1), 48–64. https://doi.org/10.1037/0096-3445.131.1.48

Wickham, H. (2011). The split-apply-combine strategy for data analysis. *Journal of Statistical Software*, *40*(1), 1–29. https://www.jstatsoft.org/v40/i01/

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis.* Springer-Verlag New York. https://ggplot2.tidyverse.org

Wickham, H. (2019). *Stringr: Simple, consistent wrappers for common string operations.* https://CRAN.R-project.org/package=stringr

Wickham, H. (2021). *Forcats: Tools for working with categorical variables (factors).* https://CRAN.R-project.org/package=forcats

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wickham, H., François, R., Henry, L., & Müller, K. (2022). *Dplyr: A grammar of data manipulation.* https://CRAN.R-project.org/package=dplyr

Wickham, H., & Girlich, M. (2022). *Tidyr: Tidy messy data.* https://CRAN.R-project.org/package=tidyr

Wickham, H., Hester, J., & Bryan, J. (2022). *Readr: Read rectangular text data.* https://CRAN.R-project.org/package=readr

Wilke, C. O. (2020). *Cowplot: Streamlined plot theme and plot annotations for 'ggplot2'.* https://CRAN.R-project.org/package=cowplot

Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision*, *4*(12), 11. https://doi.org/10.1167/4.12.11

Wolfe, J. M. (1999). Inattentional amnesia. In V. Coltheart (Ed.), *Fleeting memories: Cognition of brief visual stimuli.* The MIT Press.

Wu, S., Hitchman, G., Tan, J., Zhao, Y., Tang, D., Wang, L., & Chen, A. (2015). The neural dynamic mechanisms of asymmetric switch costs in a combined Stroop-task-switching paradigm. *Scientific Reports*, *5*(1), 10240. https://doi.org/10.1038/srep10240

Wylie, G., & Allport, A. (2000). Task switching and the measurement of "switch costs". *Psychological Research*, *63*(3-4), 212–233. https://doi.org/10.1007/s004269900003

Xu, Y. (2002). Encoding color and shape from different parts of an object in visual short-term memory. *Perception & Psychophysics*, *64*(8), 1260–1280. https://doi.org/10.3758/BF03194770

Yantis, S. (1998). Control of visual attention. In H. Pashler (Ed.), *Attention* (pp. 223–256). Psychology Press.

Yeung, N., & Monsell, S. (2003a). Switching between tasks of unequal familiarity: The role of stimulus-attribute and response-set selection. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(2), 455–469. https://doi.org/10.1037/0096-1523.29.2.455

Yeung, N., & Monsell, S. (2003b). The effects of recent practice on task switching. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(5), 919–936. https://doi.org/10.1037/0096-1523.29.5.919

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233–235. https://doi.org/10.1038/nature06860

Zhang, W., & Luck, S. J. (2011). The Number and Quality of Representations in Working Memory. *Psychological Science*, *22*(11), 1434–1441. https://doi.org/10.1177/0956797611417006

Zokaei, N., Heider, M., & Husain, M. (2014). Attention is Required for Maintenance of Feature Binding in Visual Working Memory. *Quarterly Journal of Experimental Psychology*, *67*(6), 1191–1213. https://doi.org/10.1080/17470218.2013.852232

Keele University FNS Psychology Faculty Research Ethics Committee
psychology.ethics@keele.ac.uk

Dear Stuart

Project Title:	Attentional control of visual short-term memory resources
REC Project Reference:	PS-200137
Type of Application	Amendment

Keele University's Psychology Research Ethics Committee (PSY-FREC) reviewed the above project application

Favourable Ethical opinion
The members of the Committee gave a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation, subject to the conditions specified below.

Conditions of the favourable opinion
The favourable opinion is subject to the following conditions being met prior to the start of the project

.	None

Reporting requirements
The University's standard operating procedures give detailed guidance on reporting requirements for studies with a favourable opinion including:

- Notifying substantial amendments
- Notifying issues which may have an impact upon ethical opinion of the study
- Progress reports
- Notifying the end of the study

Approved documents
The documents reviewed and approved are:

Document	**Version**	**Date**
https://drive.google.com/open?id=1nenPme2Ln6ize00oZkgYNr5hdlFOdFww		01.05.20
https://drive.google.com/open?id=1aG5Ys9ASQnsITLSGqafTjOqp0CezT-l1		01.05.20
https://drive.google.com/open?id=1m1_Zf89OK7awcHsjVT1knOGiKiPAAULO		01.05.20
https://drive.google.com/open?id=1K6UYSO33CPzi0Orm74p5-hCznIoCTHqd		01.05.20

Yours sincerely,

Dr Joseph Brooks, Chair

Figure 1. Figure showing an example trial procedure within a change detection task using coloured squares as stimuli.

Figure captions

Figure 2. Figure showing an example trial procedure within the continuous (colour) report task.

Figure 3. Probability density functions for different values of κ, the concentration parameter of the von Mises distribution. A higher value of κ indicates more precise memory representations.

Figure 4. Figure showing stimuli used in Experiments 1 and 2. Note that in Experiments 1b and 2b a change also occurred in the irrelevant feature dimension.

Figure 5. Plots displaying mean a) accuracy, b) response time, c) sensitivity (d'), and d) response bias (c) across each sequencing condition in Experiment 1 (PR = pure repetition, MR = mixed repetition, MS = mixed switch). Error bars represent the standard error of the mean.

Figure 6. Plots displaying mean a) accuracy, b) response time, c) sensitivity (d'), and d) response bias (c) across each sequencing condition in Experiment 2 (PR = pure repetition, MR = mixed repetition, MS = mixed switch). Error bars represent the standard error of the mean.

Figure 7. Figure showing stimuli used in Experiment 3.

Figure 8. Plots displaying mean a) accuracy, b) response time, c) sensitivity (d'), and d) response bias (c) across each sequencing condition in Experiment 3 (MR = mixed repetition, MS = mixed switch). Error bars represent the standard error of the mean.

Figure 9. Figure showing example trial procedures in a) Experiment 1a and b) Experiment 1b. Note that a change occurred in both the relevant and irrelevant feature dimensions. The response screen (indicated with a question mark) was blank during the each experiment.

Figure 10. Plots displaying mean a) accuracy, b) response time, c) sensitivity (d'), and d) response bias (c) across sequencing condition and interval duration for Experiment 1a and Experiment 1b. Error bars represent the standard error of the mean.

Figure 11. Figure showing example trial procedure in Experiment 2. Note that a change occurred in both the relevant and irrelevant feature dimensions.

Figure 12. Plots displaying mean a) accuracy, b) response time, c) sensitivity (d'), and d) response bias (c) across sequencing condition (CR = cue repetition, CS = cue switch, DS = dimension switch) in Experiment 2. Error bars represent the standard error of the mean.

Figure 13. Figure showing example trial procedure in Experiment 1. Note that if orientation was relevant, the wheel presented on the probe display was black.

Figure 14. Plots displaying a) mean values for each of the parameters returned by the three-component model, b) participant-averaged response error (points) with overlaid model fit (solid line), and c) mean response time across sequencing condition in Experiment 1. Error bars represent the standard error of the mean.

Figure 15. Figure showing example trial procedure in Experiment 1. Note that the same stimuli were used in both Experiment 1a and 1b with the only difference between experiments being the change magnitudes used. The change in this example occurred in the top-right stimulus.

Figure 16. Plots displaying a) fixed model fit and b) free tau model fit in Experiment 1a. Note that the fixed model provided the best fit to the data, followed by the free tau model.

Figure 17. Plots displaying the remaining model fits in Experiment 1a; a) free j-bar model, b) free p model, c) full model, d) fully-free model.

Figure 18. Plots displaying a) fixed model fit and b) free-j model fit in Experiment 1b. Note that the fixed model provided the best fit to the data, followed by the free-j model.

Figure 19. Plots displaying the remaining model fits in Experiment 1b; a) free-p model, b) free-tau model, c) full model, d) fully-free model.

Table captions

Table 1. Table displaying Bayes Factors (BF_{10}) for each model across accuracy, response time, sensitivity, and response bias analyses in Experiment 1a. The BF_{10} for each model is in comparison to the null model.

Table 2. Table displaying Bayes Factors (BF_{10}) for each model across accuracy, response time, sensitivity, and response bias analyses in Experiment 1b. The BF_{10} for each model is in comparison to the null model.

Table 3. Table displaying Bayes Factors (BF_{10}) for each model across accuracy and response time analyses in Experiment 1a. The BF_{10} for each model is in comparison to the null model.

Table 4. Table displaying variability of parameters across sequencing condition and feature dimension, log-likelihood, and AIC for each model in Experiment 1a. A $\oplus$ symbol indicates that the parameter was free to vary across sequencing condition, while a $\ominus$ symbol indicates that the parameter was not free to vary across feature dimension.

Table 5. Table displaying Bayes Factors (BF_{10}) for each model across accuracy and response time analyses in Experiment 1b. The BF_{10} for each model is in comparison to the null model.

Table 6. Table displaying variability of parameters across sequencing condition and feature dimension, log-likelihood, and AIC for each model in Experiment 1b. A $\oplus$ symbol indicates that the parameter was free to vary across sequencing condition, while a $\ominus$ symbol indicates that the parameter was not free to vary across feature dimension.