Investigating the role of spatial and non-spatial attention in visual short-term memory: A comparison of recall for colour and orientation features

Student ID: 229026073

Supervisor: Dr Doug Barrett

**Abstract**

Visual short-term memory (VSTM) is a capacity-limited system responsible for the temporary storage and manipulation of visual information. This study investigated how spatial and non-spatial attention independently influence VSTM performance, testing predictions from both slot-based and resource-based models. Participants completed a change detection task in which attention was directed to one, two, or four spatial locations via pre-stimulus cues. After a retention interval, a post-stimulus, feature-based cue (colour, orientation) indicated the relevant feature dimension to retrieve. Accuracy and reaction times were measured across a 3 (spatial cue load) x 3 (feature cue type) within-subjects design. Results revealed significant main effects of spatial load and feature cue type on both accuracy and reaction time. Accuracy declined as spatial cue load increased and was lower when orientation changes were required compared to colour. However, no significant interactions emerged, suggesting that spatial and non-spatial cues influenced VSTM performance independently. Reaction time data indicated increased retrieval difficulty with greater spatial load, and a potential speed-accuracy trade-off in trials requiring cross-dimensional attention. These findings support a resource-based account of VSTM in which spatial and non-spatial attentional resources operate in parallel but remain independently limited. The results challenge fixed-capacity slot models and highlight the role of attentional guidance in modulating both encoding and retrieval in visual memory.

**Introduction**

Visual short-term memory (VSTM) enables the temporary retention and manipulation of visual information, supporting perceptual continuity, decision-making, and interaction with dynamic environments (Postle, 2014). It facilitates core cognitive functions such as tracking moving objects, face recognition, and the integration of visual details across time. Despite its central role in cognition, the mechanisms underlying VSTM remain incompletely understood (Hollingworth et al., 2008). A central debate concerns the nature of its capacity limitations. Slot-based models posit a fixed number of discrete storage units, each capable of representing one object with high precision (Cowan, 2001). In contrast, resource-based models argue for a flexible system where memory resources are shared across items, resulting in reduced fidelity as item load increases (Bays, Catalao, & Husain, 2011). However, both perspectives have limitations. Slot models struggle to explain graded precision, while resource models often overlook strategic allocation and prioritization mechanisms. Recent work suggests that attentional processes may modulate how resources are allocated in VSTM, in particular, the interaction between spatial and non-spatial location, this may potentially reconcile aspects of both models.

Deficits in VSTM are prominent in neuropsychological disorders such as schizophrenia, Alzheimer’s disease, and Parkinson’s disease, with affected individuals often exhibiting up to a fifty percent reduction in capacity relative to healthy controls (Lee et al., 2010). Yet, the precise nature of these impairments remains ambiguous, partly due to inconsistent methodological approaches and the complexity of underlying attentional and memory interactions. Beyond clinical populations, VSTM also declines with healthy aging, with impairments in recall precision and feature binding being particularly pronounced (Matsuyoshi, Osaka, & Osaka, 2014; Peich, Husain, & Bays, 2013). As the global population continues to age, understanding VSTM deterioration has become increasingly important for informing cognitive health interventions and societal planning.

A core function of VSTM is its ability to retain visual details briefly, allowing for perceptual continuity. A long-standing and central question concern its capacity limitations; specifically, how much and what kind of information can be maintained at once. This issue has major implications, not only for everyday functioning but also for more complex cognitive tasks like reading comprehension and decision-making (Daneman & Carpenter, 1980). Despite extensive investigation, the mechanisms constraining VSTM capacity remain debated, with focus now turning to how different forms of attention may interact to influence storage precision and item retention. This shift highlights the need to integrate attentional dynamics into current capacity models.

A central question in VSTM research is whether memory operates through discrete storage “slots” or by flexibly distributing a shared resource. According to slot models, VSTM maintains information in fixed-capacity slots, with each slot representing a bound object composed of multiple features, such as colour and orientation (Hojjati et al., 2024). Research suggests that VSTM can store a maximum of four items, beyond which recall performance declines sharply (Cowan, 2001). Slot models propose that if an item falls within this capacity limit, it is recalled with perfect precision, with no loss in visual detail. However, if the number of items exceeds capacity, some objects are not stored at all, resulting in complete memory loss for those items (Donkin et al., 2013). Supporting evidence indicates that resources can be distributed among stored items, but only up to a fixed limit, after which performance deteriorates significantly (Luck & Vogel, 2013). This all-or-nothing principle remains central to slot-based theories, yet its validity continues to be debated, as competing models suggest that memory precision may decline more gradually rather than abruptly failing. Understanding these capacity limitations is crucial for refining theoretical models of VSTM and its role in cognitive processes.

In contrast to slot models, resource models conceptualize VSTM as a flexible system governed by a shared pool of cognitive resources rather than discrete storage slots. According to resource models, visual items compete for a common, divisible memory resource that is dynamically allocated across stored stimuli (Bays & Husain, 2008). This framework allows resources to be concentrated on fewer items to preserve high-precision representations or distributed more broadly, resulting in reduced fidelity (Ye et al., 2017). Unlike the fixed item limits in slot models, resource theories suggest that performance limitations emerge from variability in memory precision rather than item count. Bays, Catalao, and Husain (2011) argue that as more items are stored, fewer resources are available per item, leading to degraded precision. However, a critical unresolved issue is whether such limitations reflect internal constraints within the attentional system or are instead influenced by external factors like encoding time or perceptual load (Cappiello, 2019). These theoretical contrasts raise important questions about the nature of stored visual features, whether they are maintained as integrated objects or as separable, feature-specific representations, they also highlight the need for paradigms capable of isolating attentional contributions to VSTM performance.

The extent to which attention can mitigate VSTM capacity limitations remains a central question in cognitive neuroscience. As a gatekeeper of memory, attention determines which visual information is encoded, maintained, and retrieved, thereby influencing the fidelity and quantity of representations in VSTM. Spatial attention prioritizes specific locations in the visual field, facilitating the encoding of objects at attended positions while suppressing distractors (Matsakura et al., 2007). Conversely, non-spatial (feature-based) attention enhances the processing of object-specific attributes such as colour or orientation, regardless of their spatial location. Neuroimaging evidence suggests that these two forms of attention are supported by partially distinct cortical networks, reflecting specialized mechanisms for modulating memory based on spatial versus feature-based priorities (Heuer et al., 2016). Yet, how these attentional subsystems interact to shape VSTM performance remains poorly understood. Critically, it is unclear whether spatial and non-spatial attention operate independently or synergistically to influence how memory resources or slots are allocated. Clarifying this interaction may offer a deeper understanding of whether attentional guidance aligns more closely with a resource-based model or a slot-based model. Addressing these questions is therefore essential for elucidating the architecture of VSTM and the mechanisms that constrain its capacity.

The change detection task is a foundational paradigm for probing mechanisms and capacity constraints of VSTM (Luck & Vogel, 1997; Feldmann-Wustefeld, 2020). In this task, participants view an initial array of visual stimuli, followed by a brief retention interval, typically around one second, after which a second array (the probe display) appears. This display may be identical or contain subtle feature changes, such as colour or orientation, requiring participants to judge whether a change occurred. Performance on this task reflects the efficiency of encoding, maintenance, and retrieval processes in VSTM. Crucially, by manipulating set sizes, researchers can estimate memory capacity and assess whether performance aligns with slot-based or resource-based models. The paradigm is also well-suited for investigating attentional influences, directing spatial or non-spatial attention during encoding can reveal how attentional mechanisms modulate memory fidelity. Importantly, VSTM serves as a temporal buffer that integrates perceptual input over time, yet its limited capacity can give rise to change blindness (Sapotka et al., 2015). This phenomenon underscores the interdependence between attentional allocation and memory capacity, highlighting the importance of experimental paradigms that can disentangle these components. As such, the change detection task offers a powerful tool for investigating how spatial and non-spatial attention interact with the structural limits of VSTM.

Reaction time (RT) is frequently used alongside accuracy as a complementary measure in the investigation of VSTM, offering additional insight into the temporal and cognitive dynamics underlying attentional deployment, encoding, and retrieval processes (Awh et al., 2006; Bays & Husain, 2008). While accuracy typically reflects the precision or capacity of memory representations, RT provides a more sensitive index of processing efficiency and cognitive effort. Importantly, slot-based models predict relatively stable RTs up to capacity limits, after which performance may abruptly decline, whereas resource-based models suggest a more continuous increase in RT as memory load and cognitive demands rise (Griffin & Nobre, 2003). RT measures have therefore been instrumental in distinguishing between these theoretical accounts. However, prior studies have often examined RT without fully disentangling the contributions of spatial and non-spatial attention. By incorporating RT in the present study, which manipulates both attentional domains, it becomes possible to uncover nuanced attentional effects on VSTM that may not be evident through accuracy alone. This multimodal approach may yield deeper insights into how attentional mechanisms interact and allocate cognitive resources over time, addressing limitations in previous research and contributing to a more comprehensive account of VSTM function.

Awh and Jonides (2001) provide compelling evidence that spatial attention enhances performance in VSTM tasks by selectively prioritizing items during encoding. Their findings revealed improved memory accuracy when participants were cued to attend to specific spatial locations, supporting the notion that attention functions as a gatekeeper. Determining which items are encoded and potentially aligning with slot-based models where attention governs access to a limited number of discrete storage slots. However, the study’s narrow focus on spatial attention limits its explanatory scope; by neglecting non-spatial attention, their findings cannot fully account for how attentional allocation modulates memory precision. This oversight is particularly problematic considering resource-based models, influencing the fidelity rather than just the presence of memory representations. Addressing this gap, the current study manipulates both spatial and non-spatial attention within the same paradigm to explore how these attentional mechanisms interact. This integrative approach allows for a more nuanced test of competing theoretical models and clarifies how attentional systems jointly shape the quality and capacity of VSTM.

Heuer and Schubo (2016) investigated the combined effects of spatial and non-spatial attention on VSTM, testing whether attention directed to locations or features enhances memory accuracy and whether these attentional systems operate independently. They found that both spatial and feature-based cues significantly improved performance, with additive effects suggesting partially distinct underlying mechanisms. These findings support resource-based accounts of VSTM, which argue that attention allocates flexible cognitive resources across spatial and non-spatial dimensions to enhance memory precision. However, a notable limitation of their study is the short retention interval, which restricts conclusions about the role of sustained attention in maintaining memory over time. While their results clarify how attentional selection influences encoding, they do not address whether spatial and feature-based attention continue to support representations during maintenance. This gap is critical for understanding how attentional mechanisms shape VSTM beyond initial encoding, particularly under resource-limited conditions. The present study addresses the limitation by introducing longer retention intervals, enabling a more nuanced analysis of how sustained spatial and non-spatial attention modulate the fidelity of memory representations, advancing both slot and resource-based theoretical perspectives.

Research investigating attention during the maintenance phase of VSTM has revealed that spatial attention can dynamically modulate stored representations, enabling flexible prioritization of task-relevant information. Awh et al. (2006) demonstrated that retro-cues (spatial cues presented after encoding) significantly enhanced memory performance, with improved accuracy and faster reaction times when attention was directed toward specific item locations. Critically, even retro-cues presented later in the retention interval conferred benefits, suggesting that attentional selection can be recruited post-encoding to reallocate cognitive resources toward relevant representations. These findings align with resource-based models, which argue that attention enhances memory precision not only at encoding but also during maintenance by flexibly distributing limited cognitive resources. However, a major limitation of this and related studies is their exclusive focus on spatial attention. Neglecting non-spatial attention fails to capture the full dimensionality of attentional allocation in VSTM, particularly given evidence that feature-based cues also modulate memory fidelity. This theoretical gap limits the explanatory power of existing models and hinders a more integrative understanding of VSTM maintenance. To address this, the present study introduces novel conditions that manipulate both spatial and non-spatial attention during extended retention intervals, enabling a more comprehensive investigation into how these mechanisms interact to shape memory over time.

In the current study, attention is operationalized through two types of cueing; spatial attention is manipulated by indicating the relevant location(s) in the visual array, whereas non-spatial attention is manipulated through cues that specify a particular feature dimension (colour or orientation). Single-feature cueing refers to conditions where participants are cued to attend to either colour or orientation, while dual-feature cueing requires participants to attend to both feature dimensions simultaneously. Memory load is varied by changing the number of items in the memory array, with low set size representing minimal load and high set size reflecting increased cognitive demands. These manipulations allow for systematic examination of how attentional deployment and memory demands jointly influence VSTM performance. The change detection paradigm employed in this study is especially well-suited to disentangle the effects of spatial and non-spatial attention on VSTM. By incorporating both feature-based and location-based cues, the paradigm enables independent and combined manipulation of attentional dimensions. Furthermore, by systematically varying set size and cueing conditions, this design allows for direct testing of competing models of VSTM capacity. Reaction time and accuracy serve as complementary indices of performance, capturing not only how well participants store visual information but also the processing costs associated with attentional allocation. Together, these features make the task an ideal tool for investigating how attentional mechanisms modulate the distribution of memory resources.

This study addresses several unresolved issues in literature first, although prior work has established the individual roles of spatial and non-spatial attention in VSTM, few studies have systematically investigated their combined influence within a unified paradigm. Second, much of the existing retro-cue literature focuses on short-retention intervals and exclusively on spatial guidance, leaving unclear whether non-spatial cues can similarly modulate stored representations over time. Third, theoretical models remain divided on whether attention serves as a gateway for item selection or as a mechanism for dynamically allocating limited resources. By integrating dual-attention cueing, varying memory load, and extending retention intervals, the present study builds on previous work while overcoming methodological limitations, offering a more comprehensive view of how attention and memory interact in visual cognition.

Based on existing models of VSTM, three core hypotheses were formulated. First, it was hypothesized that memory performance would decline as set size increased, evidenced by reduced accuracy and prolonged reaction times at higher memory loads. This pattern would be consistent with both discrete slot models and resource-based models. Second, it was predicted that cueing participants to attend to both colour and orientation would result in poorer performance compared to single-feature cueing. This reflects the cognitive cost associated with the combined resource model, it was hypothesized that a significant interaction would emerge between spatial and feature-based cueing. Specifically, performance was expected to be lowest when participants were required to attend to both spatial locations and multiple features, indicating a compounding cost when cognitive resources are distributed across both spatial and non-spatial domains. This study aims to investigate how spatial and non-spatial attention interact to influence VSTM performance under varying memory loads, using a change detection paradigm with both accuracy and reaction time as outcome measures.

**Method**

Participants

Fourteen individuals (nine male, five female; M age = *20.93* years, SD = *0.27*) were recruited from the University of Leicester to take part in the study. All participants reported normal or corrected-to-normal vision and no history of neurological or visual impairments. Informed consent was obtained from all participants in accordance with the University of Leicester’s ethical guidelines.

Apparatus

The experiment was programmed and executed using MATLAB with the Psychophysics Toolbox extension and presented on a 24-inch LCD monitor. Participants were seated approximately 60 cm from the screen with their head stabilized using a chinrest to maintain a consistent viewing distance and minimize head movement. Responses were recorded using a standard keyboard.

Stimuli

The visual stimuli comprised arrays of four elliptical shapes presented on a black background. Each ellipse had a fixed aspect ratio of 1:1:8 (minor:major:axis) and was enclosed within a 150 x 150-pixel square. Stimuli were positioned at the four cardinal points (top, bottom, left, right) around a central fixation cross, with an eccentricity of 230 pixels from the screen centre, forming a consistent circular configuration across trials. Each ellipse varied along two feature dimensions: colour and orientation. Colours were selected from a 255-point virtual circle in RGB colour space, ensuring perceptual distinctiveness and balanced sampling across trials. Additionally, colours were sampled every 30 degrees along a 360-degree hue wheel. Orientation was defined by angular rotation, with values sampled from a 0 to 180-degree range in 20-degree increments, providing sufficient variability while maintaining perceptual discriminability. This stimulus configuration allowed for controlled variation in both spatial location and visual feature content, enabling precise manipulation of spatial and non-spatial attention. The design was well-suited for testing competing models of VSTM by systematically altering attentional demands across feature and spatial dimensions.

Design

The experiment employed a 3 (spatial cue load: 1, 2, or 4 locations) x 3 (feature cue type: Colour, Orientation, or both) within-subjects factorial design. The spatial cue load determined the number of spatial locations participants were cued to attend to prior to stimulus presentation. Feature cue type was manipulated via a post-retention probe cue, which instructed participants to respond based on the colour or orientation of the probed item. This design allowed for the investigation of how attentional demands across spatial and non-spatial domains independently and interactively affected VSTM performance. All participants completed trials in each of the conditions in a counterbalanced order to minimize order and fatigue effects. Accuracy and reaction time were recorded as dependent variables; however, the primary analyses focused on accuracy as the key indicator of VSTM performance.

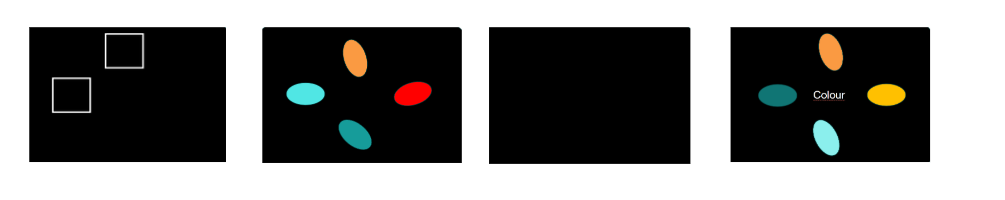


Figure 1. Example of a simulation of a singular trial. The first screen shows the spatial cue for participants to attend to; the second screen displays the memory array followed by a blank screen requiring participants to hold all information in their VSTM. The final screen is the test array; it has a non-spatial cue instructing participants which feature to attend to.

Procedure

Each trial began with a central fixation cross displayed for 250 ms, followed by a spatial cue that highlighted either one, two, or four locations for 400 ms. This cue indicated the positions participants were required to attend to in the upcoming memory array. A second fixation screen was then presented for 250 ms, reorienting participants’ attention to the centre. Next, the memory array appeared for a duration that was scaled with the set size; i.e., 500 ms, 1000 ms, or 2000 ms, for one, two, or four items respectively. This array consisted of four oval stimuli, each varying in colour and orientation, arranged at fixed, equidistant positions around the central fixation. Following the memory array, a 1000 ms retention interval was presented, during which a feature cue appeared at the centre of the screen (either the word “colour” or “orientation”). This cue instructed participants which feature dimension to evaluate in the upcoming test. A test array then appeared in the same spatial configuration as the memory array and remained onscreen until the participant responded or for a maximum of 3000 ms. Participants were instructed to indicate whether any of the items at the previously cued spatial locations had changed in the cued feature (colour or orientation) using a keyboard response. The experiment consisted of eight blocks, each containing sixty trials. Trials were counterbalanced across spatial cue load and feature cue conditions, with short breaks provided between blocks to reduce fatigue.

Data analyses

All statistical analyses were conducted using JASP software. A 3 (spatial cue load: 1, 2, 4 locations) x 3 (feature cue type: colour, orientation, both) repeated-measures ANOVA was performed on accuracy and reaction time scores to examine the independent and combined effects of spatial and non-spatial attentional cueing on VSTM performance. Mauchly’s test was used to evaluate the assumption of sphericity; where violations occurred, Greenhouse-Geisser corrections were applied. Statistical significance was determined at an alpha level of α = .05, and Partial eta squared (η²ₚ) was reported as a measure of effect size. Descriptive statistics (means and standard deviations) were also computed to support interpretation of main effects and interactions.

**Results**

Accuracy

Visual inspection of the descriptive plots confirmed that accuracy decreased systematically as spatial cue load increased across all conditions. Participants performed best when one location was cued and worst when four locations were cued. Across both feature types, accuracy was consistently higher when participants were cued to attend to colour compared to orientation, although this advantage was more pronounced in the colour change condition. The decline in performance with increased load was parallel across cue types, with no evidence of an interaction, consistent with the ANOVA results showing independent effects of spatial cue load and feature-based cueing on VSTM accuracy.

A 3 x 3 repeated-measures ANOVA was conducted on accuracy data to assess the effects of spatial and feature-based attention on VSTM performance in a change detection task. The analysis revealed a significant main effect of Cue Type, *F* (1*, 13*) = *5.44*, *p* = *.036, η²ₚ = .30*, indicating greater accuracy when participants were cued to attend to colour compared to orientation. A significant main effect of Change type was also found, *F* (*1, 13*) = *8.74, p = .011*, *η²ₚ = .40*, with higher accuracy for detecting colour changes than orientation changes.

A significant main effect of Spatial Cue Load was observed, *F* (*2, 26*) = *40.29*, *p* < *.001, η²ₚ* = *.76*, indicating that accuracy decreased as the number of cued spatial locations increased. Participants were most accurate in the 1-location condition and least accurate in the 4-location condition. All two-way and three-way interactions were non-significant: Cue x Change, *F* (*1, 13*) = *1.12*, *p = .310,* η²ₚ = *.08*; Cue x Spatial Cue Load, *F* (*2, 26*) = *0.82*, *p* = *.451, η²ₚ* = *.06*; Change x Spatial Cue Load, *F* (*2, 26*) = *1.96*, *p* = *.161, η²ₚ = .13* and Cue x Change x Spatial Cue Load, *F* (*2, 26*) = *0.24*, *p* = *.792, η²ₚ = .02*. These results indicate that the effects of cue type, change type, and spatial cue load on accuracy were independent, with no evidence of interaction effects.

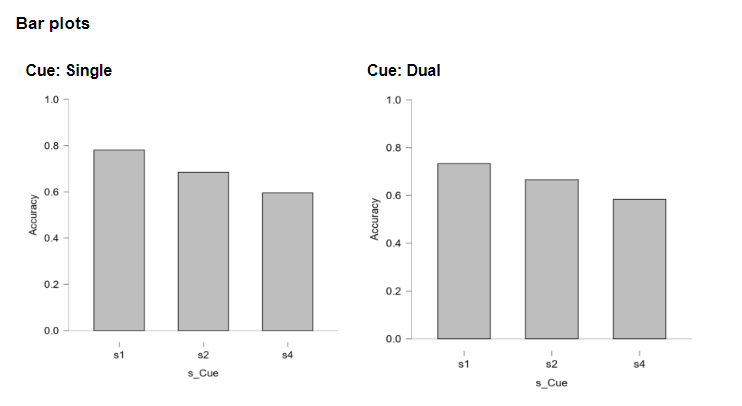


Figure 2. Accuracy results for single and dual feature conditions

Reaction Time

The descriptive plots upon inspection indicated that RTs were generally slower under dual-feature cue conditions across all spatial loads. In the colour feature condition, mean RTs ranged from approximately *0.90* seconds at set size 1 to around *1.00* seconds at set size 4 for dual cues, while single-cue trials showed faster responses, ranging from roughly *0.75* to *0.85* seconds across the same loads. A similar pattern was observed for the orientation feature, with dual-cue RTs remaining relatively stable around *0.95* to *1.00* seconds, and single-cue trials showing quicker responses between approximately *0.75* and *0.80* seconds. Across both feature types, participants responded more rapidly when attention was directed to a single feature rather than to both features simultaneously. These trends align with the ANOVA results, indicating that feature cue type influenced retrieval efficiency independently of spatial cue load.

A graph of a bar graph

AI-generated content may be incorrect.

Figure 3. Reaction Time results under increasing cognitive loads.

A 3 x 3 repeated measures ANOVA was conducted to examine the effects of spatial and feature-based attention on reaction times in the change detection task. There was a significant main effect of Cue, *F (1, 13*) = *24.17*, *p* < *.001, η²ₚ = .65*, with participants responding faster on cued trials compared to uncued trials. A significant main effect of Spatial Cue Load was also observed, *F (2, 26*) = *7.82*, *p* = *.002*, *η²ₚ = .38*, with reaction times increasing as the number of cued locations increased. The main effect of Change type was not significant, *F (1, 13*) = *0.14*, *p* = *.714, η²ₚ = .01*, indicating that the presence or absence of a change did not significantly influence reaction times. All two-way and three-way interactions were non-significant: Cue × Change, *F*(*1, 13*) = *0.46*, *p* = *.508, η²ₚ = .03*; Cue × Spatial Cue Load, *F*(*2, 26*) = *1.53*, *p* = *.236, η²ₚ = .11*; Change × Spatial Cue Load, *F*(*2, 26*) = *1.72, p = .199, η²ₚ = .12*; and Cue × Change × Spatial Cue Load, *F*(*2, 26*) = *1.28, p = .296, η²ₚ = .09*. Mauchly’s test indicated a violation of the sphericity assumption for the Spatial Cue Load factor, therefore, Greenhouse-Geisser corrections were applied where necessary. Overall, the pattern of results underscores the influence of spatial attention on processing efficiency in VSTM, with increased spatial demands leading to slower and less efficient responses.

**Discussion**

This study aimed to examine how spatial and non-spatial attentional cueing influence VSTM performance under varying memory loads. The findings demonstrate that both types of attentional guidance contributed to memory performance, but with distinct cognitive costs when attentional resources were distributed across multiple domains. Accuracy declined systematically as set size increased across all cueing conditions, consistent with resource-based models of VSTM. For example, in the attend-colour condition, accuracy decreased from M = *0.797* (s*1*) to M = *0.638* (s*4*), while in the attend orientation condition, it dropped from M = *0.736* (s*1*) to M = *0.551* (s*4*). Notably, dual-feature conditions exhibited even greater reductions in performance at higher loads (e.g., M = *0.765* attend both, orientation s*1*; M = *0.533* s*4*), suggesting increased cognitive demands when maintaining multiple feature dimensions simultaneously. Reaction time data further supported this interpretation. Across conditions, RTs increased as spatial load rose, indicating greater retrieval difficulty under higher memory demands (e.g., attend-colour condition: M = *0.660* (s*1*), M = 0.*866* (s*4*); attend-orientation condition: M = *0.914* (s*1*) to M = *0.979* (s*4*)). Crucially, participants exhibited slower RTs when attending to both features compared to a single feature, consistent with greater cognitive load and more complex retrieval processes in dual-feature conditions. These patterns suggest that the distribution of attentional resources across both spatial and feature domains imposes measurable costs on both memory precision and retrieval efficiency. Taken together, the accuracy and reaction time findings challenge traditional fixed-slot models of VSTM, the gradual decline in performance observed here is more consistent with resource-based models.

The pattern of results observed in this study show similarities with a study by Bays, Catalao, & Husain (2009) who challenged the notion of a fixed-capacity slot model by demonstrating a continuous decline in recall precision with increasing memory load. Rather than supporting a hard upper limit on the number of items that can be stored, the present data reveal a gradual decrease in accuracy as set sizes increase, most notably from M = *0.797* for attend colour, s1 to M = *0.553* for attend both condition, s4. This pattern underscores well-documented capacity limitations in VSTM and points toward models that view memory resources as flexibly allocated rather than rigidly compartmentalized. From a cognitive neuroscience standpoint, these findings contribute to an evolving understanding of how attentional mechanisms interact with memory maintenance processes. Specifically, they support hybrid and resource-based accounts, thereby influencing the fidelity of stored representations. The benefit of spatial cueing, particularly at lower memory loads, suggests that top-down attention enhances the resolution of memory representations by prioritizing relevant spatial locations during encoding. This is a notion consistent with prior work demonstrating improved memory performance following pre-encoding spatial attention (Griffin & Nobre, 2003; Awh et al., 2006). Together, these insights provide a strong foundation for interpreting the current results through the lens of resource-based models.

VSTM performance was compared across single and dual-feature conditions to explore how dividing non-spatial attention impacts memory accuracy and retrieval efficiency. Interestingly, the descriptive data revealed that accuracy scores for single and dual-feature conditions were highly comparable across all set sizes. This suggests that dividing attention between colour and orientation dimensions did not impose a substantial cost to VSTM performance. Rather than causing a dramatic decline in accuracy when cue and change features differed, participants were able to flexibly allocate attention across feature domains with only modest reductions in precision. These findings challenge strict slot-based models while aligning more closely with resource-based accounts because the observed gradual degradation in accuracy implies a continuous allocation of memory fidelity across multiple representational demands (Ye et al., 2017). Rather than an all-or-none encoding of features, VSTM resources appear to be allocated in proportion to task demands, such that precision is preserved at the expense of slight reductions in representational quality when attention is divided. This interpretation reflects key principles of resource models, in which representational fidelity is the limiting factor. Furthermore, the findings are consistent with Fan & Turk-Browne's (2013) proposal that retrieval processes reinforce attended feature dimensions while suppressing unattended ones, suggesting that attentional prioritization can dynamically shape representational strength even when resources are taxed across multiple domains.

The reaction time data provides additional support for the view that attentional control and memory precision are governed by resource-limited, rather than discretely bounded, mechanisms. RTs increased approximately linearly with increasing set size across all cueing conditions, even when spatial or non-spatial cues were available. This graded pattern contradicts slot-based models, which predict a plateau in response latency once mnemonic capacity is saturated. Instead, the data suggests that participants continued to engage in effortful retrieval even beyond conventional capacity limits, consistent with resource-based frameworks. Notably, participants responded more quickly when a single feature was cued, particularly at lower set sizes. This pattern implies that reducing the scope of feature-based attentional demands facilitates faster prioritization and comparison processes, likely due to reduced competition among stored representations. Slower RTs under dual-feature conditions may reflect interference arising from overlapping or competing representational codes within a shared attentional resource. This interpretation is supported by neuroimaging evidence showing that dual-feature maintenance engages overlapping frontoparietal networks, increasing connectivity demands and cognitive load (Gunseli et al., 2019). Furthermore, the inability of either cue type to eliminate load-related slowing suggests that attentional modulation operates within the broader constraints of a shared memory-attention system (D’Esposito & Postle, 2015). Together, these findings position attention not as an independent enhancer of performance, but as a mechanism for dynamically optimizing limited mnemonic resources under varying task demands.

The significant main effect of cueing (*p* = *.036*, *η²*ₚ = *.295*) underscores the benefit of attentional guidance on memory accuracy, suggesting that cue-driven prioritization enhances the precision of encoded representations. This finding aligns closely with resource-based models, which propose that mnemonic resources can be flexibly allocated toward task-relevant information (Bays & Husain, 2008; Ma, Husain, & Bays, 2014). The significant main effect of change presence (*p* = .011, η²ₚ = .402) further highlights the dependency of successful retrieval on the strength and fidelity of stored representations, reinforcing the notion that higher-resolution memory traces are necessary for detecting subtle changes, as also demonstrated in earlier change detection studies (Zhang & Luck, 2008). Importantly, the large main effect of spatial cue load (*p* < *.001*, *η²ₚ* = *.756*) indicates that the nature of attentional effects on VSTM are not only not only quantitative but qualitatively differentiated. It indicates that attentional modality affects representational precision, in-line with graded-resource theories.

The absence of significant interaction effects suggests that spatial and non-spatial attention mechanisms operate additively rather than interactively, potentially reflecting dissociable but parallel systems for enhancing VSTM performance (Heuer & Schubo, 2016). This layered attentional architecture could correspond to distinct temporal windows of influence, with spatial cues predominantly affecting encoding, and non-spatial cues exerting greater influence at retrieval (Griffin & Nobre, 2003). Together, these findings provide strong empirical support for resource-based interpretations of VSTM and highlight the need for future research using temporally sensitive neurophysiological methods, such as EEG to more precisely track how distinct attentional mechanisms dynamically shape memory precision across different task phases.

While the current study offers meaningful insights into how spatial and non-spatial attention modulates VSTM, several considerations highlight opportunities for future refinement. The consistent presence of spatial cues was a deliberate design choice to establish a stable reference for attentional allocation across trials; however, this meant that a true no-cue baseline was unavailable. Including an uncued condition in future studies could clarify whether spatial cues actively enhance encoding or simply preserve baseline performance (Souza & Oberauer, 2016). Additionally, non-spatial cues were presented only at the point of retrieval rather than during encoding or maintenance, reflecting an interest in their retrospective influence. However, prior research suggests that cue timing is critical, with earlier feature-based cues potentially offering stronger modulation of memory representations (Griffin & Nobre, 2003). Future work should therefore manipulate cue timing to directly assess whether proactive attentional orienting toward features during encoding improves VSTM precision more effectively than retrospective selection. Finally, while the modest sample size (N = 14) is consistent with within-subjects designs in cognitive neuroscience, it naturally limits generalizability. Although the observed effects were statistically robust, replication with larger samples would improve confidence in the findings and reduce susceptibility to individual variability (Button et al., 2013). Addressing these design considerations will allow future research to more precisely map the temporal dynamics of attentional selection and its role in resource allocation within VSTM.

Building on current findings, future research should examine how individual differences in attentional control shape VSTM performance. Although the present study focused on group-level effects, substantial variability likely exists in individuals’ ability to allocate attention under high memory load or when maintaining multiple features (Fukuda & Vogel, 2009). Traits such as working memory capacity, cognitive flexibility, and attentional breadth may modulate the effectiveness of both spatial and feature-based cueing strategies (Unsworth & Engle, 2007). Incorporating neurocognitive assessments and eye-tracking measures could help reveal whether individuals preferentially engage spatial versus feature-based selection mechanisms, offering more fine-grained insights into internal attention dynamics (Myers et al., 2017). Moreover, adaptive cueing paradigms tailored to individual attentional profiles could optimize VSTM performance, providing a potential avenue for personalized cognitive interventions. Pursuing these directions would not only refine resource-based models of VSTM but also clarify how attention-memory interactions vary across individuals, contributing to a more nuanced, mechanistic understanding of cognitive architecture.

This study advances understanding of how spatial and non-spatial attention interact to shape VSTM performance. By systematically manipulating cue types and memory load, the results show that spatial attention markedly enhances memory accuracy and retrieval speed, while non-spatial cues offer additional, though smaller benefits. These findings align with resource-based models, highlighting the flexible allocation of cognitive resources based on attentional demands. Critically, the study demonstrates that spatial and non-spatial cues jointly optimize memory efficiency even when deployed at different task stages, refining current models of attentional control in VSTM. Future research should investigate how individual differences in attentional capacity further modulate these mechanisms.

**References**

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>

Awh, E., Vogel, E., & Oh, S. (2005). Interactions between attention and working memory. *Neuroscience, 139*(1), 201-208. <https://doi.org/10.1016/j.neuroscience.2005.08.023>

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. <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>

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafo, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience, *Nature Reviews Neuroscience, 14*(5), 365-376. <https://doi.org/10.1038/nrn3475>

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences,* *24*(1), 87-114. <https://doi.org/10.1017/S0140525X01003922>

Cappiello, M. (2019). Reevaluation of formal model comparison between slot and resource models of visual working memory. *eScholarship*. <https://escholarship.org/uc/item/4xg0n4zt>

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*(4), 450-466. <https://doi.org/10.1016/s0022-5371(80)90312-6>

D’Esposito, M., & Postle, B. R. (2015). The cognitive neuroscience of working memory. *Annual Review of Psychology, 66*(1), 115-142. <https://doi.org/10.1146/annurev-psych-010814-015031>

Donkin, C., Nosofsky, R. M., Gold, J. M., & Shiffrin, R. M. (2013). Discrete-slots models of visual working-memory response times. *Psychological Review, 120*(4), 873-902. <https://doi.org/10.1037/a0034247>

Fan, J. E., & Turk-Browne, N. B. (2013). Internal attention to features in visual short-term memory guides object learning. *Cognition, 129*(2), 292-308. <https://doi.org/10.1016/j.cognition.2013.06.007>

Feldmann-Wustefeld, T. (2020). Neural measures of working memory in a bilateral change detection task. *Psychophysiology, 58*(1). <https://doi.org/10.1111/psyp.13683>

Fukuda, K., & Vogel, E. K. (2009). Human variation in overriding attentional capture. *Journal of Neuroscience, 29*(27), 8726-8733. <https://doi.org/10.1523/JNEUROSCI.2145-09.2009>

Griffin, I. C., & Nobre, A. C. (2003). Orienting attention to locations in internal representations. *Journal of Cognitive Neuroscience 15*(8), 1176-1194. <https://doi.org/10.1162/089892903322598139>

Gunseli, E., Fahrenfort, J. J., Van Moorselar, D., Daoultzis, K. C., Meeter, M., & Olivers, C. N. L. (2019). EEG dynamics reveal a dissociation between storage and selective attention within working memory. *Scientific Reports, 9*(1), Article 13486. <https://doi.org/10.1038/s41598-019-49577-0>

Heuer, A, A., Schubo, A., & Crawford, J. D. (2016). Different cortical mechanisms for spatial vs. Feature-based attentional selection in visual working memory. *Frontiers* *in Human Neuroscience, 10*. <https://doi.org/10.3389/fnhum.2016.00415>

Hojjati, F., Motahharynia, A., Adibi, A., Adibi, I., & Sanayei, M. (2024). Resource or slot model in visual working memory: Are they different? *BioRxiv.* <https://doi.org/10.1101/2024.02.08.579494>

Hollingworth, A., Richard, A. M., & Luck, S. J. (2008). Understanding the function of visual short-term memory: Trans saccadic 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>

Landman, R., Spekreijse, H., & Lamme, V. A. (2003). Large capacity storage of integrated objects before change blindness. *Vision Research, 43*(2), 149-164. <https://doi.org/10.1016/s0042-6989(02)00402-9>

Lee, E., Cowan, N., Vogel, E. K., Rolan, T., Valle-Inclan, F., & Hackley, S. A. (2010). Visual working memory deficits in patients with Parkinson’s disease are due to both reduced storage capacity and impaired ability to filter out irrelevant information. *Brain, 133*(9), 2677-2689. <https://doi.org/10.1093/brain/awq197>

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>

Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: From psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences, 17*(8), 391-400. <https://doi.org/10.1016/j.tics.2013.06.006>

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>

Makovski, T., Sussman, R., & Jiang, Y. V. (2008). Orienting attention in visual working memory reduces interference from memory probes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(2), 369-380. <https://doi.org/10.1037/0278-7393.34.2.369>

Matsukura, M., Luck, S. J., & Vecera, S. P. (2007) Attention effects during visual short-term memory maintenance: Protection or prioritization? *Perception & Psychophysics, 69*(8), 1422-1434. <https://doi.org/10.3758/bf03192957>

Matsuyoshi, D., Osaka, M., & Osaka, N. (2014). Age and individual differences in visual working memory deficit induced by overload. *Frontiers in Psychology, 5*, Article 384. <https://doi.org/10.3389/fpsyg.2014.00384>

Myers, N. E., Stokes, M. G., & Nobre, A. C. (2017). Prioritizing information during working memory: Beyond sustained internal attention. *Trends in Cognitive Sciences, 21*(6), 449-461. <https://doi.org/10.1016/j.tics.2017.03.010>

Nobre, A. C., Griffin, I. C., & Rao, A. (2008). Spatial attention can bias search in visual short-term memory. *Frontiers in Human Neuroscience, 1.* <https://doi.org/10.3389/neuro.09.004.2007>

Peich, M., Husain, M., & Bays, P.M. (2013). Age-related decline of precision and binding in visual working memory. *Psychology and Aging,* *28*(3), 729-743. <https://doi.org/10.1037/a0033236>

Postle, B. R. (2014). The cognitive neuroscience of visual short-term memory. *Current Opinion in Behavioral Sciences, 1*, 40-46. <https://doi.org/10.1016/j.cobeha.2014.08.004>

Sapkota, R. P., Pardhan, S., & Van Der Linde, I. (2015). Change detection in visual short-term memory. *Experimental Psychology, 62*(4), 232-239. <https://doi.org/10.1027/1618-3169/a000294>

Shomstein, S., & Gottlieb, J. (2016). Spatial and non-spatial aspects of visual attention: Interactive cognitive mechanisms and neural underpinnings. *Neuropsychologia, 92,* 9-19. <https://doi.org/10.1016/j.neuropsychologia.2016.05.021>

Souza, A. S., & Oberauer, K. (2016). In search of the focus of attention in working memory: 13 years of the retro-cue effect. *Attention Perception & Psychophysics, 78*(7), 1839-1860. <https://doi.org/10.3758/s13414-016-1108-5>

Taylor, R., & Bays, P. M. (2018). Efficient coding in visual working memory accounts for stimulus-specific variations in recall. *Journal of Neuroscience, 38*(32), 7132-7142. <https://doi.org/10.1523/JNEUROSCI.1018-18.2018>

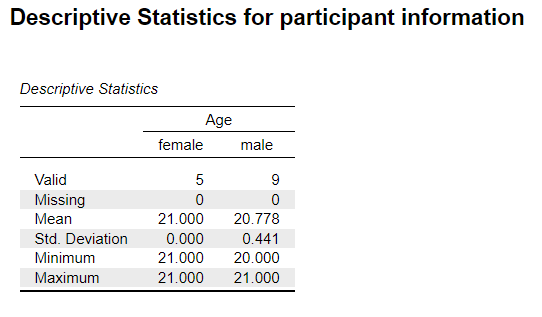
Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review, 114*(1), 104-132. <https://doi.org/10.1037/0033-295X.114.1.104>

Ye, C., Hu, Z., Li, H., Ristaniemi, T., Liu, Q., & Liu, T. (2017). A two-phase model of resource allocation in visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*(10), 1557-1566. <https://doi.org/10.1037/xlm0000376>

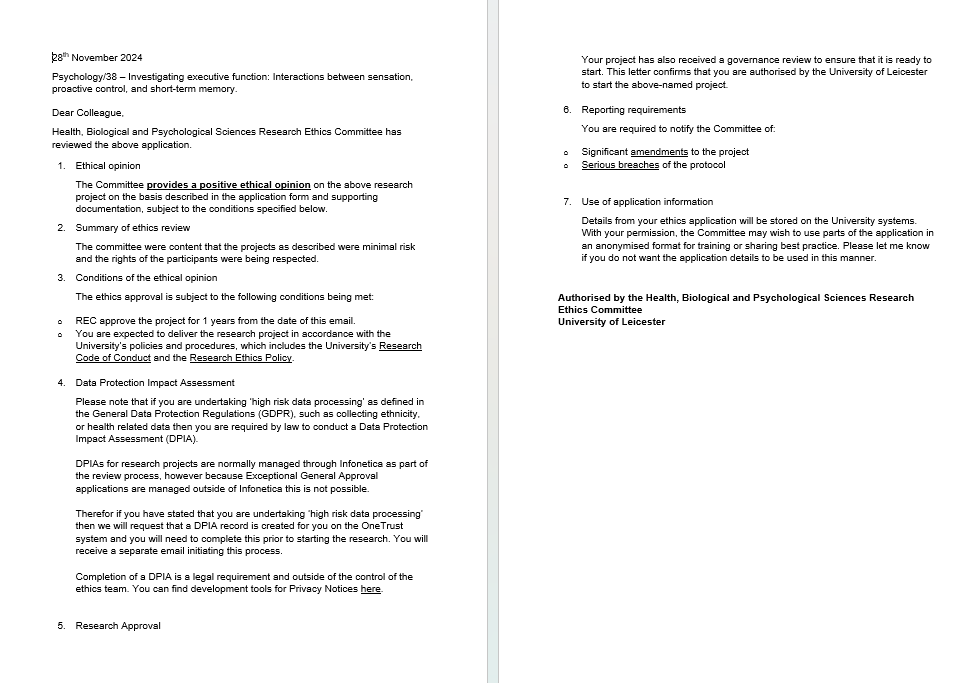
Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature, 453*(7192), 233-235. <https://doi.org/10.1038/nature06860>

**Appendices**

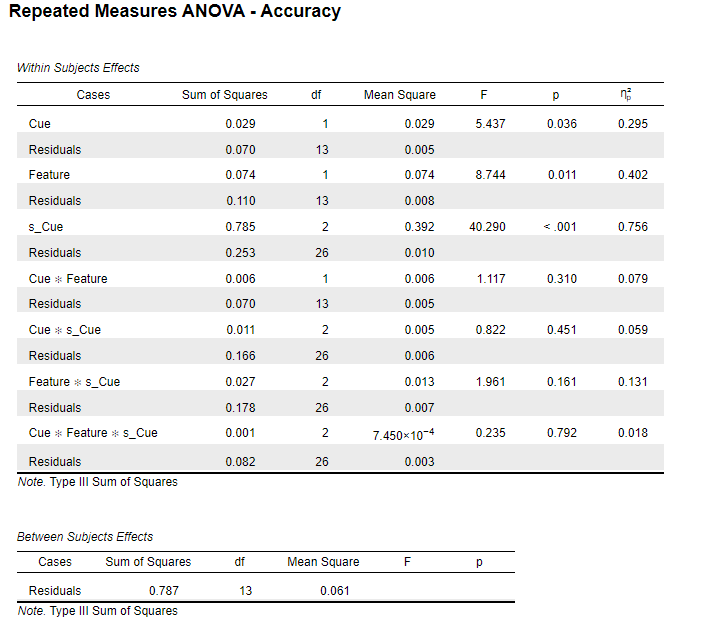
Appendix A – participant information



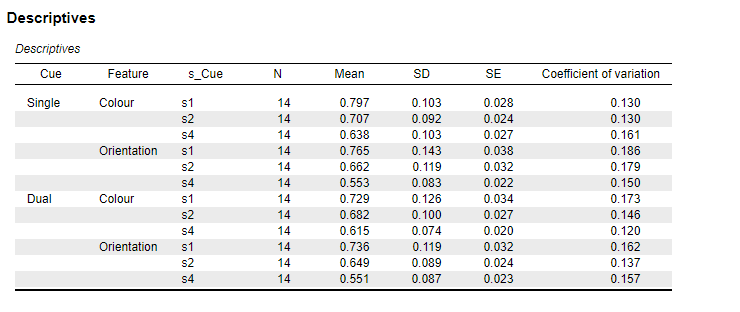
Appendix B – Ethics Approval Letter



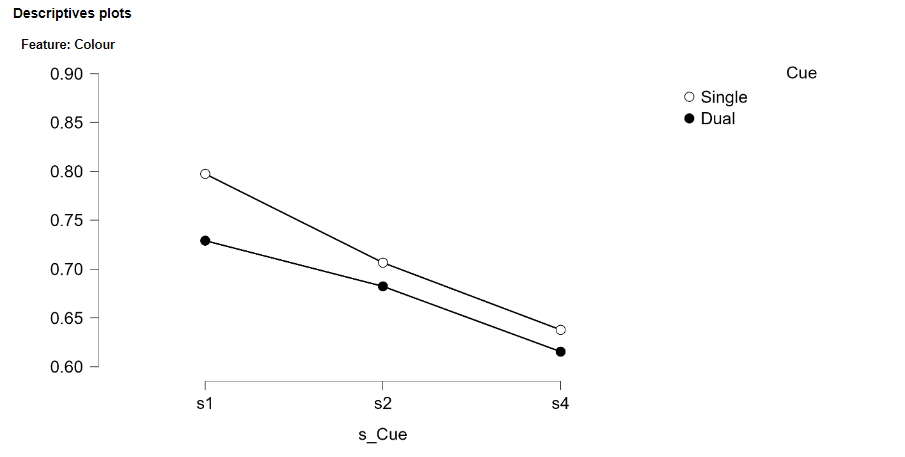
Appendix C – Repeated-Measures ANOVA for accuracy



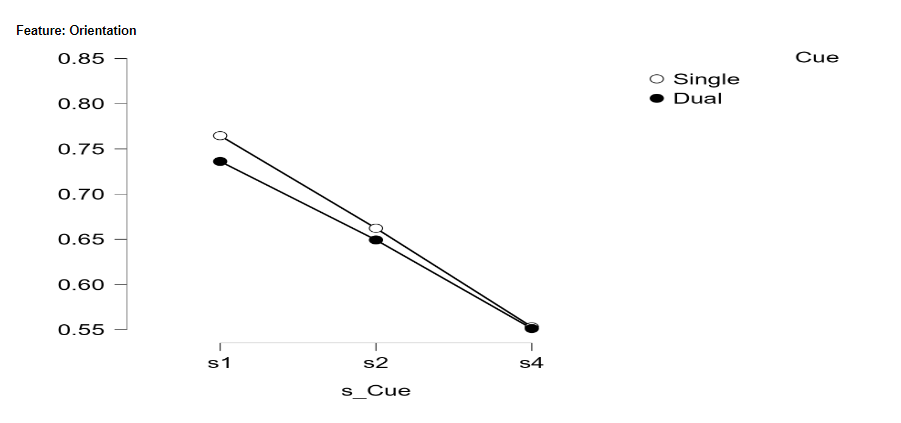
Appendix D – Descriptive Statistics for Accuracy



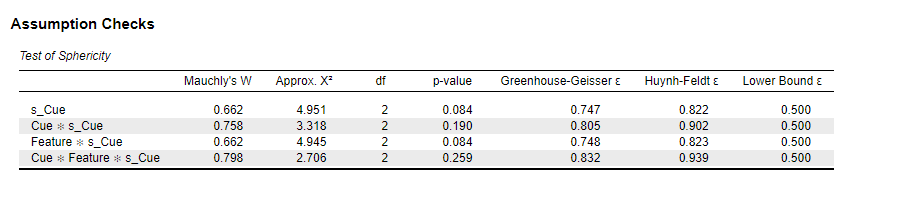
Appendix E – Descriptive plots for Accuracy in the Colour feature



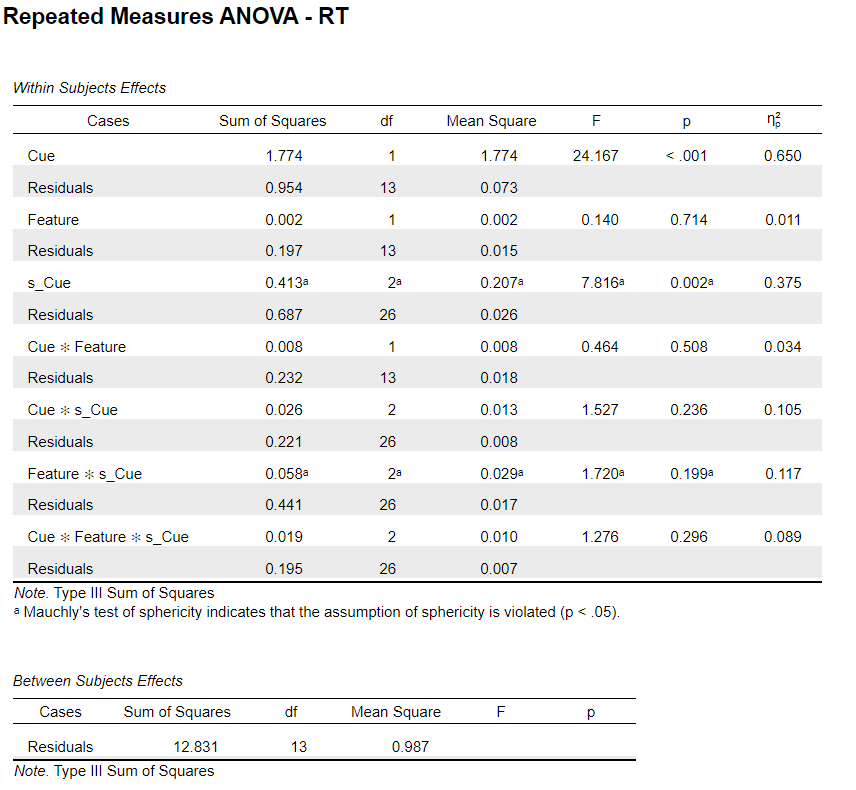
Appendix F – Descriptive plots for Accuracy in the Orientation feature



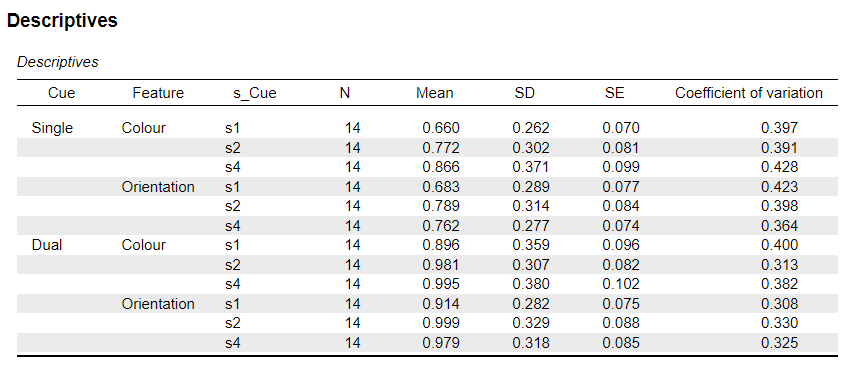
Appendix G – Assumption checks for accuracy



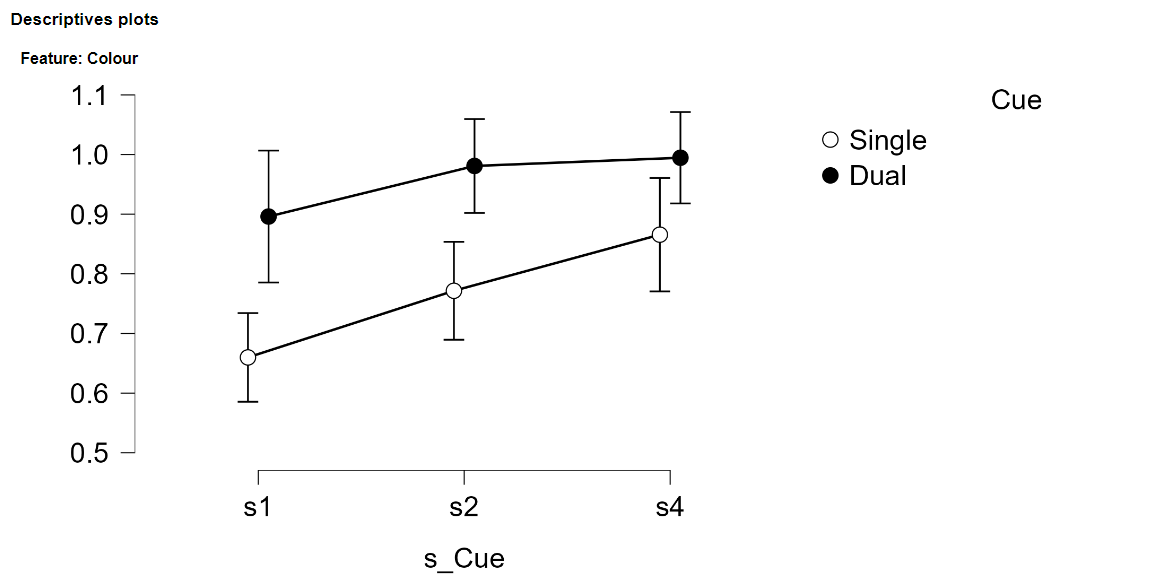
Appendix H – Repeated-Measures ANOVA for Reaction Time



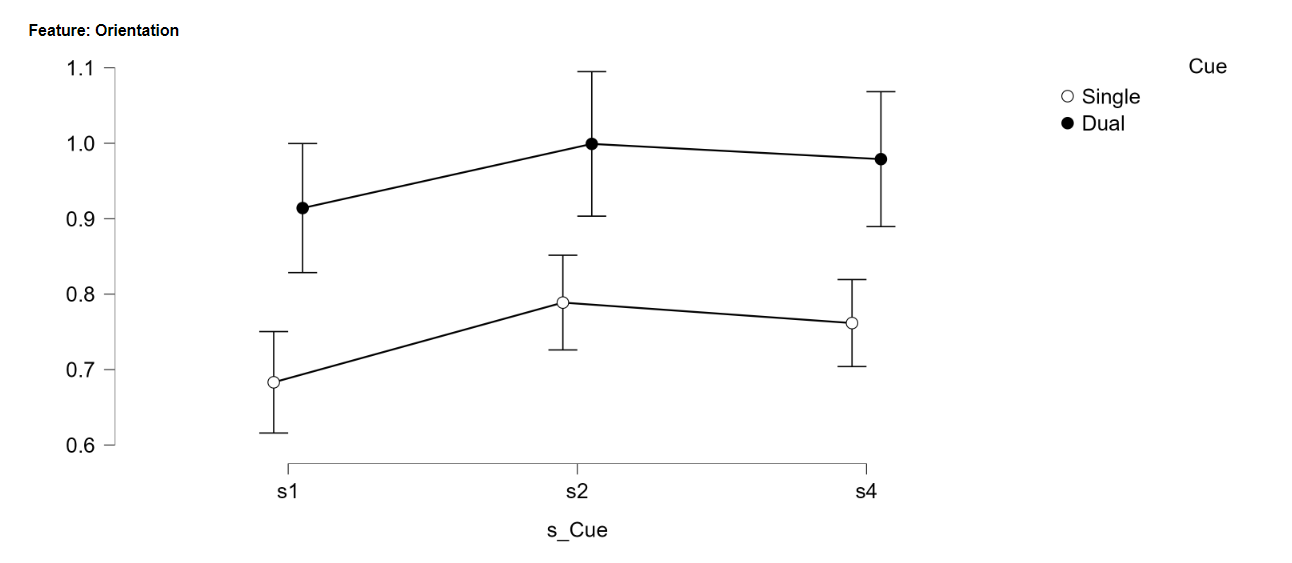
Appendix I – Descriptive statistics for Reaction Time



Appendix J – Descriptive plots for Colour feature in Reaction Time



Appendix K – Descriptive plots for Orientation feature in Reaction Time



Appendix L – Assumption checks for Reaction Time

