Comics Cognition: Working Memory and Situation Models in Visual Narrative Processing

Daniel Matkin

This thesis has been submitted in partial fulfilment of the requirements of
Sheffield Hallam University
for the degree of Doctor of Philosophy

November 2024

Declaration

I hereby declare that:

1. I have not been enrolled for another award of the University, or other academic or professional organisation, whilst undertaking my research degree.

2. None of the material contained in the thesis has been used in any other submission for an academic award

3. I certify that this thesis is my own work. The use of all published or other sources of material consulted have been properly and fully acknowledged.

4. The work undertaken towards the thesis has been conducted in accordance with the SHU Principles of Integrity in Research and the SHU Research Ethics Policy, and ethics approval has been granted for all research studies in the thesis.

5. The word count of the thesis is 53855 words

Name	*Daniel Matkin*
Date	*November 2024*
Award	*PhD*
Research Institute	*Social and Economic Research Institute*
Director(s) of Studies	*Dr Paul Aleixo*

Acknowledgements

First and foremost, I should thank my supervisors. Thanks go to Dr Jane Morgan for trying in vain to organise the chaos. Thanks as well go to Dr Diarmuid Verrier, who has been an excellent cheerleader and support throughout. But special thanks should go to Dr Paul Aleixo for giving me an opportunity that has changed my life.

There are so many people who deserve thanks for their parts in seeing me through this marathon. Thanks, of course, need to go to Summer, for putting up with the turmoil, pushing me forward on this journey and reminding me of purpose when I lost sight of it. Special thanks also go to Noel and Amanda, who have been a constant source of inspiration, motivation and a font of excellent advice over the years.

Enormous thanks should also go to Dominika, who saw me through the final and most challenging parts of this thesis and without whom, I would have been lost. Thanks also need to go to my fellow PhD grads, Dr Olly Hawkins and Dr Elin (Hawkins) Ivansson, for being the best PhD companions anyone could ask for. I'm so grateful to have shared this journey with you both.

Thanks should also go to my friends—Will, Elliot, Torin, Bennett, Joe, Steve-o, Cam, Owen, James, Jane, Sami, and Lu—who provided escape and a sense of normality at the pub, and who, like true friends, kindly never asked, *'Aren't you finished yet?'*

It would also be criminal not to give thanks to Lee and Jordan for dealing with my years of emails and facilitating my way through this work, as well as the SHU Psych, Soc and Crim community. To Kendal, Zoe, Piyali, Titus, Petra, Daniel, Ben, Liz, Hannah and Lewis, thank you for all the support, coffees and pints over the years.

And finally, thanks to my brother Nick, his wife Eve and my niece and nephew, Josh and Erin, who generously put up with me as I completed this project. And of course, thanks to Betty the Wiszla for providing copious thought-generating walks.

Abstract

Recent research in visual narrative processing, significantly informed by Visual Language Theory (Cohn, 2013), highlights the parallels between linguistic and visual modalities, particularly the role of grammar and structural principles in comprehension. While psycholinguistic approaches have advanced our understanding of visual narrative processing as akin to language, less attention has been paid to the cognitive mechanisms that facilitate this comprehension. This thesis adopts a cognitive science perspective, employing the multicomponent model of working memory and the situation model processing framework to investigate the interplay of visuospatial and verbal working memory in visual narrative processing. By addressing gaps in previous research, including the contrasting findings of Magliano et al. (2016) and Zhao et al. (2024), and testing novel stimuli designed specifically for this study, the thesis seeks to illuminate the roles of working memory and mental imagery in constructing and comprehending visual narratives.

The empirical work then investigates the roles of working memory and situation model processing in visual narrative comprehension through a series of experiments using the dual task paradigm as well as established situation model metrics. Experiments 1–3 reveal that visuospatial working memory, rather than verbal working memory, underpins the processing of visual narrative sequences when individual differences in working memory are controlled. Experiment 4 confirms that visual narratives are processed within the situation model framework, evidenced by faster reading times for coherent comics compared to incoherent ones and situation model updating behaviors observed in recognition tasks. However, Experiments 5 and 6 suggest that neither

visuospatial nor verbal working memory directly contribute to situation model processing, indicating that working memory supports aspects of narrative comprehension but is not integral to forming or updating situation models. Finally, Experiment 7 finds that participants with better mental imagery ability exhibited longer reading times, suggesting deeper engagement with visual narratives. These findings highlight the complexity of cognitive processes in visual narrative comprehension and the distinct roles of working memory and mental imagery.

This thesis makes significant contributions to understanding visual narrative processing by offering novel insights into the roles of working memory and mental imagery. The findings refine theoretical perspectives on the cognitive processes underpinning narrative comprehension, demonstrating that visuospatial working memory predominantly supports visual narrative processing, while verbal working memory plays a negligible role. Moreover, the unexpected relationship between vivid mental imagery and slower reading times suggests that mental imagery enhances narrative engagement, contributing to a deeper but more resource-intensive comprehension process. These results not only affirm the applicability of situation model frameworks to visual narratives but also advance methodological approaches for studying multimodal comprehension, providing a robust foundation for future investigations into the cognitive underpinnings of visual narrative processing.

Table of Contents

Table of Figures

List of Tables

Chapter 1 – Introduction

1.1 Opening Context and Overview

Research into visual narrative processing has recently expanded considerably, and has been significantly advanced by visual language theory (Cohn, 2013). This theory posits that comics are constructed and understood through language-like principles, and broadly speaking follow patterns of linguistic processing. While this growing field of research bridges linguistics and cognitive science, most studies have primarily adopted a psycholinguistic framework. The focus of this body of research has primarily been on demonstrating that the brain processes visual sequences similarly to language. Cohn and colleagues have shown that semantic, syntactic, and grammatical principles are crucial in visual narrative processing. However, comparatively little research has examined the specific cognitive processes and behaviours underlying this comprehension. The current project aims to address this gap by adopting a cognitive perspective and incorporating theoretical frameworks from cognitive science to enhance our understanding of visual narrative processing.

In this thesis, the multicomponent working memory model (Baddeley & Hitch, 1974) is used as a framework for online processing. Several experiments in which working memory, situation models and mental imagery are considered as processes through which cognitive and narrative understanding is achieved. Only two studies have previously investigated working memory in cognitive processing during visual narrative processing (Magliano et al., 2016; Zhao et al., 2024). Though the explicit focus of Magliano et al. (2016) was to explore inference during visual narrative processing, the findings of their experiments are interpreted as implicating both visuospatial and verbal working memory in visual narrative processing. In contrast, Zhao et al. (2024), after measuring their participants' visuospatial working memory ability, report no significant association of visuospatial working memory ability in processing visual narratives. As such, it remains unclear how important working memory is during this process.

The focus of the first three experiments is the contribution of different working memory components during online visual narrative processing. These experiments employed the dual-task paradigm in which participants were tasked with maintaining stimuli of a particular modality and were simultaneously asked to process a comic. By adapting the methods used by Magliano et al. (2016), the cognitive load imposed during the dual task procedure was tailored to the participants using a titration measure before the experiment. The results of these experiments demonstrate that visuospatial working memory processes are considerably affected by the requirement to process a comic, indicating that visuospatial resources are implicated in visual narrative processing. Conversely, verbal working memory was not affected by the requirement to process a comic when accounting for individualised working memory load limit. Thus, in both agreement and contrast to Magliano et al. (2016), the findings of these experiments demonstrate that verbal working memory is not implicated in visual narrative processing.

It is broadly accepted that the process by which narratives or discourse are comprehended is by the generation of mental models which represent the communicated content (van Dijk &

Kintsch, 1983; Gernsbacher et al., 1990; Zwaan et al., 1995; Zacks & Swallow, 2007). Until recently, work explicitly testing situation model processing in comics had been relatively rare. However, theorists have recently used the situation model framework to understand visual narrative processing (Cohn & Kutas, 2015; Cohn & Wittenberg, 2015; Cohn & Bender, 2017), demonstrating that situation model processing mechanisms also apply to visual narratives. This has been achieved using equivalent behavioural metrics and physiological measures that have previously established situation model-based processing in text. This research has then demonstrated that the distinct characteristics associated with situation model processing are associated with visual narratives.

The visual narratives used across the experiments in this thesis were short comic strip-based stimuli, which conveyed simple narratives. Given that these were novel stimuli, not previously used in experimentation, the fourth experiment sought to validate the application of the situation model processing framework to these specific comics stimuli. To do this, a coherence manipulation paradigm was implemented in which participants were presented with both coherent and incoherent comics, with the latter being generated from the base comics stimuli. This then provides the basis for the fifth and sixth experiments. To this end, the fourth experiment directly tests whether behaviourally based metrics predict situation model processing behaviours in the comics used in this project. Participants' reading times for coherent and incoherent comics were recorded. The participants were subjected to a recognition test of the panels they had seen. The results of this experiment indicated several interesting findings. Primarily, the reading and recognition time results verify that processing the current stimuli reflects characteristic situation model processing behaviours.

Having established that the situation model processing framework accounts for comics reading behaviours, Experiments 5 and 6 explore the roles of visuospatial and verbal working memory in visual narrative situation model processing. These experiments use a novel methodology combining the dual-task paradigm used in Experiments 1 to 3, with the coherence manipulation used in Experiment 4 to explore whether visuospatial or verbal working memory is explicitly activated in the visual narrative situation model creation process. Though the findings serve to advance and replicate those of Experiments 1 to 3, the findings did not provide support for the hypothesis that visuospatial or verbal working memory is activated during visual narrative situation model processing. These findings are then discussed in relation to the potential need for adjustments to the role of working memory in theories that attempt to outline visual narrative situation model-based processing.

As the role of working memory was determined to be non-significant in situation model processing, the final experiment explored the potential role of mental imagery ability in visual narrative processing. Mental imagery potentially has an important relationship with visual narratives as it has been demonstrated to contribute to the comprehension and functional processing of situation models. The relationship between mental imagery and situation model processing is interesting to consider. It is reasonably well established that situation models are dynamic mental representations of content that, when derived from text, directly depict the events described (Bower & Morrow, 1990; Haenggi et al., 1995; Glenberg et al., 1997). Mental imagery has essentially only been implicitly assumed to play a role in mental representations of narrative and discourse. Though there is a clear implicit link between situation model processing and mental imagery, few studies have explored the relationship, and fewer still have explored the relationship between mental imagery ability and visual narrative processing. The final experiment presented in this thesis then investigated the influence of mental imagery ability on coherent and incoherent visual narrative sequence reading speeds.

Together, the project's findings present several interesting contributions to the understanding of visual narrative processing as it attempts to tie together several areas of research that previously have received little attention in combination. Firstly, it provides contrary evidence to the current understanding of visuospatial and verbal working memory activation in visual narrative processing. Where previous research has indicated that verbal and visuospatial working memory both contribute to visual narrative processing, this project presents evidence across two studies that verbal working memory does not appear to be associated with processing wordless visual narratives. The ramifications of this are discussed in relation to visual narrative situation model processing and resultant models created from different modalities. The combination of the dual-task and coherence paradigms similarly produced interesting findings that further suggest that the role of working memory in visual narrative situation model processing should be considered more carefully. The final contribution of the thesis is the novel link demonstrated between reading speeds and mental imagery ability.

1.2 Theoretical Background

Two theoretical frameworks underpin the experiments presented in this thesis. These are the multicomponent theory of working memory and situation model processing. The final study also incorporates the concept of depictive mental imagery. The relationship between the working memory and situation model frameworks is interesting to consider from the perspective that they have often been linked with each other (Zwaan & Radvansky, 1998; Zacks et al., 2007; Zacks & Swallow, 2007; Gernsbacher, 1990; van Dijk & Kintsch, 1988; Johnson-Laird, 1983), but, as is outlined in the introduction of experiment 5, the evidential basis for relationships and links between the theories is less substantial than may be immediately intuitive or obvious. The same is arguably true of the relationships between mental imagery and situation model processing. This section then details the theoretical bases of the frameworks in a general way. The introductions preceding each experiment then outline specific gaps addressed by the experiments.

1.2.1 Working Memory

Working memory describes a specific type of active memory that is considered to be implicated in processing, holding and manipulating information over short periods. Several theories have addressed the concept of an activated portion of working memory, but the most productive of these has been the multicomponent theory (Baddeley & Hitch, 1974; Baddeley, 2000; Baddeley, Hitch & Allen, 2021). This theory outlines a "limited capacity system for the temporary storage and processing of information required for complex cognition" (Baddeley et al., 2021, p.11).

The multicomponent theory of working memory described by Baddeley and Hitch (1974) comprised three components. The phonological loop and the visuospatial sketchpad are subsystems controlled by a central executive component. As indicated by their names, the subcomponents are suggested to relate to specific types of information input. The visuospatial sketchpad is primarily responsible for processing visual and spatial information, whereas the phonological loop specialises in handling auditory information. Extending this idea, the phonological loop is also proposed to manage all forms of verbal input, whether spoken, written, or otherwise. Although it may seem intuitive to assume that written language would be processed through the visuospatial working memory—given its initial visual nature—written language is fundamentally conveyed through graphemes. Reading involves decoding these graphemes into phonemes, their auditory equivalents. This decoding process is facilitated by verbal working memory, as evidenced by its predictive power across various aspects of reading. This is demonstrated by several studies which have, for example, revealed links between phonological (or verbal) working memory and individual differences in reading (Daneman & Merikle, 1996), early reading acquisition (Gathercole et al., 1992; Wagner et al., 1997) and second language acquisition (Papagno & Vallar, 1995).

Visuospatial working memory, like the phonological loop, is a subsystem within the multicomponent working memory model. It is proposed to be responsible for temporarily storing and manipulating visual and spatial information, serving an equivalent function to the phonological loop but for incoming information in these modalities. Similarly to the way that verbal working memory is critical to tasks which relate to reading and linguistic comprehension, visuospatial working memory is implicated in functions such as spatial navigation (Gyselinck et al., 2007; Gyselinck et al., 2009; Meneghetti et al., 2009), nonverbal reasoning (Allen et al., 2019; Mamarella et al., 2018) and mental animation (Sims & Hegarty, 1997). The relative complexity of this component compared to the phonological loop is arguably reflected in the wide-ranging nature of what might be subsumed under the umbrella of "visuospatial processing".

The third component initially outlined by the multicomponent working memory model is the central executive (Baddeley & Hitch, 1974; Baddeley, 2000). The central executive was theorised to coordinate the other subsystems' activities, manage attention, and integrate information from long-term memory, thus playing a key role in higher-order cognitive processes (Baddeley, 2003; Miyake & Friedman, 2012). However, this is arguably the least well-understood and most generalised component of working memory (Baddeley, 2012). More recently, the central executive has been considered somewhat redundant by some prominent researchers who argue that the higher cognitive functions that this component was considered responsible for seem likely to be managed by an array of cognitive functions instead of simply one (Logie, 2016; Baddeley, 2007).

The central executive has often been described as a 'catch-all' component, managing complex processing that could not be categorised under the other subsystems. Despite significant research efforts, it remains a broad concept, representing the complexity of higher-order processing in relation to working memory, while allowing researchers to focus more concretely on the verbal and visuospatial components (Baddeley et al., 2021). In this thesis, while the central executive is acknowledged as a critical interface within working memory, the experiments do not address it in isolation. Instead, its functions are considered as underlying mechanisms supporting several experiments. For instance, the central executive is proposed to play an active role in resource distribution during dual-task experiments (experiments 1-3, 5, 6), acting as the overarching control mechanism. It is also implicated in interfacing with long-term memory (Oberauer, 2009; Baddeley, 2003) and, therefore, seems likely to be involved in situation model processing.

In 2000, the episodic buffer was introduced into the model (Baddeley, 2000). This component was added to account for the limitation that subsystem-based modality processing could not account for integrated verbal, visual, and spatial episodes activated during and created by working memory. The episodic buffer was then suggested as a similarly limited-capacity storage unit capable of combining verbal and visuospatial working memory information into a coherent representation. Along with the update to this model, it was also suggested that each subsystem had a bi-directional relationship with working memory information, meaning that information could flow both ways between the central executive and the subsystems, allowing for both the integration of incoming information and the retrieval of stored information to influence cognitive processes. In contrast, previously, this was considered only to be related to the role of the central executive. The updated model then regards each working memory subsystem as bi-directional in that each processes incoming information while fluidly drawing from long-term memory. As the episodic buffer is a relatively recent addition to the multicomponent working memory model, comparatively little research has been undertaken to understand it. Though this thesis does not directly address the episodic buffer component, comics and visual narrative processing that incorporate both visuospatial and verbal or phonological content arguably represent further unrealised research opportunities in which comics stimulus could provide significant benefit.

While both the central executive and the episodic buffer are important to the broader theory of working memory, the primary focus of this thesis is the role that the visuospatial and verbal subsystems play in visual narrative processing and visual narrative situation model processing. Several facets of working memory research underpin this decision. Firstly, visuospatial and verbal working memory roles are already well defined in that these subsystems are dedicated to processing distinct input types (Baddeley, 2000; Baddeley et al., 2021). As suggested, the central executive is purported to be related to overarching control. However, significant disagreement exists about the processes it controls and its base function or role (Repovš & Baddeley, 2006). The same is arguably true of the episodic buffer. While its role in providing the capacity for integrating information from the subsystems (Baddeley, 2000) is better defined than the central executive's role (which is somewhat surprising given that the central executive is one of the original components), the imprecise functionality of the episodic buffer component makes it significantly more difficult to isolate experimentally (Baddeley, 2012). Comparatively, many studies have demonstrated clear functionality of the visuospatial and verbal subsystems, and several established and widely used tests directly relate to these components (Lin & Matsumi, 2021; Gyselinck et al., 2008; Shah & Miyake, 1996). As such, the visuospatial and verbal subsystems' essential clarity makes it significantly simpler to devise experiments in which their isolated function and utility can be investigated during specific processes.

Though it is not universally agreed upon that working memory can be delineated into the subsystems outlined by the multi-component model, possibly one of the only widely agreed upon principles of working memory is its limited capacity. Historically, the concept of strictly limited conscious, active memory arises from Miller's (1956) 7 +/- 2. This has then been extended to each subcomponent and is evidenced by findings demonstrating that visuospatial and verbal working memory are responsible for different abilities. Studies, for example, have shown that visuospatial working memory is associated with spatial navigation tasks (Gyselinck et al., 2009), whereas verbal working memory is linked to linguistic comprehension tasks (Daneman & Merikle, 1995). Given that the verbal and visuospatial subsystems are dedicated to processing various types of information, the literature has raised the question of whether processing these distinct types of information requires differential resources (Logie, 2011; Baddeley & Repovs, 2006; Wickens, 2002). While some suggest that working memory draws from a single pool of resources regardless of the type of input (Cowan & Morey, 2005; 2006; Cowan & Saults, 2007; Morey, 2018), evidence indicates that, in practice, the system operates with a degree of specificity, where certain resources are more suited to particular types of information (Cocchini et al., 2002; Duff & Logie, 2001; Gyselinck et al., 2007). These findings illustrate that the limited capacity of working memory is reflected in the division and specialisation of its resources for different types of information..

Numerous studies have demonstrated selective interference between verbal and visuospatial working memory components. A common method used to explore this phenomenon is the dual-task paradigm, a long-standing experimental approach in which participants perform two tasks simultaneously, each with different processing and storage demands (Pashler, 1994). The rationale behind this paradigm is rooted in the concept of capacity limitation, making it a valuable tool for investigating working memory. When two tasks draw on the same capacity-limited resource, performance typically declines on at least one task. Conversely, if the tasks rely on separate resources or capacities, the secondary task has little to no effect on performance. This concept is also referred to as the selective interference paradigm. For example, studies involving pairwise combinations of verbal, visuospatial (e.g., perceptuomotor tracking), and visual working memory tasks have often found minimal disruption between tasks (Cocchini et al., 2002; Guérard & Tremblay, 2008; Soemer & Saito, 2016; Jarrold et al., 2011; Jarrold & Towse, 2006; Jarvis & Gathercole, 2003). However, evidence also exists for cross-domain interference, where tasks in different working memory domains affect each other, suggesting some overlap in resource demands (Ricker et al., 2010; Morey & Cowan, 2004; 2005). Despite this, there is general agreement in the literature that a central capacity is supplemented by modality-specific processes, allowing for some degree of separation in resource allocation.

Working memory development in children has been extensively researched, given its value as an educational marker (Gathercole et al, 2004; Alloway et al., 2009). One important finding is that younger children (around 4–5 years old) tend to rely more heavily on visual recall rather than auditory or verbal recall, compared to older children (Hitch et al., 1988). This supports the idea that visuospatial working memory capacities develop earlier than linguistic and verbal working memory capabilities. This developmental sequence is likely influenced by the fact that humans primarily intake information through visual and spatial modalities. Interestingly, visuospatial ability in children has been shown to improve independently of both age and verbal ability

(Bobrowicz et al., 2024). Such findings suggest that visuospatial and verbal working memory components have distinct developmental trajectories, with a gradual shift from visuospatial to verbal dominance as children grow. These differences align with the multicomponent model of working memory, which proposes complementary but distinct visuospatial and verbal systems.

Age-related decline in working memory is another well-established area of research. Numerous studies have demonstrated significant declines in both verbal and visuospatial working memory across age groups. For example, verbal working memory has been shown to experience age-related reductions, as demonstrated by meta-analyses (Bopp & Verhaeghen, 2005). Similarly, visuospatial working memory also shows declines with age (Brockmole & Logie, 2013; Markostamou & Coventry, 2021). However, these two subsystems appear to follow distinct patterns of decline. Research consistently shows that visuospatial working memory often exhibits a sharper decline in performance compared to verbal working memory, which tends to decline at a slower rate (Cansino et al., 2013; Cansino et al., 2020; Swanson, 2017; Levi & Heled, 2024). These findings mirror the differential developmental trajectories observed in children, further emphasising the unique and complementary nature of the visuospatial and verbal working memory systems.

The distinction between verbal and visuospatial working memory is also supported by neurophysiological evidence. At a basic level, considerable evidence indicates that separable brain areas are responsible for phonological (and consequently verbal) and visual and spatial processing. The prefrontal cortex has been widely demonstrated to be associated with working memory functionality (Smith & Jonides, 1999; Miller & Cohen, 2001; Kane & Engle, 2002; Wager & Smith, 2003). Further research has demonstrated a significant and consistent left-right split in the processes associated with verbal and visuospatial working memory (respectively). This has then been supported by transcranial magnetic stimulation (TMS) findings in which temporary lesions are created on specific parts of the brain to assess functionality. TMS findings have indicated, for example, that disruption caused to the left dorsolateral prefrontal cortex significantly affects verbal working memory ability (Osaka et al., 2007). Contrastingly, visuospatial working memory functionality is associated with right hemispheric activity (Jonides et al., 1993; Smith et al., 1996). Overall, there is a significant accumulation of evidence from behavioural, developmental, and neurophysiological sources that together strongly suggest that visuospatial and verbal working memory processing are, at least to some degree, domain-specific components within the broader construct of working memory.

The position adopted in this thesis, from a methodological standpoint in relation to working memory, is that the verbal and visuospatial working memory subsystems are independent but capable of integrated function. As discussed above, studies across behavioural, developmental and neurophysiological fields demonstrate this independence. Behavioural studies (Salway & Logie, 1995; Gyselinck et al., 2007; Cocchini et al., 2002) show clear patterns of selective interference by matched modality stimuli (i.e. Verbal-Verbal) but not unmatched modality stimuli (Verbal-Visuospatial). Developmentally, different trajectories characterise verbal and visuospatial working memory capacities (Gathercole et al., 2004). Moreover, neurophysiological studies underscore the separation of areas related to the functional use of each subsystem (Jonides et al., 1993). However, while the independence of these subsystems is well supported, it is equally important to acknowledge that in their typical use, the subsystems frequently interact. The concept of visuospatial bootstrapping exemplifies this interaction, occurring for instance when spatial cues serve to enhance verbal working memory (Darling et al., 2020; Darling & Havelka, 2010) and demonstrating that verbal working memory can be supplemented by visuospatial capacities.

In conclusion, the evidence points towards the requirement for a nuanced and practical understanding of working memory. The perspective taken on working memory as a processing framework in this thesis then reflects that independence in using dual-task experiments designed to assess the use of verbal and visuospatial working memory subsystems in visual narrative processing. This method allows for the ability to isolate and subsequently analyse each subsystem's functionality and contribution to visual narrative processing.

1.2.2 Situation Models

The term 'situation model' refers specifically to mental representations derived from narratives or discourse (van Dijk & Kintsch, 1983). These models are integral to understanding how individuals process and comprehend complex texts or spoken language. Over time, discourse processing theory has evolved and been incorporated into theories that account for general lived experience. This has been a particularly influential model of discourse comprehension and understanding, with newer terms emerging such as 'event model' (Zacks et al., 2007; Zacks & Swallow, 2007), of the event model theory, being used to refer to mental representations of not only narratives, but also consciously experienced episodes and related memories. Both situation model and event model fundamentally describe a 'mental representation created from experience whether derived from narrative, discourse or lived experience. Importantly, though the theory has evolved, and the processes described by different narrative and discourse comprehension theories have been explored and expanded upon, the foundational principles have largely remained similar across models.

Arguably, one of the most influential frameworks in this area is the Construction-Integration model proposed by van Dijk and Kintsch (1983). This model describes two fundamental processes that underpin the understanding of all narratives and discourse: construction and integration. In the construction phase, readers activate relevant, pre-existing information that helps build a basic contextual framework for comprehension. Following this in the example of a prose-based narrative, as reading progresses, the integration process occurs as activation based on the initial network of information then further activates associated networks. Importantly, this process is considered in that they occur simultaneously and repeatedly throughout reading. The similarly critical nature of this concept becomes clearer in later models of narrative comprehension. The model of associative activation is well evidenced by classical priming studies (Mayer & Schvaneveldt, 1971; Xavier Alario et al., 2000), in which, essentially, processing functions are shown to be facilitated by the activation of broad semantically related information networks. Importantly, this concept is also suggested to underpin the process of inference, which is a similarly critical feature of situation model processing, and is described later in this section.

The Construction-Integration model has also been influential for its characterisation of narrative processing, which is suggested to occur over three progressively abstract levels. As the large majority of research into this concept has used prose-based narratives, the example here follows the typical procedure of situation model creation from this medium. Surface-level processing is the initial level and describes processing the actual text as it appears. This level of processing is considered to engage lower-level perceptual processing. Though in prose-based narratives, the surface level of processing is considered to play a relatively minor subsequent role in the resulting situation model, the findings of Chapters 3 and 4 in this thesis indicate that the surface form plays a much more significant role in visual narrative processing. From the Construction-Integration perspective, surface-level processing captures the encoding of word forms, syntax, and basic structural elements of the text. These elements are retained only briefly during comprehension but provide the foundational building blocks for deeper semantic and inferential processes. The findings suggest that, in visual narratives, these surface-level details may not only persist longer but actively shape the integration of meaning across the narrative structure.

The propositional level of processing is where semantic processing and associated information networks are activated. For instance, the propositional components of the sentence, "Alfred returned with the keys to the sports car," such as "returned," "keys," "sports car," and "Alfred," each invoke broader concepts through interaction with the reader's long-term memory and pre-existing knowledge. For example, "returned" evokes the concept of re-arriving at a particular location, while "keys" and "sports car" activate their respective networks of associated information. In the context of the current thesis, the entity "Alfred" might also trigger pre-existing knowledge related to the Batman universe[1], illustrating the significant role of contextual spreading activation in comprehension. This example highlights how propositional processing integrates textual information with prior knowledge to construct meaning.

Finally, the highest level of comprehension, the situation model, represents a coherent mental representation of the situation constructed from the propositional level (Zwaan & Radvansky, 1998; Zwaan et al., 1998). At this level, readers' understanding extends beyond a series of connected propositions, integrating activated information from long-term memory and contextual knowledge into a dynamic and holistic representation of the communicated narrative. This process enables readers to track events, causal relationships, and characters' goals within the unfolding narrative, creating a mental simulation of the described scenario. This aspect of the Construction-Integration model (van Dijk & Kintsch, 1988) has been extensively studied and is now widely regarded as the primary mechanism through which narratives are processed. Crucially, the situation model allows for flexibility in comprehension by adapting to context, enabling the reader to fill in gaps, resolve ambiguities, and infer unstated information to fully understand the narrative's meaning.

Several theories have followed the Construction-Integration theory and have specifically explored the situation model level of processing. Building on van Dijk & Kintsch's work, the Structure-Building framework (Gernsbacher, 1990; Gernsbacher et al., 1998) sets out a similar staged theory of narrative processing. The procedure initially comprises laying a foundation for the

[1] For readers uninitiated in comics, Alfred Pennyworth is the reliable confidant of Bruce Wayne (Batman).

mental representation of the content, and then, as the reader progresses, new information from the narrative is mapped onto the foundation. Finally, when new information is incongruous with the existing structure of the mental representation, readers shift, creating a new structure onto which new congruous information is mapped. The initial construction of the framework is predicated on research, which demonstrates that reading times increase at the beginning of narratives, or when a new model must be created and this remains a long-standing finding related to situation models (Zwaan et al., 1998; Bestgen & Vonk, 2000; Rinck & Weber; 2003; Swets & Kurby, 2016).

As proposed by Zwaan and Radvansky (1998), the Event-Indexing model offered a significant advancement in narrative processing theory. This model built on the foundational Structure-Building framework by incorporating the concept of continuous updating, which adapted the concept of situation models as narrative processing to reflect the idea that they represent an ongoing, dynamically updated situation. This theory conceptualises situation models as "what is happening now" in a narrative. Further to this, the event indexing model proposed that this updating process operates dynamically across five key dimensions: temporal, spatial, causal (essentially why events have occurred, and this specifically includes physicality of events in a narrative), goal-based (completion of or re-establishment of goals) and protagonist based. Tracking these dimensions is suggested to enable readers to create a coherent situation model by continuously updating the model with new information over time along these core dimensions. This demonstrates that the adequate formation of situation models is vital for engagement with a given narrative.

Further building on the foundation of situation model-based narrative processing, Event Segmentation theory (Zacks & Swallow, 2007; Kurby & Zacks, 2008) offers a framework for understanding how we process the continuous stream of conscious information and organise it into conscious memory. Zacks and colleagues suggest that conscious experience is naturally segmented into distinct events which help organise and manage the constant stream of sensory and cognitive input. Although attempting to provide a framework for understanding more general conscious experiences, Event Segmentation Theory acknowledges that how we process narratives mirrors how we process incoming information streams and how they are structured in memory. Consequently, findings from the theory generalise from mental model processes and, as such, offer equally useful insights into narrative processing.

Overall, situation model theory has generated significant and wide-reaching insights into how individuals construct meaning from narratives. This is illustrated by the theory's progression from discourse and text-based narratives to informing how we create meaning from events and store them in memory. Not only has the theory covered these diverse areas, but as is demonstrated below, it has been applied to processing visual narratives with findings that interestingly mirror those related to text-based narratives, suggesting a modality-independent framework of narrative comprehension. The next section outlines the significant body of research which has explored the parallels between linguistic and visual narrative processing.

1.2.3 Visual Narratives and Situation Models

Recently, understanding graphic communication has been significantly advanced by the concept that visual narrative processing overlaps significantly with text—or prose-based linguistic processing. Visual Language Theory (Cohn, 2013) proposes that when graphic communication is presented in a structured sequence, linguistic principles such as grammar are innately applied, thus governing the structure of the communication. This underpins the suggestion that visual narratives are written in visual languages. A large body of research has since been given to the concepts of both a visual lexicon (Cohn & Foulsham, 2022; Cohn et al., 2016; Cohn, 2021) and understanding visual narrative grammar (Cohn & Kutas, 2017; Cohn et al., 2014; Cohn & Bender, 2017).

The grammatical principles applied to visual narratives, specifically comic panels, have had significant utility in demonstrating the similarities between text and visual narratives and, consequently, have demonstrated the application of situation model processing to visual narratives. In visual language theory, panels are considered to have specific structural functions that relate to the progression of a narrative arc. Establishers, Initials, Peaks and Releases are described by Visual Language Theory (Cohn, 2013) respectively as setting up a scene (Establisher), initiating the action of a narrative arc (Initial), marking the height of the narrative arc in terms of tension (Peak) and resolving the action of a narrative arc (Release). Another important concept to note is modification panels. In this category, the concept of Refiner panels refers to those that do not add to the narrative but hone information. Critically, they must relate to or refine information previously presented in another panel (Cohn, 2013)[2]. Under the framework of this theory, Cohn and colleagues have provided substantial evidence, using various methods, that visual narrative processing adheres to principles of situation model processing.

The findings of Cohn and Wittenberg (2015) highlight how the characteristics of visual narrative processing strongly resemble those of text-based situation model processing. These authors used a self-paced viewing time paradigm in which participants viewed sequences of panels, controlling the pace by pressing a button to proceed through the sequence. The sequences included both coherent sequences and 'scrambled' sequences. The panels of the scrambled sequences were considered to be semantically scrambled in that the panels of the stories were connected by narrative, but they were not in the 'correct' order. The 'Peak' elements of the comics were also manipulated by substituting them with either action stars or blank placeholder panels. Participants' viewing times and coherence ratings were recorded to assess visual narrative processing. Most clearly in line with situation model processing was the finding that coherent sequences promoted faster reading times. This sequential facilitation aligns with a large proportion of previous research on situation model processing, where coherence and continuity are demonstrated to facilitate processing and the creation of mental representations.

Given that inference is a critical part of situation models, anaphor is an important concept in narrative processing, and this is particularly interesting to consider in relation to the visual narrative situation model processing. In text, anaphor refers to incomplete indexical referents.

[2] Please note this is not an exhaustive list.

These references require being set up in previous clauses to be understood, and include terms such as he, she, or they. These referents maintain coherence without necessitating an overt link back to their referent. The use of anaphora in visual narratives is not an immediately intuitive concept because visual narrative communication is, by nature, directly depictive. Consequently, indexical reference must naturally refer explicitly to the narrative component in a visual lexicon, meaning that arguably, a visual narrative reference cannot be ambiguously indexical in the same manner as an anaphoric text reference. However, Cohn et al. (2024) argues that reading times for 'refiners' indicate that these panels act as anaphora in visual narratives, given that they refer back to something previously and add detail to that component (Cohn, 2013; Cohn et al., 2024). The role of refiners as a kind of visual narrative anaphora is then demonstrated by findings showing that reading times, which indicate processing continuity, are significantly affected by both placing the anaphoric refiner in the incorrect sequential position and similarly by placing this type of panel too far away from its base referent. This highlights how visual narrative processing mirrors linguistic processing in its reliance on referential cohesion but adapts to the visual medium by employing its own distinct systems of representation. Together, reading time findings (Cohn et al., 2024; Cohn & Wittenberg, 2015; Brich et al., 2024) clearly demonstrate a unified narrative processing framework applicable across mediums.

As outlined above, situation models are characterised as dynamic representations of 'what is happening now'. As narratives unfold, it is necessary and important that new information is integrated into the situation model to reflect the current state of the narrative events accurately. This ongoing integration is illustrated by findings demonstrating a downward trend in reading times across panels during discrete events or narrative arcs (Cohn et al., 2024; Cohn & Wittenberg, 2015; Brich et al., 2024). As the process is iterative, where newly presented, relevant details prompt an update in the situation model, this also triggers further predictive processing. This updating to prediction has been demonstrated using reaction times whereby coherence in visual narratives has been shown to facilitate faster reaction times to target panels (Cohn et al., 2012) in contrast to semantically scrambled sequences (as above) in which readers cannot predict the next panel. This effect then aligns with the predictive mechanisms inherent to situation model processing. In turn, this highlights both the applicability of situation model processing theory to visual narrative comprehension and the efficacy of using behavioural methods with visual narratives to test facets of this theory.

A substantial amount of electrophysiological evidence further supports the application of situation model processing to visual narratives. Electrophysiological studies of text-based narratives support the effects of coherence-facilitated prediction and updating in situation models. The N400 is an event-related potential observed in the brain under electroencephalography (EEG) (Kutas & Hillyard, 1980). This response is considered to provide a neurophysiological marker of semantic incongruence. For example, the N400 amplitude related to the sentence "the laptop was too slow to cook" would be significantly larger than that of "the laptop was too slow to use." Findings related to the N400 support the situation model processing concept because this neural response reflects a neuro-physical marker of these predictive processes in action. Specifically, the amplitude of the N400 demonstrates how readers actively predict upcoming information based on the coherence of the narrative context, allowing researchers to track comprehension as it unfolds.

This marker has been demonstrated in visual narrative processing (Cohn et al., 2012; West & Holcomb, 2002), which further reinforces the suggestion that visual narratives are processed under the same framework as linguistic processes.

Further evidence for this has also come from a physiological marker related to situation model updating. The P600 response is a similar electrophysiological response, previously considered to represent syntactic anomaly (Ousterhout & Holcomb, 1992). However, more recently, the interpretation of the P600 response has been reconsidered as potentially reflecting semantic integration processes or—in terms related to situation model processing—updating (Brouwer et al., 2012; Delogu et al., 2018; Brouwer et al., 2016; Kuperberg et al., 2020). This event-related potential has been demonstrated in visual narrative processing across several studies (Cohn et al., 2014; Cohn & Kutas, 2015; 2017; Cohn et al., 2017). However, this interpretation should be tempered by the fact though this response is suggested to represent situation model updating, this is potentially difficult to differentiate from syntactic anomaly. More convincingly, the P600 has recently been associated with visual narrative comprehension proficiency (Coderre & Cohn, 2024). That P600 amplitudes were differentially affected by proficiency arguably presents some of the most compelling evidence for the concept that visual narratives engage mechanisms similar to language for visual narrative situation model processing, as this parallels the proficiency effect in text-based situation model processing for second languages (Leon Guerrero et al., 2021; Zwaan & Brown, 1996).

Overall, compelling evidence suggests that situation model processing theory provides a unified cognitive framework that spans both linguistic and visual modalities. Through visual language theory, researchers have demonstrated that structured visual narratives adhere to grammatical principles that facilitate and permit cohesive processing where they are applied. Other key findings demonstrate that visual narratives can utilise linguistic processing components like anaphora and the parallels with text-based processing mechanisms for the predictive facilitation and integration processes that situation model processing generates. Neurophysiological markers then further corroborate these findings. Together, this research presents a strong case for the concept of a shared cognitive architecture for visual and textual narratives, leading to the creation of dynamic mental representations.

1.2.4 Working Memory, Situation Models and Visual Narratives

The working memory and situation model processing frameworks, particularly the event model processing perspective, fit together intuitively. Working memory provides a framework for active maintenance and manipulation of incoming information that encompasses both phonological/verbal and visuospatial information. Situation model processing describes how we actively maintain the current state of a narrative, iteratively updating the contents of the model when pertinent information is registered. Perhaps because this link is intuitively clear, there is a generally accepted assumption amongst situation model literature that working memory underpins situation model processing. This assumption, and the scarcity of evidence available, is further outlined in the introduction to Experiment 5 (Chapter 4).

In the previous section on situation models and visual narrative processing, evidence demonstrates clear situation model-based processing parallels between linguistic and visual narrative processing. Given that situation model processing is assumed to be underpinned by working memory, it is reasonable to suggest that visual narrative processing can also be assumed to be underpinned by the working memory framework. Only two published studies have explored the link between working memory, visual narratives, and situation model processing.

Magliano et al. (2016) reports three experiments. In the first, participants were asked to view three-image visual narratives. These narratives contained what they term, beginning states, bridging events and end states. The sequences were manipulated where they were presented to the participants, either with or without the bridging event. Without the bridging event, the participants were presented with a blank panel. Their panel viewing times were recorded, and these indicated a clear effect of the requirement to generate inference when the bridging event was missing. This was ascribed based on the finding that end-state panels were viewed longer when the bridging event was missing.

Their second experiment introduced the concept of working memory and used the dual-task paradigm. In this experiment their participants were required to complete the same procedure as the first experiment, but additionally, they were assigned to either the verbal, visuospatial or control conditions. Participants in the verbal or visuospatial conditions were required to encode a verbal or visuospatial working memory stimulus and maintain this as a concurrent cognitive load while viewing the comics. This experiment's dependent variable was the time taken to view the panels. The results of this experiment indicated that maintaining either a verbal or visuospatial working memory load reduced the time the participants viewed the end-state panels. The authors then interpret these reduced viewing times with a working memory load present to indicate that fewer working memory resources were available to the participants, meaning that participants were less able to make inferences between the images because of the concurrent load. Consequently, they conclude that both verbal and visuospatial working memory are implicated in generating inferences in visual narrative processing and, by extension, that both types of resources are recruited throughout visual narrative processing.

Using the same paradigm, the third experiment reported by Magliano et al. (2016) explores the role of subvocalisation by using this as a concurrent task. However, in this case, their results indicated that subvocalisation caused no more significant interference than the control task to end-state panel reading times. The findings of Magliano et al. (2016) are collectively interpreted as indicating that both linguistic and visuospatial working memory resources contribute to bridging inferences in visual narrative processing. By extension, these findings also indicate that both verbal and visuospatial working memory resources are more generally implicated in visual narrative processing. This is on the basis that both resources are implicated in inference and, therefore, seem likely not to be limited only to inference.

One further experiment has been conducted on the use of working memory in visual narrative processing. Zhao et al. (2024) adopt the methods and metrics of Magliano et al. (2016) in combination with the methods of Cohn and Wittenberg (2015). Where the comics used by

Magliano and colleagues had an absent bridging effect, the comics used by Zhao et al. (2024) used complete beginning, bridging and end state sequences or panels with the bridging event replaced with an action star. Actions stars are a common technique used in comics whereby the action of a panel is occluded. This is typically used to symbolise impact or an altercation, however in Zhao et al., (2024) action stars were used to provoke the participants to infer content. Their study focused on the ability of older adults to perform inference generation in visual narratives with action stars in place of the narrative peak. In combination with measuring the participants' panel reading times, Zhao et al. (2024) also measured their participants' visuospatial working memory ability using a Corsi-type sequence. Their results replicated those of Magliano et al. (2016) in that the different bridging event panels (actions stars or sequence congruent panels) caused a significant difference in final panel viewing times. However, the findings of a linear mixed effects regression model did not indicate any significant impact of visuospatial working memory on viewing times or any interaction between the type of bridging panel and visuospatial working memory ability.

Consequently, the only currently available literature that has explored the concept of working memory in visual narrative processing presents contrasting findings concerning its role in processing. This contrast highlights a gap in our understanding of how different components of working memory contribute to visual narrative processing. On the one hand, Magliano et al. (2016) present findings which indicate that both verbal and visuospatial working memory components are implicated in visual narrative processing. Conversely, Zhao et al. (2024) challenge this, finding no significant effect of visuospatial working memory in visual narrative processing. As is outlined in the following section which identifies the thesis aims, this gap was then a focus of the current project.

1.3 Thesis Aims and Objectives

Research on visual narrative processing has undergone a significant recent evolution driven by the psycholinguistic perspective that visual narrative processing aligns with well-established language-based structures. Cohn's Visual Narrative Grammar (2013) has consistently been used to argue that visual narratives rely on adjacent processing principles to language. More than a decade of research on this topic has generated significant and consistent evidence that visual and verbal narrative understanding share a complex but overlapping cognitive processing infrastructure.

Despite significant advances, research applying cognitive frameworks to visual narratives remains limited, leaving gaps in our understanding of the processes that support visual narrative comprehension. While psycholinguistics research has demonstrated structural similarities between language processing visual narrative situation model processing, further research has only explored how different memory systems contribute to visual narrative processing. For example, very few studies have tested how individuals retain and manipulate information in real-time as they interpret visual narratives, and this leaves the role of working memory in processing these types of narratives particularly underexplored. One study has demonstrated that both visuospatial and verbal working memory play roles in visual narrative processing (Magliano et al., 2016), but there has also not been any experimental validation of these findings.

The concept of situation models is central to visual narrative processing, yet little research has examined the role of working memory in this context. No studies have directly bridged the gap between situation models and working memory in visual narratives, leaving a crucial area of cognitive processing underexplored. Cognitive theories propose that situation models enable individuals to construct mental representations of narratives. Similarly, theories of narrative processing emphasise working memory as the cognitive foundation that supports situation model processing. This connection is logical, given that the primary functions of working memory are to maintain and manipulate information for processing. However, research combining working memory paradigms with situation model processing remains scarce, even in studies focused on text-based narratives. In the field of visual narratives, these investigations are even more limited.

Several key research questions then guide the current thesis. It intends to investigate cognitive mechanisms that underpin visual narrative processing, focusing on visuospatial and verbal working memory. Specifically, the thesis aims to address how the distinct verbal and visuospatial working memory components are engaged during visual narrative processing. In investigating these processes and using these components in visual narrative processing, the thesis aims to contribute to understanding the link between working memory and situation model processing in visual narrative processing.

To achieve this, the thesis has several specific objectives addressed by each chapter detailing empirical work. The first objective was to examine and clarify the activation of verbal and visuospatial working memory components in real-time visual narrative processing. This involved determining the degree of activation of different working memory components in visual narrative processing using the dual-task paradigm but with the important methodological addition of titrating individual differences to minimise their influence. It was considered that fulfilling this objective would contribute to validating previous findings about using working memory in visual narrative processing.

The second objective of the thesis was to confirm that situation model processing theory predicted behavioural responses to the stimuli used in the current project. This aim was critical for several reasons. Firstly, it was important to establish that the stimuli were suitable for use in experiments designed to investigate the use of working memory in visual narrative situation model processing. Secondly, as these stimuli were created specifically for the thesis, it was important to verify that they conformed to the expectations of processing behaviours outlined by previous situation models and visual narrative processing research.

The final objective of the thesis was to investigate the use of visuospatial and verbal working memory in visual narrative situation models. To this end, the dual-task paradigm used in the initial set of experiments and the coherence manipulation paradigm—used to establish that the comics were being processed as expected under the situation model processing framework—were combined.

Overall, then this thesis aims to explore and investigate the cognitive mechanisms that underpin visual narrative processing with a particular focus on the roles of visuospatial and verbal working memory. Drawing on the foundational frameworks of working memory and situation models, it aims to contribute a perspective on the fundamental processing behaviours related to visual narratives.

Chapter 2

2.1 Experiment 1 - Visuospatial Working Memory in Comics Processing

2.1.1 Introduction

As discussed in the previous chapter, research on the role of working memory in visual narrative processing remains sparse. In line with the framework proposed by Magliano et al. (2016), this initial experiment investigates the contribution of visuospatial working memory to visual narrative processing using the dual-task paradigm. However, this experiment introduces several methodological refinements to enhance the approach. The experiment had two key objectives. Firstly, it intends to examine whether visuospatial working memory plays a critical role in processing visual narratives and thereby aims to validate the findings of Magliano et al. (2016) in this regard.

Magliano et al. (2016) explore the use of visuospatial and verbal working memory in generating bridging inferences during visual narrative sequence comprehension. Their study used the dual-task paradigm in combination with reading times. In their first experiment, participants viewed visual narrative sequences with or without bridging events. Their reading times for the visual narrative element that followed the missing bridging events were longer than when the bridging event was complete. This serves as the basis for the findings of the second experiment in which their participants were required to encode and then maintain either a visuospatial or verbal working memory load while they, again, viewed visual narrative sequences with or without bridging events. The participant's panel reading times were recorded and analysed as an indicator of the use of visuospatial or verbal working memory. Their results indicated that the additional requirement of either a visuospatial or verbal working memory stimulus reduced the viewing time effect of their first experiment, which is interpreted as suggesting that both visuospatial and verbal working memory systems contribute to visual narrative processing. This conclusion is drawn from the idea that increasing the cognitive load in either working memory domain interferes with the processes required to create bridging inferences and, therefore, demonstrates a reliance on both visuospatial and verbal working memory resources. In other words, the suggestion of Magliano et al. (2016) is that the participant's ability to generate bridging inferences was reduced by the requirement to maintain a visuospatial or verbal working memory load, and thus, their panel viewing times were significantly reduced.

In relation to the use of visuospatial working memory, a recent article presents findings that present something of a contradiction to those of Magliano et al. (2016). Zhao et al. (2024) similarly explores visuospatial working memory in visual narrative processing. In their experiment, participants reading times of sequences similar to those of Magliano et al. (2016), in which bridging events were present or replaced with action stars, were recorded. Action star here refers to a device typically used in comics to indicate collision or impact where the actual narrative depiction is replaced by a jagged shape to represent this. As well as this, the participants also completed the Corsi block test. This test of visuospatial working memory sees participants view flashing block sequences of increasing length, which they are then required to repeat. In contrast to Magliano et al. (2016), their findings indicated that, though the presence of a bridging event significantly affected the participant's reading times, this was not related to the participant's visuospatial working memory ability as measured by the Corsi block test. This is interpreted as indicating that visuospatial working memory is not heavily relied upon during visual narrative processing. As such, the currently available research presents contrasting views on the contribution of visuospatial working memory to visual narrative processing. As such, the discrepancy between these findings is important to explore and resolve.

The dual-task paradigm has been used extensively in previous research, which explores the functions and mechanisms of working memory (Gyselinck et al., 2007; 2009; Shah & Miyake, 1996; Della sala et al., 1999; Lin & Matsumi, 2021). The underlying theory of this experimental paradigm is straightforward: a primary and a secondary task are paired and then completed at the same time. Performance is then measured on the primary task and compared against a single-task condition where only the primary task is completed. The degree to which performance on these two tasks is different is referred to as the dual-task cost. The absence of dual-task cost suggests that the primary and secondary tasks rely on distinct processes and operate independently of each other. Alternatively, the presence of dual-task cost is taken to indicate that primary and secondary tasks share a common resource or process. Thus, the nature of the interference can be used to infer the activation of capacities or resources required by different kinds of processing behaviours. This is referred to as the capacity-sharing approach (Pashler, 1994).

Capacity limitation is one of the fundamental facets of working memory that is almost universally agreed upon by theoretical models (Oberauer et al., 2016; Halford et al., 2007; Logie et al., 2021). Further to this, capacities in both verbal and visuospatial domains are well established to vary by individual (Oberauer et al., 2008; Feldman-Barrett et al., 2008; Jarrold & Towse, 2006). These differing capacities have then been linked to differing levels of cognitive ability. Greater verbal working memory capacity has, for example, been shown to predict more effective reading comprehension in children (Seigneuric et al., 2000) as well as faster prediction of what words mean thus facilitating language comprehension (Li & Qu, 2024; Huettig & Janse, 2016). Comparatively, much less research has been conducted on visuospatial working memory. However, greater capacity in this domain has been shown to predict a significant amount of variance in children's mathematics ability (Giofrè et al., 2018; Li & Geary, 2013). Individualised capacity is then a critical feature of verbal and visuospatial working memory domains as both have been demonstrated to affect cognitive functioning.

A logical extension of individualised capacities in verbal and visuospatial working memory domains is then to account for this in dual-task procedures. However, the concept of tailoring working memory load during dual-task experiments has only recently come to prominence. Recent studies have demonstrated that cross-modality interference (i.e. the effect of a primary task in one modality and completing a secondary task in another modality) only begins to detrimentally affect working memory maintenance when the maintenance task is at the limit of an individual's capacity (Doherty & Logie, 2016; Izmalkova et al., 2022). In other words, if an individual's digit span score is five items, but they are asked to remember seven while simultaneously performing a secondary task, their accuracy in remembering the seven-digit items will automatically suffer because the task already exceeds their capacity limit. Whatever interference effect or dual-task cost then arises from a task like this would, to some degree, be confounded by the initial working memory component overload. Conversely, the same would also be true if the participant was underloaded. Consequently, in dual-task studies, a form of working memory capacity titration is arguably necessary to more accurately capture the processes at work.

This gives some cause for caution in interpreting the results related to working memory domain implication presented in Magliano et al. (2016) for two reasons. Firstly, in their second experiment, which explores visuospatial working memory activation in visual narrative processing, their analysis primarily relates to the completion of the secondary task (panel reading times) as opposed to the participant's ability to complete the primary working memory task. This represents something of an issue in that focusing on the secondary task in a dual-task paradigm may not directly capture the cognitive load or capacity demands of the primary task, which arguably should be the focus of analysis when assessing working memory engagement. In using the secondary task reading time metric, the study then ultimately comes to rely on indirect measures of cognitive capacity.

Secondly, there is a methodological issue related to the blanket use of seven-item length working memory stimuli. Although the authors of the study report having piloted four and seven-length stimuli to account for variation in individualised working memory capacity, their uniform application of a seven-length stimulus across all participants in their dual-task trials is likely to have automatically caused a level of interference in some proportion of their participants. This is evidenced by their findings, which show that participants could only recall the verbal working memory and visuospatial working memory sequences used as dual-task loads with average accuracies of 67% and 55%, respectively. Consequently, the interference demonstrated by their dual-task approach, which they suggest reveals the use of both verbal and visuospatial working memory domains in processing image sequences, may instead reflect the lack of sufficient account for individual differences in working memory capacity.

Arguably, a more effective and direct method of determining the use of different working memory resources in comics processing would be to first tailor working memory loads to participants and then, following a dual-task procedure, measure the accuracy with which the participants were able to recall a verbal or visuospatial working memory stimulus. This then broadly was the method of the first experiment presented here, which aimed to test the use of visuospatial working memory in image sequence processing. It was hypothesised that if visuospatial working memory contributes to processing visual narrative sequences, the maintenance of a visuospatial working memory stimulus will be significantly affected by the requirement to simultaneously process a comic strip.

2.1.2 Method

2.1.2.1 Participants

40 participants (28 Female) completed this study with a mean age of 20 (*SD* = 3.83). Participants were recruited using a course credit mechanism, and thus, all were enrolled in an

undergraduate psychology course. All participants had normal or corrected to-normal vision. All participants gave informed consent to participate in the study and were naïve to the purposes and intentions of the study. The experiment was approved by the University's Research Ethics Committee in line with the ethical procedures.

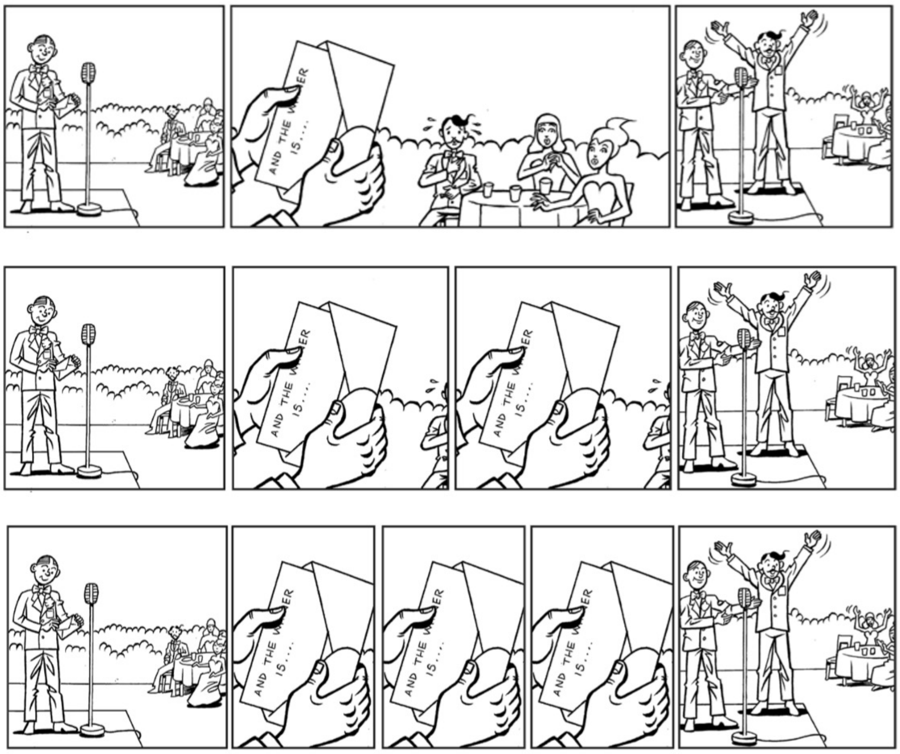
2.1.2.2 Materials

Six comic strip stimuli were used in this experiment (Arrow, Award, Boomerang, Boxing, Bus, Dog, Sailor – All comics stimuli are depicted in Appendices A - M). The Award comic was used consistently as the practice trial material. It is important to highlight here that the experimental materials used in this project were initially designed to investigate the concept of time in comics and its relationship to the use of panels. Originally, this was one of the primary aims of the project as a whole. However, as the research progressed, the focus evolved to testing the concepts of working memory activation during visual narrative processing instead. This shift in focus explains why the panel variant manipulation (described below) was included in the experimental designs of this and the subsequent two experiments presented in this chapter. While the relationship between panel manipulation and working memory will be analysed to ensure it did not inadvertently affect visuospatial working memory, no further commentary will be provided on this aspect.

Given the project's initial aim, the comic materials were drawn specifically with the intention of testing time-based concepts in visual narratives and, thus, variants of each comic were created, which manipulated the number of panels but were also designed to minimise the variable effects of both the physical space that the comic occupied as well as the semantic content. To this end, each of the comics presented a simple narrative which depicted or inferred motion as well as an indeterminate amount of time passing across three separate event units. Under visual narrative grammar, the function of these panels would be described as Initials, Peaks and Releases (Cohn, 2013). Initials, also acting somewhat in the role of establishers, setup the action. The Peaks or middle component of the comics depict the moment of maximum tension of the sequence, and finally, the Releases resolve the tension of the narrative. The design philosophy underlying these was to create complete visual narratives that were as simple as possible.

Structurally, the first and last panels of the comics were created to be equivalently sized. The configuration of the middle panels, however, was varied to allow for testing while also providing control over both the semantic components of the narratives and the physical space that the comics occupied. In each of the comics, the primary content of the middle panel remained consistent; however, in the four and five-panel variants, the content was repeated, increasing the number of panels in the comic, with these panels being decreased in size. This allowed for an investigation of the effect of panel number on processing while not significantly altering the narratives. Figure 1 shows an example of each comic length.

Figure 1 *Panel Variants of the Comics Stimuli used in Experiments 1 – 3*



The comics used in this experiment were organised into three blocks to manage their presentation. Each block included all six unique experiment phase comics (Arrow, Boomerang, Boxing, Bus, Dog, Sailor) and two comics for each panel length variant. Figure 1 depicts the panel variants. For example, each block contained two comics with three panels, two with four panels, and two with five. Each participant was shown only one block, meaning they were presented with all six unique comics and two examples of each panel-length variant. This organisation ensured participants experienced all content variations and a balanced distribution of panel lengths within their assigned block.

All participants completed the experiment on a HP Elitebook laptop with a screen resolution of 1920 × 1080 with a connected mouse, which the participants were specifically instructed to use instead of the trackpad for anything that required mouse-based input.[3]

2.1.2.3 Measures

The Visual Language Fluency Index was used to measure participants' comics reading experience. While this is not presented in this first experiment, it was used to ensure parity between the participants who completed Experiments 1, 2 and 3 in their experience with visual languages. This survey uses Likert-style responses and asks participants to respond to a number of questions about the frequency with which they read comics both "while growing up" as well as how frequently they read comics currently. Answers from the questions are computed into the participant's Visual Language Fluency Index. This experiment treats this variable as indicating experience with comics, with higher scores indicating more experience. Cohn (2012) proposes that responses to this survey can be classified into groups, with 0 to 11 labelled as low visual language fluency, 12 - 22 representing an average score and 22 + representing high visual language fluency.

The Corsi block test, was used in this experiment as a well-established measure of visuospatial working memory. This test was developed as a counterpart to verbal working memory

[3] Due to the initial uncertainty around whether this experiment could take place in person given concerns around the circumstances of the COVID pandemic, interoperability was a primary concern given the experiment program may have needed to be operable on a variety of computers. To meet these compatibility requirements, the program was designed to use 33% of the screen height that the program was being run on. This was then used to determine the aspect ratio that the comics would be presented in. While it was possible for this experiment to be conducted on campus thus consistently on the same laptop computer, the aspect ratio of the comics will inevitably be different when the program is run on another computer.

span tasks (Milner, 1971; Vandierendonck et al., 2004; Kessels et al., 2008) and has since been used numerous times as a metric of visuospatial working memory. The task involves participants viewing a spatial sequence whereby, in the original physical version of this test, a grid of blocks is presented to the participant, some of which are then tapped in a specific order. The participant's task is then to replicate the block sequence as accurately as possible. Recently, research interest has turned to the efficacy of computerised versions of this task and has primarily found that accuracy and ability on the physical iteration largely translate to computerised versions (Brunetti et al., 2014; Robinson & Brewer, 2016). Though some research has found some disparity in performance between computerised and physical forward versions of the task (Claessen et al., 2015), computerised versions are now standard practice. In computerised versions, the blocks are displayed on a screen and flash in a specific order that the participant is then required to replicate using the mouse.

All of the experiments in this thesis were implemented using the Python programming language. As part of this work, the author developed a Corsi block test in Python. Demonstration versions of the code used to run the cognitive tests, including the Corsi block test, are available in a GitHub repository: https://github.com/danmatkin/Cognitive-Tests.

2.1.2.4 Design

This experiment used a between-groups design. To this end, participants completed either the experimental dual-task procedure or the single-task control condition, in which there was no secondary task but a time-matched stimulus maintenance period. The experiment's independent variable was whether the participants experienced a comic as interference. The dependent variable was operationalised as the accuracy with which the participant could recall the working memory stimulus.

In line with previous research, to ensure parity between the participants in terms of their working memory ability and account for the individual differences that are often found in working memory capacity, a titration procedure following the suggestion by Logie and Doherty (2016) was implemented. This meant that participants were first required to complete repeated trials of the Corsi block task before completing the experiment, which gradually increased in difficulty by one block to the point that they failed two trials. The point at which they failed two titration trials was considered to be above the participant's visuospatial working memory capacity limit, and the level of dual-task interference they experienced was then set at one block less. This individualised capacity limit was then utilised in the subsequent dual-task trials to ensure that each participant's cognitive load was appropriately tailored to their working memory capacity. For example, a participant who reached a length six block sequence but failed two trials at this level would be assigned a length five block sequence in the dual task procedure. The maximum sequence length that the participant could reach was nine blocks. This method allowed for the standardisation of the cognitive load across participants and consequently enhanced the reliability and validity of the accuracy results by minimising the effect of working memory variability. The same procedure was implemented for the single-task version.

For simplicity, a positional accuracy scoring procedure was used. In this method, the accuracy of a participant's response is determined by the correct positional order of items. For example, if the participant were shown a sequence of ABCD and registered a response of ABDC, this would return 50% accuracy, given that only A and B are in the correct positions. This procedure was used to determine the level of interference that the participant experienced in the dual-task experiment.

2.1.2.5 Procedure

Participants completed the experiment in booths that were designed to be distraction-free. After reading the information sheet and consenting to participate in the study, they were asked to provide their age, gender details and email address[4]. Following this, they were required to complete the visuospatial titration procedure outlined above. This task component initially presented the participants with written instructions on-screen outlining what they were expected to do during the Corsi block task. The participant then completed two practice trials at a sequence length of three. The researcher was present throughout the completion of the practice trials to resolve any issues or confusion about the task's requirements. On completion of the practice trials, the participants were asked to confirm verbally that they understood what they were required to do. After confirming this, they began the titration procedure.

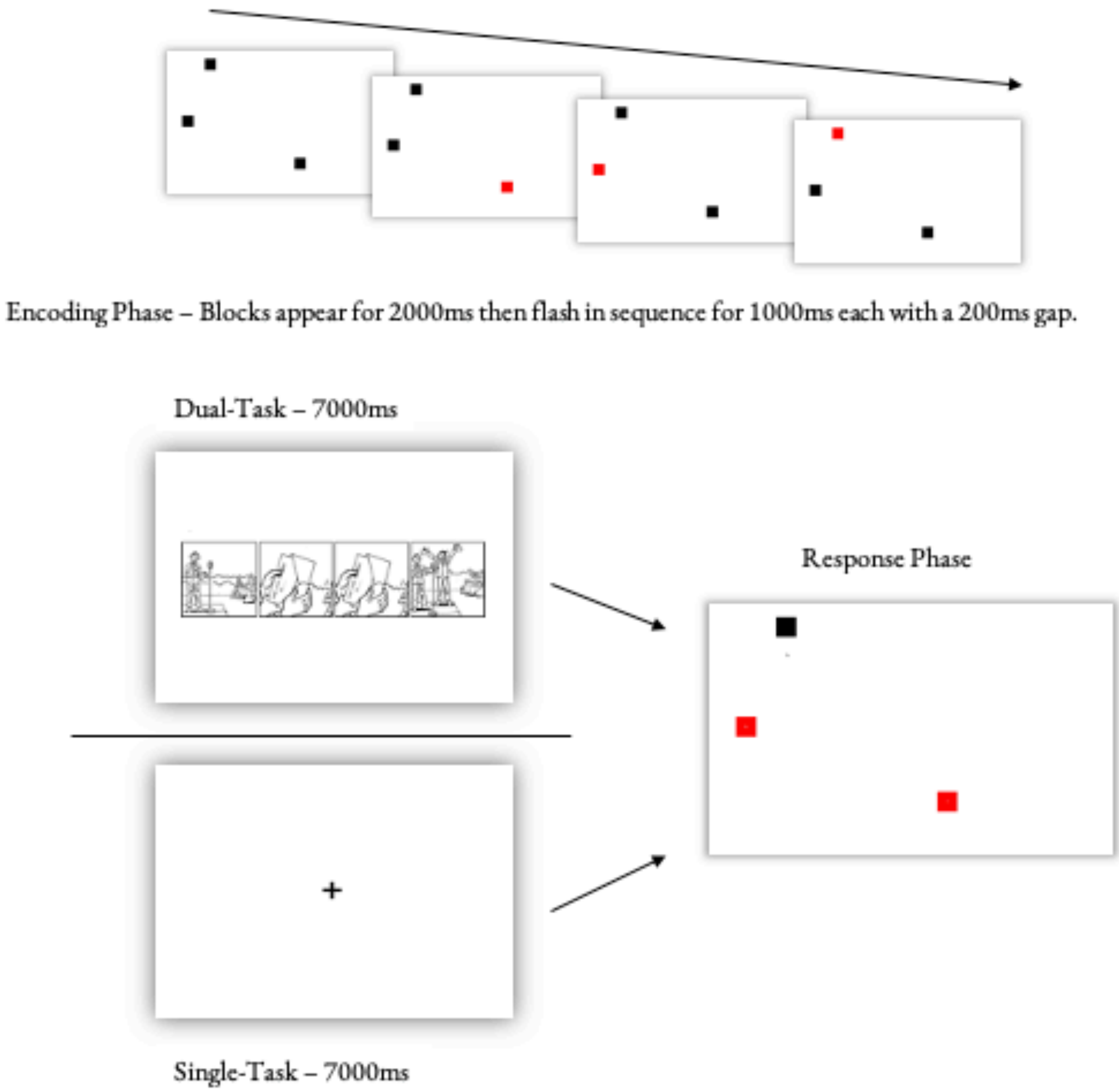
[4] This was used only to email the participants with the debrief materials and was not saved by the program.

2.1.2.6 Corsi Block Test Procedure

The procedure for the actual Corsi block test was as follows. The participants were first presented with an on-screen prompt telling them to 'Get Ready'. This was displayed for 3000ms. At the offset of this instruction, an interstimulus cross was presented for 2000ms to focus their attention on the middle of the screen. At the offset of the cross, the blocks appeared on the screen. These were placed in pseudo-random spaces on the screen. This was considered pseudo-random because the blocks were programmed to be reasonably evenly spaced so as not to encroach on each other. The participant was then given the Corsi block sequence to encode and maintain. In this sequence, the blocks appeared on screen for 2000ms, after which each block flashed red for 1000ms each, with a short gap of 200ms between each flash. This represented the specific order that the participant was required to remember. The block stimulus offset 1000ms after completion of the flashing sequence and was replaced by the presentation of an instruction to 'Repeat the Sequence', which appeared for 2000ms. The blocks then reappeared on the screen and remained there until the participants had clicked the blocks and registered a response. Upon finishing the Corsi block titration procedure, the participants were asked to complete a computerised version of the VLFI. After this, the participants were provided with a 3-minute break. Figure 2 graphically represents the Corsi block trials.

Following the break, participants in the control condition completed one practice trial of the single-task version. After confirming again that they understood the task's requirements, they completed six trials of the single-task version. The procedure of the control dual-task trials was essentially the same as the titration trials but with two major differences. First, participants immediately saw a block sequence at their titrated capacity length. At the offset of the Corsi stimulus, the participants were simply presented with an interstimulus cross for 7000ms. At the offset of the comic, the blocks were returned to the screen, and the participants were required to recall the block sequence.

Figure 2 *Graphical Representation of Experiment 1 Procedure*



Participants in the experimental condition completed the same procedure; however, instead of simply maintaining the block sequence for 7000ms, they were presented with a randomly selected comic strip from the predetermined sets mentioned above. Following the offset of the comic strip, the blocks returned to the screen, allowing the participants to input their responses.

After the participants had registered their responses to the working memory test, they were asked to recall the comic strip and estimate the amount of time that had passed in the comic. It was specifically explained during the practice trial that this question referred to the amount of time that passed within the scene depicted in the comic and not to the actual amount of time that the comic was on screen. The participants were informed that they would be asked this question after every comic. This question was included to promote participant's engagement with the comics content. A text entry box registered responses to this question with the instruction to input a number and a time frame reference, such as "5 m/5 minutes".

To further verify engagement with the comic materials, participants who completed the dual-task procedure were asked to provide written summaries/descriptions of the content of the comics, as well as indicate how many panels they remembered the comic having. Participants entered their responses into a series of text entry boxes at their own pace. Following this phase, the participants then answered six questions which pertained to specific details about each of the comics they had seen in the experiment. These questions were presented sequentially in reverse relative order to the presentation order of the comics. Following this procedure, the participants were verbally debriefed and informed that they would be emailed a copy of the relevant debrief information. The total time required to complete the experiment varied between 25 to 35 minutes.

2.1.3 Results

2.1.3.1 Corsi Block Descriptives and Assumptions

All analysis reported across this project was conducted in R (R Core Team, 2022). The primary variable of analyses in this experiment was the accuracy with which participants were able to recall the Corsi block sequence. Participant's overall accuracy was generated by averaging across each of the six trials. This variable was then used to assess the dual-task cost to visuospatial working memory when processing a comic. Shapiro-Wilk tests of normality indicated that average accuracy in both the experimental ($W = 0.91, p = 0.09$) and the control condition ($W = 0.93, p = 0.15$) met the parametric assumptions of normality. Levene's test for homogeneity of variance similarly confirmed homoscedasticity ($F(1, 38) = 2.21$, p = 0.15). Table 1 displays means and standard deviations for single and dual-task Corsi block accuracy.

Table 1 *Corsi Block Single and Dual-Task Accuracy Descriptives*

Condition	N	M	SD
Dual-Task	20	0.609	0.188
Single-Task	20	0.842	0.108

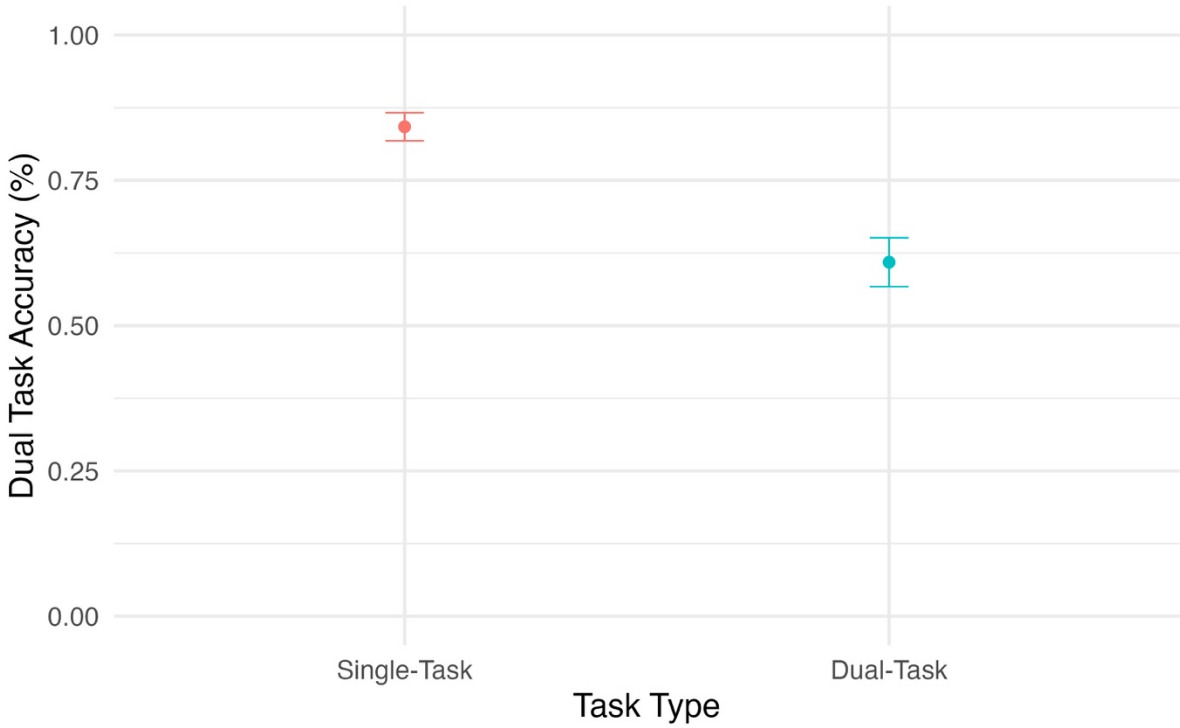
Note. Number of participants (N), Means (M) and Standard Deviations (SD) for both the Corsi block dual-task and single-task conditions.

Although comprehension was not directly tested within the main task, participants demonstrated high levels of accuracy in a subsequent free recall task (M = 79%), indicating that they had generally understood and retained the comic narratives. This was calculated by scoring each response for binary accuracy (I.E presence of the main events or an acceptable description of the comics narrative). A fuller discussion of this check and subsequent limitations is provided in the discussion of Experiment 3.

2.1.3.2 Corsi Block Dual Task Accuracy Analysis

An independent samples t-test indicated that the average difference in Corsi block accuracy between participants in the dual-task condition (*M* = 0.61, *SD* = 0.19) and single-task condition (*M* = 0.84, *SD* = 0.11) was statistically significant $t(38) = 4.81$, p = <.001, 95% *CI* [0.14, 0.33] and this was associated with a very large effect size Cohen's $d = 1.52$, 95% *CI* [0.79, 2.24]. This indicates that the comics had a significant impact on the participant's ability to maintain an individualised visuospatial working memory load. Figure 3 displays the differences in accuracy for the single and dual-task Corsi block conditions.

Figure 3 *Corsi Block Accuracy for Single and Dual-Task Conditions*



2.1.3.3 Corsi Block Dual Task Accuracy by Panel Variation

As outlined in the materials section above, the effect of panel quantity on visuospatial working memory load was then tested to ensure that this manipulation did not significantly affect visuospatial working memory load. First, the residuals of the ANOVA model were assessed for normality using the Kolmogorov-Smirnov test of normality and indicated no significant issues ($D = 0.07$, $p = 0.862$). Additionally, Mauchly's test indicated that there was no violation of the sphericity. The results of the one-way repeated measures ANOVA then revealed that there was no significant effect of comic length on dual-task accuracy $F (2, 38) = 1.373$, $p = 0.266$, with a small effect size ($\eta^2 = 0.0674$). From this, we can then conclude that there was no significant effect of the panel manipulation on visuospatial working memory and in turn this leads to the conclusion that this manipulation did not significantly affect the outcomes of the main dual task vs single task analysis.

2.1.3.4 Visual Language Fluency and Corsi Block Accuracy Correlations

Pearson correlations were conducted to examine the relationship between the participants single and dual-task accuracy with their Visual Language Fluency Index scores (VLFI). The mean VLFI in the experimental group was 8.06 (SD = 5.08). For this group, the analysis revealed a moderate, statistically significant positive correlation between dual-task Corsi block accuracy and VLFI score, $r(18) = .46$, $p = .041$, 95% CI [.02, .75]. This suggests that better performance on the Corsi block dual-task is associated with higher VLFI scores.

The single-task group had a slightly lower mean VLFI score of 6.27 (SD = 6.93), with slightly greater variability than the experimental group. the analysis of this group revealed a small-to-moderate, non-significant positive correlation, $r(18) = .26$, $p = .263$, 95% CI [−.20, .63]. This suggests that, in the control condition, there was no reliable relationship between performance and VLFI. Overall, only the dual-task condition demonstrated a statistically significant association between accuracy and visual language fluency.

2.1.4 Discussion

This experiment aimed to investigate the use of visuospatial working memory resources in processing visual narratives. If processing a comic uses visuospatial working memory, then participants will perform worse on a Corsi block test when they have to process a comic at the same time (dual-task) compared to when they only perform the Corsi block test (single-task). This experiment represents only the third investigation into visuospatial working memory in the context of visual narrative processing and is the first to implement a titration procedure and individually tailor the working memory loads to participants. By accounting for individual differences in working, this approach reduces variability and ensures that the working memory demands are both equitable and precise across participants. Consequently, findings offer greater rigour and validity in comparison to previous research in the field.

The results of the experiment demonstrate significant activation of visuospatial working memory in visual narrative processing. This is derived from the difference in visuospatial working memory accuracy in the dual-task and single-task conditions, supporting the hypothesis. As can be seen in figure 3 above, the participants' ability to maintain a visuospatial working memory sequence while also having to process a comic in the dual-task condition was significantly reduced in comparison to the participants in the single-task condition. This then indicates that maintenance of a visuospatial sequence was interfered with by the requirement to process a comic and thus indicating that visuospatial working memory plays a critical role in processing comics.

In relation to existing research, the findings of the experiment provide support for the position of Magliano et al. (2016) which indicates involvement of visuospatial working memory in visual narrative processing. Further to this, the current experiment arguably provides enhanced evidence by assessing the direct impact to visuospatial working memory maintenance of the requirement to process a comic. Conversely, the findings contradict those of Zhao et al. (2024), which suggest that visual narrative processing relies less on visuospatial working memory. This contradiction is likely due to differences in the metrics used between the two experiments. Zhao et al. (2024) relied on reading times, which are limited in their ability to isolate specific cognitive mechanisms, as this metric reflects a broad combination of processes rather than specific contributions of working memory. In contrast, the current experiment used a dual-task paradigm, which directly measures the involvement of visuospatial working memory in processing visual narratives. This methodological difference highlights the importance of precise tools to assess specific cognitive resources involved in complex tasks.

While the findings of this experiment provide reasonably robust evidence for the use of visuospatial working memory in comics processing they are limited by a number of factors. Firstly, though the findings indicate the use of this working memory component, they do little to indicate how this resource is actually deployed during visual narrative processing. Secondly, the findings are also limited in that currently, the use of visuospatial working memory has been demonstrated only by comparing participants' performance on two tasks to participants performance on a single task. A feasible alternative explanation for the findings of this experiment is the idea that the dual-task cost demonstrated was simply due to having to perform two tasks simultaneously. Consequently, further research using tasks which tap different components of working memory are required to rule out this possibility.

Overall, the results of this experiment indicate that visuospatial working memory contributes significantly to visual narrative processing. The results support those of Magliano et al., (2016) but contradict those of Zhao et al., (2024). In relation to the latter, this discrepancy is explained by the metrics used. While this is a relatively unsurprising finding, this experiment is the first to have used direct assessment of the visuospatial working memory contribution to processing static image sequences. Experiment 2 then turns towards the question of the differential use of visual and spatial working memory during visual narrative processing.

2.2 Experiment 2 - Visual Working Memory in Comics Processing

2.2.1 Introduction

The results of Experiment 1 indicate that visuospatial working memory (VSWM) resources are activated during image sequence processing. This conclusion is supported by the observed interference effect when participants were required to parse a comic strip while maintaining a Corsi block sequence. However, previous research has suggested that visuospatial working memory can be divided into distinct visual and spatial sub-components (Klauer & Zhao, 2004; Pickering, 2001; Logie & Pearson, 1997). For example, the visual component is thought to be related to the maintenance of shape or colour, whereas the spatial component is considered responsible for maintaining sequence order (Logie & Pearson, 1997). While both components are likely critical to image sequence processing, no research has yet explored the specific use of visual and spatial components of working memory in image sequences. If these processes are differentially affected by the requirement to process a comic, this could suggest differential utility of these components in visual narrative processing.

Various lines of research support the distinction between visual and spatial components of visuospatial working memory. A key demonstration of this comes from the dual-task paradigm. Della Sala et al. (1999) paired the Corsi block test with the visual pattern test in a dual-task experiment alternating the two as the primary and secondary tasks. The visual pattern test presents matrices of black and white squares, which gradually increase in dimensions. As in the Corsi block test, the typical use of this sees the participant memorise a visual pattern matrix which they are then asked to recall. The results of this indicated that the Corsi block test and the visual pattern test do not produce significant dual-task costs on one another and as such, this provides dissociative evidence for the suggestion that the visual and spatial components of VSWM have some independence of function. Further evidence for this distinction comes from findings which demonstrate, for example, that the successful completion of the Corsi block test is dependent to some degree on oculomotor movements, whereas, in contrast, the visual pattern test is not (Pearson et al., 2014). These findings point towards distinct underlying mechanisms related to visual and spatial working memory.

The Corsi and visual pattern tests have also been used to demonstrate developmental differences between visual and spatial working memory. One study reports that in 5 - 6-year-olds, there was no significant difference between visual and spatial working memory task ability. However, a significant difference between visual and spatial ability was present in 11 – 12-year-olds, with this age group able to complete visual pattern recall significantly more accurately than recalling spatial sequence (Logie & Pearson, 1997). This difference further supports the distinction, demonstrating that these two capacities appear to develop along different trajectories.

Lastly, this distinction is also supported by findings related to the abilities that visual and spatial components of working memory underpin. Much research has demonstrated that visuospatial working memory is associated with ability in mathematics (Raghubar et al., 2010; Fanari et al., 2019). Interestingly, this research has also recently moved beyond the relationship of VSWM to pure mathematical performance and has begun to explore the role that visuospatial working memory plays in mathematics-related self-efficacy (Cuder et al., 2023; živković et al., 2023). In evaluations of the relationship between mathematics ability and VSWM, research indicates that spatial components of visuospatial working memory relate to difficulty in learning maths (Mamarella et al., 2018). This further underscores the distinction between visual and spatial components in VSWM and highlights their unique contributions to different cognitive abilities.

In relation to the relative use of visual and spatial working memory in processing sequences of images, it was expected that image sequence processing would more strongly rely on the use of spatial working memory. This is because the semantic relationships between panels of a comic are a product of the spatial relationship and composition of what is depicted in the panels. Comics require that the reader first understands the spatial arrangement of the panels to communicate information across panels. This is related to the concept of closure outlined by McCloud (1993), in which readers must first comprehend the information in panels and then connect the information between them. This process requires the reader to track and understand the progression of time and action to comprehend the depicted content. Furthermore, given that actions in comics are conducted across space, comics seem likely to engage spatial working memory in perceiving and tracking this dimension of the content.

Given that visual working memory is suggested to be related to more basic perceptual features of processing, such as recognising and perceiving visual form, the requirement to use these resources is likely lower than the requirement to use spatial working memory. This relates to the ease with which visual content can be processed (Potter et al., 2014; Hagmann & Cohn, 2016). In other words, understanding and recognising what is in the panels seems likely to be less resource-intensive than the process of comprehending the entirety of a sequence, which, as described, seems intuitively likely to require the use of spatial working memory. If image sequence processing requires spatial resources to a greater extent than visual resources, then the interference caused by a spatial working memory task will be greater than the interference caused by a visual working memory task.

This experiment's objective was to further investigate the contribution of visuospatial working memory to visual narrative processing by exploring the roles of visual and spatial sub-components. In the first instance, it is important to validate the use of visual working memory in visual narrative processing. Secondly, this experiment aimed to investigate the relative contributions of the visual and spatial components of working memory. To explore this, the dual-task condition accuracy scores of participants who completed the Corsi block test will be compared with those who completed the visual pattern test in this experiment. This comparison is made possible using the working memory titration procedure (Doherty & Logie, 2016). Because of the titration procedure, the participants were completing the dual-task experiment at the limit of their capacity. This then allowed for a comparison of accuracy in these conditions that would determine the relative use of these components.

In line with the findings of Experiment 1, it was hypothesised that the requirement to process a comic while maintaining a visual working memory stimulus would generate significant dual-task costs, meaning that visual pattern test dual-task accuracy would be significantly lower than visual pattern test single-task accuracy. While no findings explicitly support this suggestion, as outlined above, it was considered that spatial working memory would be more significantly affected by the requirement to process a comic than visual working memory. This is based on the characterisations of visual working memory acting as a basic perceptual recognition mechanism for shapes and the spatial working memory component acting as a store for sequence-based information (Logie & Pearson, 1997; Klauer & Zhao, 2004). In relation to the relative contributions made by visual as opposed to spatial working memory, spatial processing in visual narratives is considered to encompass both the spatial processing of the scene as well as likely being implicated to a greater degree in processing the narrative on the basis that comics communicate narrative over space. Thus, it was hypothesised that spatial working memory would be significantly more impacted than visual working memory during visual narrative processing.

2.2.2 Method

The method used in this experiment was the same as that used in Experiment 1, except that the Corsi block test was replaced with the visual pattern test.

2.2.2.1 Participants

40 participants (30 Female) completed this experiment with an average age of 20 (*SD* = 2.41). All participants were rewarded with course credits on an undergraduate Psychology course for taking part in the experiment. The same participant pool was used for this experiment as the first, but importantly, no participant took part in more than one component of the study. As in

Experiment 1, all participants had normal to corrected vision and were naïve to the purposes and intentions of the study. The University's Research Ethics Committee approved the experiment in line with the ethical procedures.

2.2.2.2 Design

This experiment used a between-groups design with participants completing either dual- or single-task conditions. The independent variable in this experiment was whether the participants completed either the dual-task or the control condition of the visual pattern test and the dependent variable was the accuracy with which participants completed the task. To present a comparison of the differential activation of spatial versus visual working memory during comics processing, the data from participants who completed the dual-task version of Experiment 1 is included in the analysis used in this experiment. Consequently, for the analysis presented in section 2.3.2.4, half of the participants took part in Experiment 1.

2.2.2.3 Materials

The comics materials, as well as the method of presentation, were also the same as in Experiment 1. Importantly, the panel variant manipulation described in Experiment 1 was also implemented in this experiment. As before, this data will be analysed and presented to ensure that this did not significantly affect the main analysis content but is not considered further.

The visual pattern test was implemented to assess the impact of the interference of visual narrative processing on visual working memory. This followed the typical procedure established in previous research (Della-Sala et al., 1999; Logie & Pearson, 1997; Brown et al., 2006) and was implemented using a version created in Python by the author. A demo version of the VPT created by the researcher is available in this GitHub repository: https://github.com/danmatkin/Cognitive-Tests. .

Given that higher levels of comics reading experience have been shown to influence the way that they are comprehended (Hagmann & Cohn, 2016), the Visual Language Fluency Index or VLFI (Cohn et al., 2012) was used, as in the previous experiment, to measure the participant's exposure to visual narratives.

The visual pattern test used in this study followed the procedure established by previous research (Della Sala et al., 1999). In this test, participants are presented with matrices of black-and-white squares and are subsequently asked to recall the pattern on a blank matrix.

2.2.2.4 Procedure

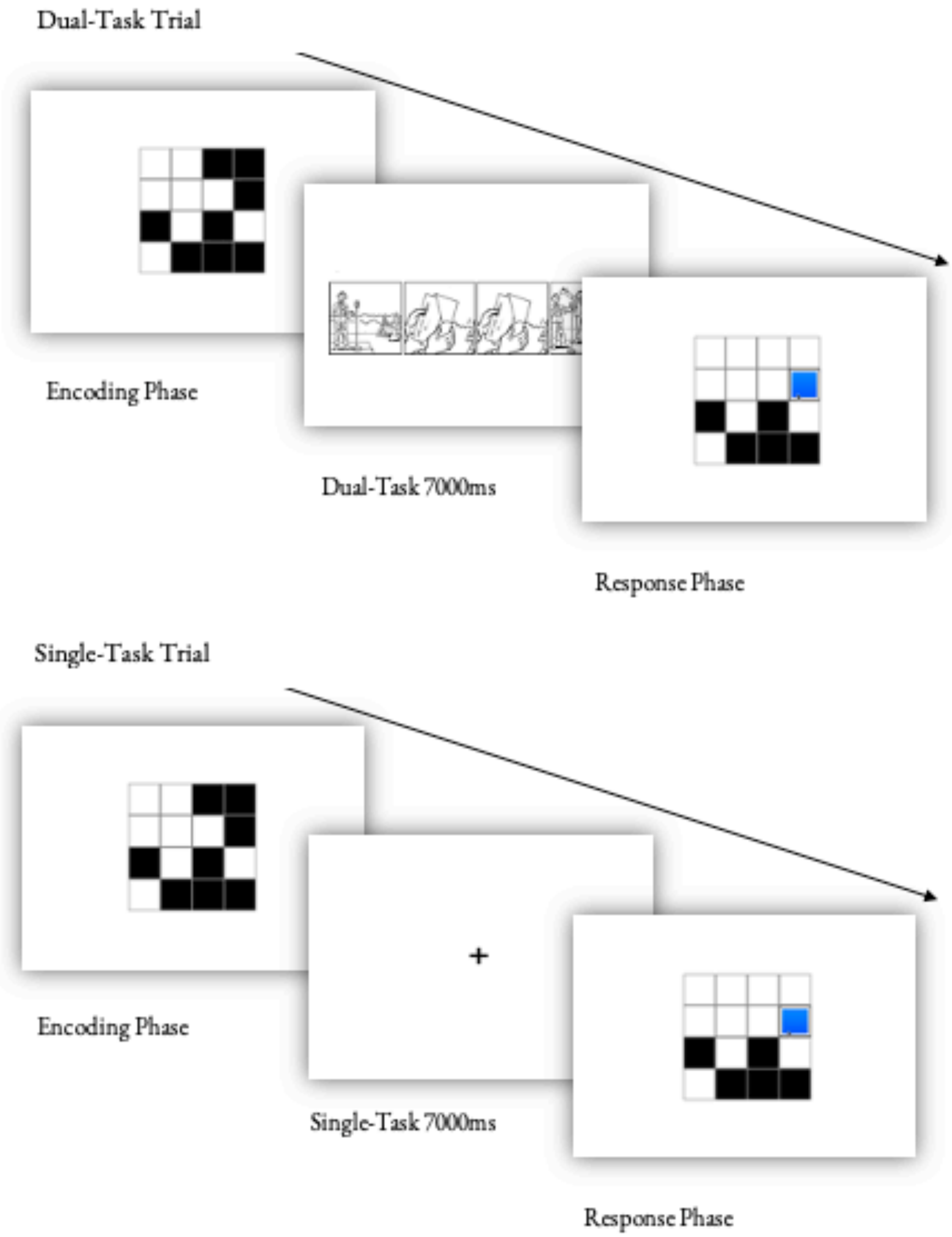
Participants completed the experiment in booths that were designed to be distraction-free. At the outset of the experiment, participants were asked to provide their age, gender details and an email address. The presentation of the information and consent checks followed this. Following this, the participants completed the titration phase of the visual pattern test. In this phase, participants were initially presented with on-screen instructions about how to complete the visual pattern test, followed by two practice trials. When these were complete, further on-screen instructions were presented, indicating that the procedure was about to begin properly. Visual pattern test trials began with an on-screen instruction for the participant to 'Get Ready' for 3000ms. An interstimulus cross was then presented for 2000ms. At the offset of the cross, the matrix appeared. Half of the cells of the matrix were filled in black, and half were left white. The matrix then remained on screen for 2000ms. This duration was consistent throughout all visual pattern test trials regardless of matrix size. During the titration trials, at the offset of the matrix, the participants were presented with an on-screen instruction to "Repeat the Pattern" displayed for 2000ms until it was replaced with a blank, clickable matrix onto which the participants were asked to recall the previously seen matrix pattern.

The practice trials of the titration task began with a matrix size of 4. The actual trials then began with a matrix size of 6. As in the previous experiment, participants were required to complete at least two successful trials at each level of the visual pattern test. However, if participants failed a trial at any level, they were required to register two more successful trials at that level. Two matrix blocks were added on successfully completing any given visual pattern matrix level. At the point that the participant failed two trials at the same level, the titration phase finished. The maximum level was 24 blocks (12-block matrix pattern).

After the titration phase, the participants completed the Visual Language Fluency Index and were given a three-minute break. After this, the participants were given further instructions for the following part of the experiment. Depending on the condition, this was either the dual-task or the single-task procedure. The participants then completed a practice trial of either the dual-task or the single-task. These trials ran as the titration trials did, with two exceptions. First, they were always at the level of working memory load set by the titration phase, and second, instead of an on-screen instruction to 'Repeat the Pattern," the participants were presented with a comic in the dual-task condition or an interstimulus cross in the single-task condition. The comic or cross remained on screen for 7000ms.

In the dual-task condition, after the participant's matrix response had been registered, the participants were asked the time question ("Approximately how much time passed in the comic"). After completing the dual-task trials, the participants completed the comics questions phase. Finally, participants were verbally debriefed and informed that the information, debrief, and consent forms had been emailed to them upon completing the experiment and closing the program. Figure 4 despicts a visual illustration of the VPT trials.

Figure 4 *Graphical Representation of Visual Pattern Task Procedure Used in Experiment 2*



2.2.3 Results

2.3.2.1 Visual Pattern Test Descriptives and Assumptions

The same preparation procedures were applied to the visual pattern matrix responses: average accuracy for participants was computed from the six dual-task single-task trials they completed. This dependent variable is expressed as a percentage. Shapiro-Wilk tests of normality indicated that while the distribution of accuracy in the dual-task condition could be approximated to normal ($W = 0.98, p = 0.86$). However, the single-task condition was not normally distributed ($W = 0.88, p = 0.01$) and was significantly negatively skewed (-0.80). Levene's test for homogeneity of variance indicated that the data met the assumption of homoscedasticity $F(1, 38) = 2.83, p = 0.10$.

Table 2 *Visual Pattern Test Single and Dual Task Accuracy Descriptives*

Condition	N	M	SD
Dual-Task	20	0.759	0.110
Single-Task	20	0.935	0.063

Note. Number of Participants (N), Means (M) and Standard Deviations (SD) for both the visual pattern dual-task and single-task conditions.

As in Experiment 1, participants comprehension was checked through the use of the free recall task they completed at the end of the experiment procedure. The participants in this experiment also demonstrated high levels of accuracy in a subsequent free recall task (M = 79%), indicating that they had generally understood and retained the comic narratives.

2.3.2.2 Visual Pattern Test Dual Task Accuracy Analysis

Log, square root and reciprocal transformations were applied to the VPT single-task condition to reduce the normality violation. However, none of these effectively reduced the Shapiro-Wilk statistic to a level that was not significant (all *p*s < 0.05). Because of this, a non-parametric approach was selected to compare the VPT dual-task condition against the control condition. A Wilcoxon Rank Sum test revealed that the difference in accuracy with which participants were able to maintain a matrix pattern in the dual-task condition (*Med* = 0.75) and the control condition (*Med* = 0.93) was statistically significant *W* = 363, *p* = <.001, with a large associated effect size r = .70. Consequently, the hypothesis that image sequence processing requires the use of visual working memory was supported. Figure 5 depicts this significant difference.

Figure 5 *Visual Pattern Test Accuracy for Single and Dual-Task Conditions*

2.3.2.3 Visual Pattern Test Panel Variant Analysis

As in Experiment 1, VPT dual-task accuracy was compared by panel variants to ensure that this had not affected the contribution of visual working memory. Normality (D = 0.09, p = 0.640) and sphericity (W = 0.98, p = 0.831) of this model were first assessed and indicated no significant violation. A one-way, repeated measures ANOVA revealed no statistically significant effect of panel variation on visual pattern task dual-task accuracy $F(2, 38)$ = 0.350, p = 0.707, with a negligible effect size (η^2 = 0.0181).

2.3.2.4 Visual Language Fluency and Visual Pattern Test Accuracy

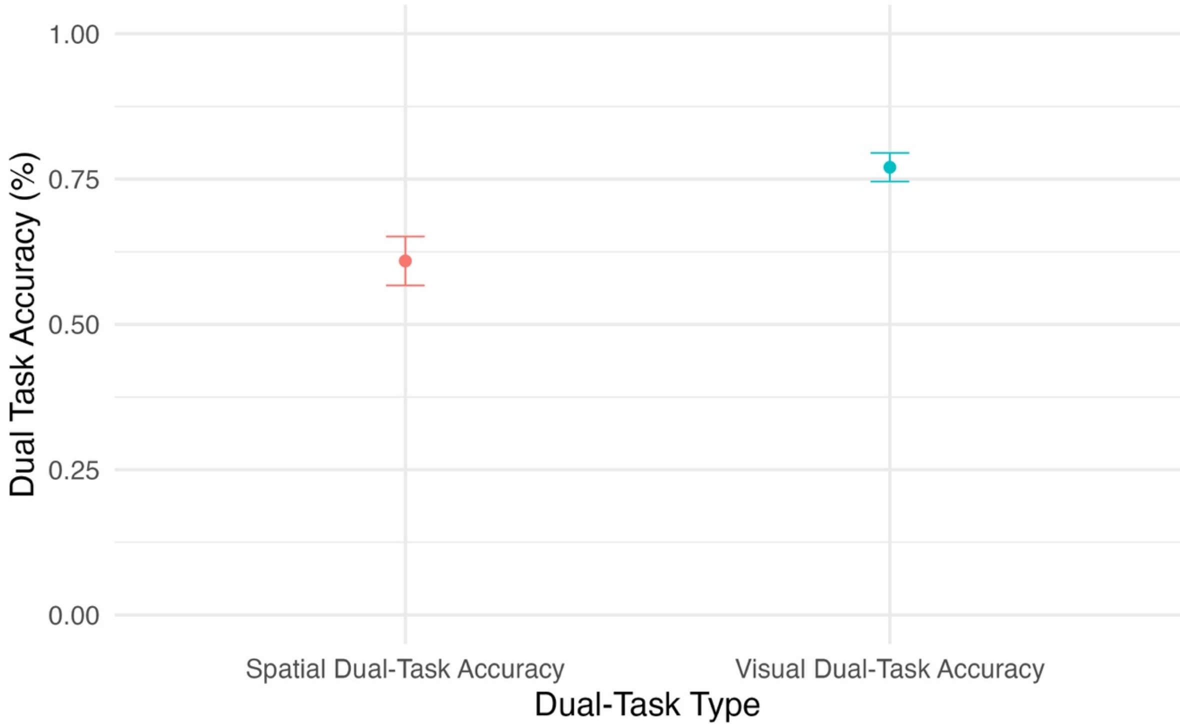
As in experiment 1, Pearson correlations were conducted to examine the relationship between single and dual-task accuracy with VLFI scores. The mean VLFI score in the dual-task VPT group was 6.06 (SD = 4.42). While there was a moderate positive correlation between participant's accuracy and their VLFI scores in the dual-task condition, this did not reach statistical significance, $r(18)$ = .36, p = .116, 95% CI [−.10, .69] This suggests a possible positive association, although the evidence was not strong enough to rule out the null hypothesis.

In the single-task condition, the participants had a slightly lower mean VLFI score of 5.51 (SD = 6.55). The correlation between accuracy and VLFI in this group was small and non-significant, $r(18)$ = .12, p = .610, 95% CI [−.34, .54]. Overall, while the dual-task condition showed a trend toward a positive association between accuracy and VLFI, neither condition demonstrated a statistically reliable relationship.

2.3.2.5 Contributions of Visual and Spatial Working Memory

Given that experience of visual narratives has been demonstrated to affect processing efficiency, it was important to determine that there were no significant differences between the participants in the spatial working memory dual-task group and the visual working memory dual-task group. The Shapiro-Wilk test of normality indicated that the VLFI scores of both the Corsi block group ($W = 0.880$, $p = .018$) and the visual pattern test group ($W = 0.832$, $p = .003$) were not normally distributed. Levene's test for homogeneity of variance showed no significant difference in variance between groups for the VLFI scores, $F(1, 38) = 0.52$, $p = .477$. To test for significant differences, a Mann-Whitney U test was conducted to compare VLFI scores between the spatial working memory dual-task group and the visual working memory dual-task group. This detected no significant difference between the groups, ($W = 254.5$, $p = .144$) suggesting no strong evidence for different levels of comics experience between the groups which completed the visual and spatial working memory tasks. These analyses confirm that the spatial and visual working memory dual-task groups were similar in their baseline characteristics relevant to visual narrative processing and, therefore, support the validity of subsequent comparisons.

Figure 6 *Comparison of Visual Pattern Test and Corsi Block Accuracy*



Given that the analysis in the previous experiment demonstrates that Corsi block dual-task accuracy was normally distributed, and the analysis above in the current experiment indicates the same for VPT dual-task accuracy, an independent measures t-test was used to evaluate whether there was any statistically significant difference between visual and spatial dual-task costs. This revealed a statistically significant difference between the groups $t(38) = -3.31$, $p = .002$, 95% *CI* [-0.26, -0.06]. Participants who completed the visual pattern test dual-task ($M = 0.77$, $SD = 0.05$) significantly outperformed participants in the Corsi block dual-task condition ($M = 0.61$, $SD = 0.05$), with a mean difference of approximately 16% ($MD = 0.16$). This was associated with a large effect size, Cohen's $d = 1.05$ 95% *CI* [-1.70, -0.38]. The difference in dual-cost to spatial working memory in comparison to visual working memory is depicted in Figure 6. This finding then tentatively indicates that spatial working memory underpins visual narrative processing to a greater degree than visual working memory.

2.2.4 Discussion

Consistent with the previous experiment, it was hypothesised that visual working memory would be activated during visual narrative processing. The findings of the experiment supported this hypothesis. Furthermore, an exploratory hypothesis regarding the differential activation of visual and spatial working memory was tested. The results revealed a significant difference, indicating that visual narrative processing engages spatial working memory more extensively than visual working memory.

These findings reinforce the notion that visual working memory plays a role in visual narrative processing. Specifically, the observed detriment to visual working memory maintenance when simultaneously processing a comic strip highlights this involvement. This aligns with the conclusions of Magliano et al. (2016), which suggest visuospatial working memory activation during such tasks. Conversely, the results contradict Zhao et al. (2024), who argued that visuospatial working memory is not implicated in visual narrative processing.

These findings gain additional significance when considered in conjunction with the findings of Experiment 1. The analysis comparing the effects of comics processing on spatial and visual working memory reveals that spatial working memory appears to be more strongly impacted. This could indicate that spatial working memory plays a more significant role in processing visual narratives. Theoretically, this aligns with the fundamental role that the spatial and sequential relationships between discrete panel units play in conveying narratives through static images. The greater reliance on spatial working memory underscores the idea that the comprehension of visual narratives relies more on processing and integrating spatial relationships. However, though these results provide preliminary support for the predominant role of spatial working memory in comics processing, it is important to acknowledge methodological limitations that may constrain the interpretation of these findings.

While the findings of this experiment are compelling, they should be interpreted with caution due to a number of limitations. Similar to the previous experiment, the results provide limited insight into how working memory is engaged during narrative processing, primarily because the study did not involve any direct experimental manipulation of narrative elements. As such, the conclusion that narrative comprehension in comics is predominantly reliant on spatial working memory, although appealing, may overstate the implications of the findings. Future research should address this gap by incorporating controlled manipulations of narrative structure to more precisely determine the role of working memory components in narrative processing.

Further, some consideration should be given to the characteristics of the Corsi block and visual pattern tasks, arguably impacting the interpretation of the performance comparison between these tasks. Corsi block sequences were significantly longer to complete than the visual pattern test. This was because when Corsi sequences increased, the trials' length also increased. In contrast, the visual pattern test trials were fixed at 2 seconds per trial, regardless of the test level.

This necessary but fundamental difference in task design seems likely to have affected the cognitive demands placed on the participants. For instance, the increased trial duration in the Corsi block task seems likely to have led to greater opportunities for forgetting than in the visual pattern test. This is supported by the basic premise of the Time-Based Resource-Sharing theory of working memory (Barouillet et al., 2004; Vergauwe et al., 2021), which emphasises that strict time constraints act upon working memory whereby held information decays rapidly if it is not actively refreshed using attention.

As arguably demonstrated by this experiment, the separation of visual and spatial working memory capacities and the application of the visual pattern test to demonstrate mechanisms in visual narrative processing holds the potential for interesting findings. However, the differential time profiles of each task highlight a critical methodological barrier in comparing their performance outcomes and justifying their interpretation. Addressing this methodological consideration will enable future research to provide more nuanced and accurate demonstrations of the interplay between visual and spatial working memory in visual narrative processing.

The differences in time profiles, cognitive demands in relation to span, and performance outcomes between the Corsi block task and the visual pattern test highlight significant methodological disparities that complicate meaningful comparisons. These inconsistencies introduce a risk of confounding factors, which could potentially compromise the validity of the findings. To ensure analytical consistency and maintain comparability across experiments, the visual pattern test was excluded from subsequent studies in order to prioritise methodological rigour and minimise the potential for bias in future analyses.

Overall, this line of research offers significant potential in advancing our understanding of the cognitive mechanisms underlying visual narrative processing in relation to the roles of visual and spatial working memory components. The current findings indicate that spatial working memory may be more critical than visual working memory in comics processing. This could be taken as a reflection of the importance of sequential and spatial relationships in how narratives are constructed across sequential images. However, the current experiment is limited in the conclusions that can reasonably be drawn from it. Consequently, the most appropriate conclusion to be drawn from the current experiment is that the findings serve to confirm the results of Experiment 1 in the indication that visuospatial working memory is implicated in visual narrative processing. The first two experiments in this thesis demonstrate that visuospatial working memory underpins visual narrative processing. The final experiment in this chapter then turns to investigate verbal working memory in visual narrative processing.

2.3 Experiment 3 - Verbal Working Memory in Comics Processing

2.3.1 Introduction

Experiments 1 and 2 used the dual-task paradigm to investigate the contribution of visuospatial working memory to visual narrative processing. Experiment 1 indicates a clear use of visuospatial working memory in visual narrative processing. While the findings of Experiment 2 serve to outline that spatial working memory may be used during visual narrative processing to a greater degree than visual working memory, this interpretation should be considered with caution, given potential issues related to the equivalence of the tests. In line with Magliano et al.'s (2016) investigation into both of the multicomponent subsystems, the aim of Experiment 3 was to explore the use of verbal working memory in visual narrative processing.

Currently, only Magliano et al. (2016) have explored the role of verbal working memory in visual narratives. Although their findings suggest that verbal working memory does contribute to this process, it is argued that this conclusion is somewhat counterintuitive and may have been influenced by methodological limitations. As discussed in the introduction to Experiment 1, the experiments presented in this chapter incorporate two key methodological improvements to address the potential shortcomings of Magliano et al. (2016).

First, these experiments place emphasis on the specific working memory components involved in visual narrative processing as the primary task of the dual-task setup. This contrasts with previous experiments' focus on reading times and their use in establishing the contribution of working memory components. Secondly, and perhaps most importantly, the incorporation of working memory titration procedures represents a substantial methodological improvement. This enhancement serves to mitigate the influence of individual differences, thereby increasing the validity of these experiments in identifying the distinct contributions of different working memory components to visual narrative processing.

The focus on inference in the study by Magliano et al. (2016) is important to consider here. Their experiments were designed to emphasise inference generation by omitting bridging events in their visual narratives. This compelled their participants to generate these inferential connections to understand the narratives. In this regard, their finding that verbal working memory is implicated in visual narrative processing arguably finds some support in previous research. One previous study has explored the idea that verbal and visuospatial working memory have specific roles in narrative processing with two experiments (Friedman & Miyake, 2000). Participants in their study read texts that were manipulated to vary in their spatial and causal complexity, and they then responded to probes related to these two dimensions throughout reading. These two components of the text were found to be correlated with the participant's visuospatial and verbal working memory abilities, respectively. These results suggest that visuospatial working memory supports spatial understanding in a text, whereas verbal working memory is purported to support understanding of causality. By extension, this would also include the concept of inferential connection. Though these experiments were conducted using prose, the findings of Magliano et al. (2016) align with this concept and support the suggestion that verbal working memory is implicated in inference during visual narrative processing in a similar fashion to text understanding reported by Friedman and Miyake (2000).

Magliano et al. (2016) justify the exploration of verbal working memory in visual narrative processing on the basis of visual language theory (Cohn, 2013), which argues that the cognitive mechanisms which underpin visual narrative processing share common features with linguistic processing. Research has demonstrated, for example, that grammatical sequencing principles govern the way that comics are both constructed and understood (Cohn et al., 2012; 2014). This conceptualisation of verbal working memory contains an implicit assumption that its use is fundamentally linked to principles of linguistic processing. Verbal working memory, as part of the multi-component model (Baddeley, 2000), primarily deals with the temporary storage and manipulation of verbal and auditory information. This component is then considered to handle phonological input. Though there is still debate about verbal working memory use and whether it holds responsibility for linguistic processing, it is clear that this component is related to traditional language processing from the perspective that this system handles, at the very least, basic inputs for linguistic processing.

In contrast, visual narratives' basic input is visuospatial rather than phonological. As demonstrated in the previous two experiments of this chapter, processing visuospatial narratives engages visuospatial working memory. This aligns with the concept of visuospatial working memory as set out by the multicomponent working memory model (Baddeley, 2000). Tasks such as maintaining and manipulating images, understanding spatial relations between visual elements, and visual recognition are all cognitive processes that are self-evidently critical to visual narrative processing.

The reliance on verbal working memory to support visual linguistic processing posited by Magliano et al. (2016) arguably overlooks the modality-specific nature of cognitive processing and potentially draws unwarranted conclusions about the requirement for verbal working memory in comics processing. Though visual narratives may share broad conceptual features of linguistic processing, such as reliance on sequence, grammatical principles, and even a form of syntax, it seems much more likely that processing a static visual narrative would rely predominantly, if not entirely, on visuospatial processing capacity.

Overall, the use of verbal working memory to support visual narrative processing seems somewhat conceptually unsound. While visual language theory highlights the similarities between visual narrative processing and verbal language processing (Cohn et al., 2014; 2016), this does not necessarily extend to the cognitive mechanisms of the working memory component involved in processing them. The contrast in base input modality, where verbal working memory handles phonological input and visuospatial handles visual input, likely plays a critical role in determining which working memory components are engaged to process these components. As such, using the same titrated dual-task paradigm as the previous two, this experiment aimed to directly assess verbal working memory's contribution to visual narrative processing. It was hypothesised that if verbal working memory is required by visual narrative processing, the requirement to process a comic while maintaining a verbal working memory stimulus would result in significant dual-task costs. In other words, if verbal working memory is implicated in visual narrative processing, verbal working memory dual-task accuracy will be significantly lower than verbal working memory single-task accuracy.

The secondary aim of this experiment was to compare the contributions of both verbal and visuospatial working memory. In this regard, it was considered that visuospatial working memory would contribute to visual narrative processing much more than verbal working memory. To this end, it was hypothesised that visuospatial working memory dual-task accuracy would be significantly lower than verbal working memory dual-task accuracy.

2.3.2 Method

The method used in this experiment matched the method of the first and second, with the exception that the digit span task was used to occupy verbal working memory.

2.3.2.1 Participants

40 participants (27 Female) completed this study with an average age of 20 (*SD* = 3.55). Participants were rewarded with course credits on an undergraduate Psychology course for completing the study. As in Experiment 1, all participants had normal-to-corrected vision and were naïve to the purposes and intentions of the study. The experiment was approved by the University's Research Ethics Committee and was in line with ethical procedures.

2.3.2.2 Design

This experiment used a between-groups design with participants completing either the dual- or single-task conditions. The independent variable in this experiment was whether the participants completed either the dual-task or the control condition of the digit span task and the dependent variable was the accuracy with which participants completed the task.

2.3.2.3 Materials

The comic materials and presentation method were also the same as in Experiments 1 and 2. Importantly, the panel variant manipulation described in Experiment 1 was also implemented in this experiment. As before, this data will be analysed and presented to ensure that this manipulation did not significantly affect the main analysis content but is not considered further. This experiment also used the Visual Language Fluency Index (VLFI) again to measure the participants' experience with comics.

To assess the extent to which verbal working memory is implicated in visual narrative processing, the dual-task paradigm was employed in the same way as in Experiments 1 and 2, using the digit span task as the working memory stimulus. This was again implemented using a version of this test created in Python by the researcher. The maximum sequence that the participant can accurately recall is considered to be their verbal working memory capacity. The code to run this test as a demo version is available to view in this GitHub repository: https://github.com/danmatkin/Cognitive-Tests.

The digit span test is used variously in IQ measures (Wilde et al., 2004) as well as in experimental studies of verbal working memory capacity (Olsthoorn et al., 2014; Hintz et al., 2024; Darling et al., 2020). It is designed to tap only the phonological component of working memory. In the task, participants are presented with sequences of digits and are then required to recall them in the same order in which they were presented. Where Magliano et al., (2016) used word sequences to occupy their participants' verbal working memory, this experiment employed a digit span sequence aligning established protocols of the literature. Digits were chosen because words inherently carry semantic content, which is automatically activated during processing. This not only facilitates memory for words but also engages cognitive processes beyond the intended phonological suppression of the digit span task and articulatory suppression.

The same issue arises with letters. Individual letters, as semantically charged graphemes can easily be transformed into words, enhancing memorability and reducing the phonological load required to retain them. For example, the letter string dpwagt can be interpolated with vowels to form “dope waget”, a more easily recallable and less phonologically demanding structure. Even when letter sequences do not directly form recognisable words, they can still be grouped or modified in ways that improve recall. In contrast, digits are much less likely to be manipulated in the same way. Unlike letters, digits remain discrete units that cannot be blended into more memorable sequences. At most, each digit requires one or two syllables to articulate, but they cannot be shortened, grouped, or transformed in a way that enhances retention. This makes digits a more effective choice for ensuring that the task primarily taxes phonological working memory without the confounding effects of semantic processing.

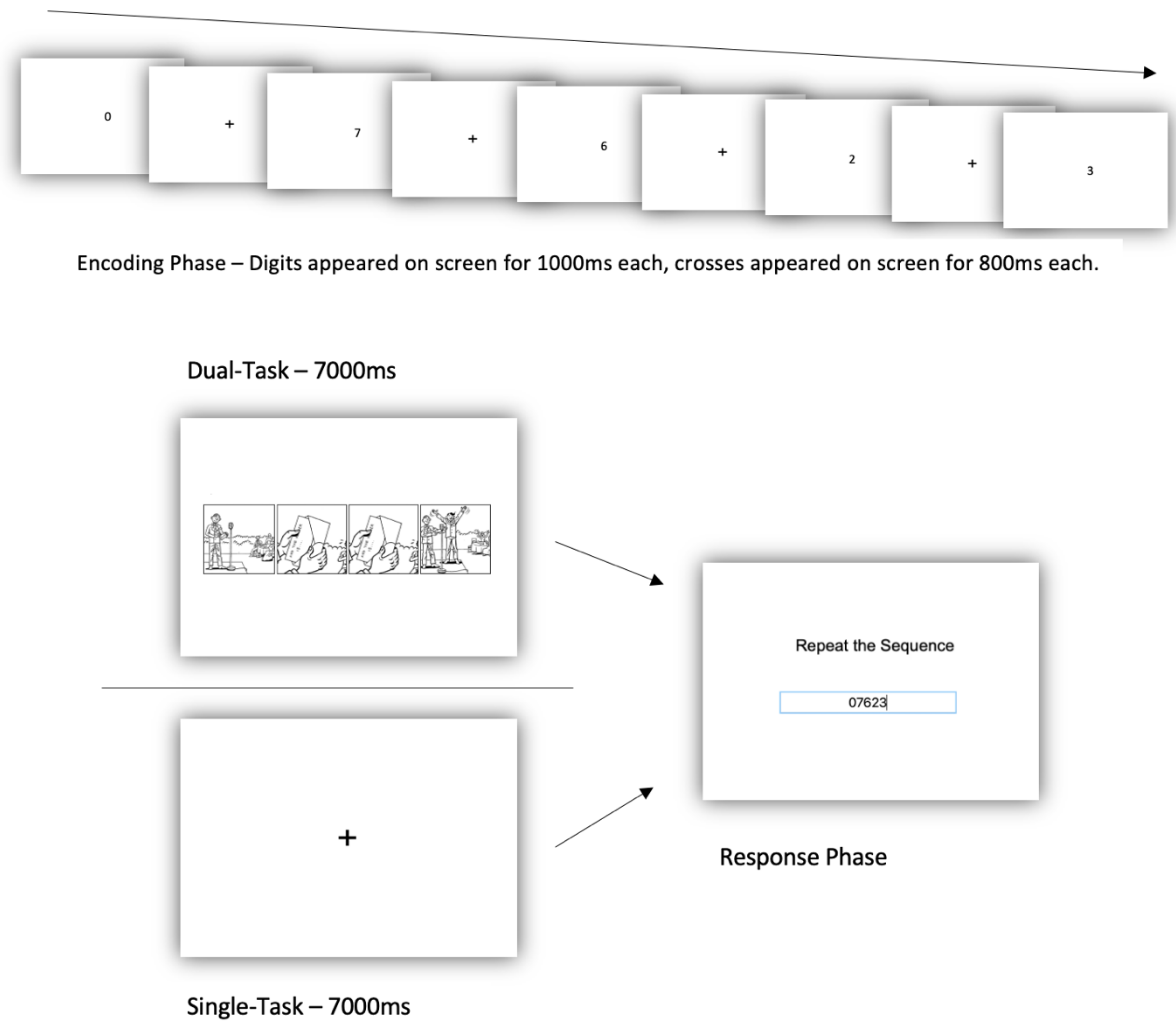
2.3.2.4 Procedure

Participants completed the experiment in booths that were designed to be distraction-free. At the outset of the experiment, participants were asked to provide their age, gender details, and an email address. The presentation of the information and consent checks followed this. Participants then completed the titration phase of the experiment for the digit span task. Participants were first presented with on-screen instructions about how to complete the digit span task. This was followed by two practice trials. When these were completed, further instructions were presented, indicating that the titration phase was about to begin properly. Digit span task trials started with an on-screen 'Get Ready' instruction. This was followed by the presentation of an interstimulus cross for 2000ms. At the offset of the cross, the first digit in the sequence appeared in the centre of the screen. Each digit in the sequence was presented for 1000ms, followed by an 800ms interstimulus cross. Following the final digit in the sequence, the participants were presented with an on-screen instruction to 'Repeat the Sequence' for 2000ms. This was replaced with a text entry box into which the participants entered their responses.

As in the Corsi block test, a positional scoring method was used whereby, for each number entered as part of the participant's response, they were considered correct if they were in the correct position. For example, a response of 12354 for the sequence 12345 would be regarded as 60% accurate based on the misplacement of 5 and 4. The programmed version of this test was designed with certain mitigations to ensure that the sequence did not repeat numbers as much as possible or contain the same numbers within three indices of each other.

As in the previous experiments, the titration phase increased in difficulty by increasing the sequence length if the participant registered two correct responses at any trial level. If they registered an incorrect response at any given level, they were then required to register three correct trials to progress. Alternatively, the titration phase ended, and their verbal working memory interference was then set at the level below the point that they had registered two incorrect responses to trials.

Figure 7 *Graphical Representation of the Digit Span Procedure used in Experiment 3*



Once the titration phase finished, the participants completed the Visual Language Fluency Index and were given a three-minute break. After this, they were given further instructions for the following part of the experiment. Depending on their assigned condition, the participants completed either dual- or single-task trials. As in the previous experiments, dual-task trials replaced the ‘Repeat the Sequence’ instruction with a comic strip presented for 7000ms. The single-task condition replaced this instruction with an interstimulus cross for 7000ms.

After each dual-task trial, the participants completed the time-span question related to the comic they had just seen. In this condition, the participants also completed the comics recall and question phases after completing the six dual-task trials. Participants were then verbally debriefed and informed that the information, debrief, and consent forms had been emailed to them upon completing the experiment. This procedure took around 25 – 35 minutes. Figure 7 depicts the digit span trials.

2.3.3 Results

2.3.3.1 Digit Span Descriptives and Assumptions

The same data preparation procedures were applied to the digit span dual-task and single-task data with average accuracy across six trials for either dual-task or control procedures computed for each participant. Shapiro Wilk normality tests of both the dual-task ($W = 0.86$, $p = 0.006$) and control ($W = 0.89$, $p = 0.02$) procedures indicated that neither variable could be approximated to a normal distribution. However, homoscedasticity was achieved ($F(1, 38) = 0.96$, $p = 0.33$). Descriptive statistics for each condition revealed that a ceiling effect was likely the cause of the non-normality, with the high levels of accuracy displayed in both the control ($M = 0.88$, $SD = 0.09$) and dual-task digit conditions ($M = 0.88$, $SD = 0.12$). Participants in this experiment showed similar comprehension and recall accuracy with an average free recall score of 86%.

Table 3 *Digit Span Single and Dual Task Accuracy Descriptives*

Condition	N	M	SD
Dual-Task	20	0.886	0.123
Single-Task	20	0.879	0.088

Note. Number of Participants (N), Means (M) and Standard Deviations (SD) for both the Digit Span dual-task and single-task conditions.

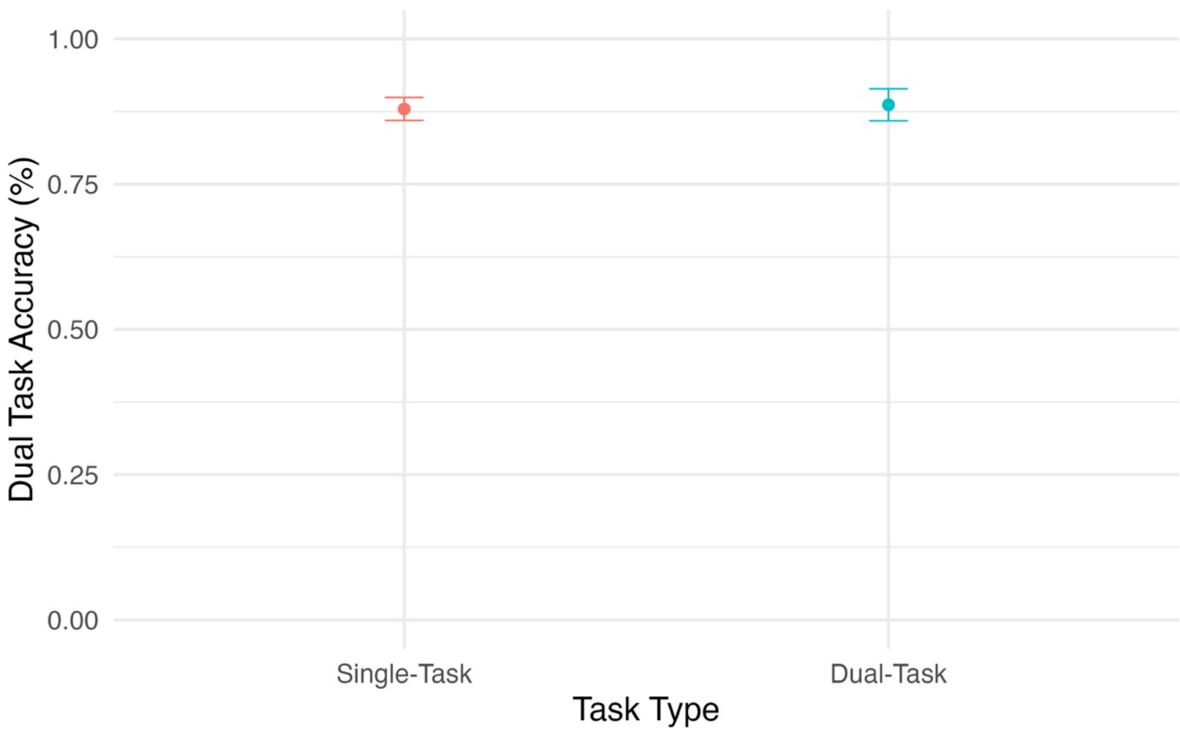
2.3.3.2 Digit Span Dual-Task Accuracy Analysis

Because both the experimental and control variables were significantly not normally distributed, non-parametric analysis procedures were adopted to evaluate the difference in the performance of the digit dual-task and control procedures. A Wilcoxon Rank Sum test confirmed that the difference of -0.03 (95% *CI* [-0.07, 0.06]) between performance in the digit span dual-task and digit span single-task conditions was not significantly different ($W = 172$, $p = 0.45$). This analysis suggests that verbal working memory resources do not significantly contribute to visual narrative processing. Figure 8 depicts the estimated marginal means of single and dual-task Digit span accuracy.

2.3.3.3 Digit Span Panel Variant Analysis

The normality ($D = 0.32$, $p = <.001$) and sphericity ($W = 0.662$, $p = 0.024$) tests of the digit span dual-task accuracy by panel variant model both indicated significant violations of these assumptions. Because of these violations, the non-parametric Friedman's analysis was used. This revealed no significant effect of panel variant on dual-task accuracy χ^2 (2) = 1.1364, $p = 0.566$.

Figure 8 *Digit Span Accuracy for Single and Dual-Task Conditions*



2.3.3.4 Visual Language Fluency and Visual Pattern Test Accuracy

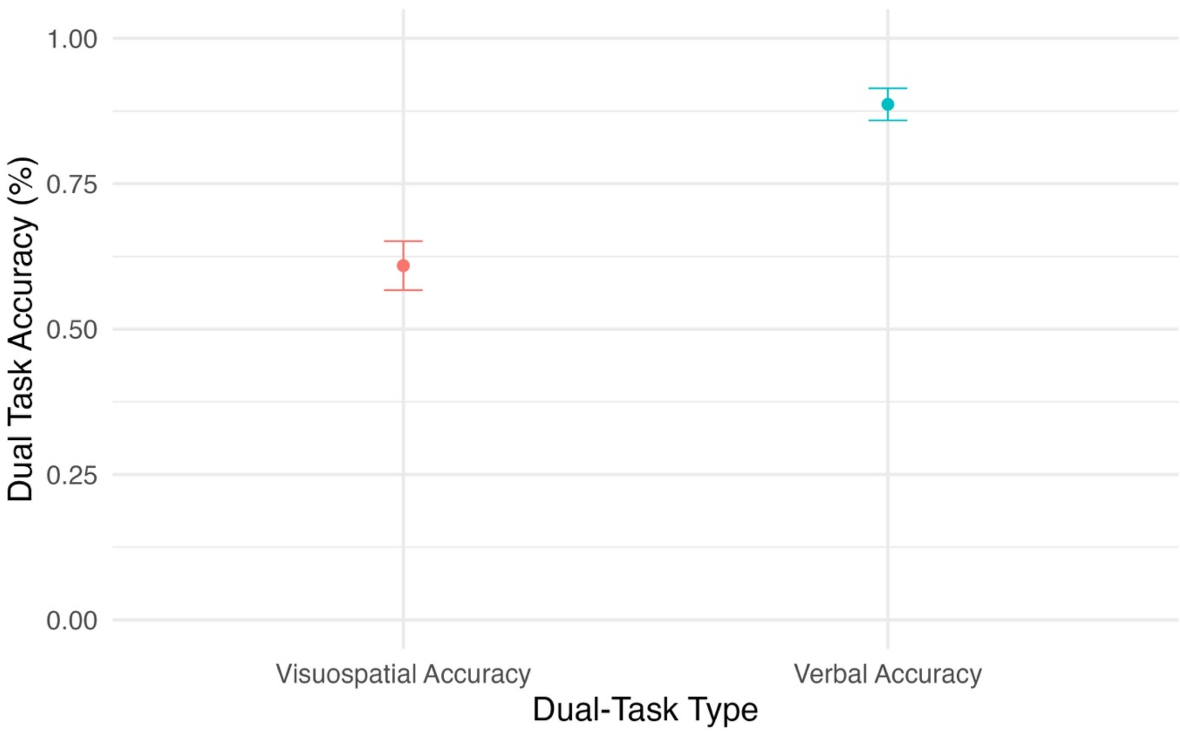
As in experiments 2 & 3, Pearson correlations were conducted to examine the relationship between single and dual-task accuracy with VLFI scores. The dual-task group had a relatively high mean VLFI score of 10.06 (*SD* = 7.04). Comparatively, the single-task group had a relatively low average VLFI score of 5.95 (*SD* = 4.21). In both the dual-task $r(18) = -.13$, $p = .576$, 95% CI [−.54, .33] and single-task conditions $r(18) = -.07$, $p = .775$, 95% CI [−.50, .39]. The correlation between performance and VLFI was small, negative, and not statistically significant. This suggests no reliable linear relationship between these variables. Overall, the results indicate that there was no meaningful or statistically reliable association between accuracy and VLFI in either condition.

2.3.3.5 Contributions of Visuospatial and Verbal Working Memory

A Shapiro-Wilk test was used to check the normality of the Corsi dual-task and the digit span dual-task participants' visual language fluency scores. Where the Corsi dual-task groups' experience with comics was found not to be normally distributed ($W = 0.88, p = 0.018$) with a moderate to large skewness value (0.86), the digit span groups' experience with comics was found to be normally distributed ($W = 0.92, p = 0.094$) with a lower skewness value (0.68). Levene's test was conducted to examine the homogeneity of variance across the groups for VLFI scores. The results showed no significant difference in variance between the group $F(1, 38) = 2.56, p = 0.118$. Because of the non-parametric nature of the data, a Wilcoxon signed ranks test was used to evaluate any differences between the participants' comics experience. This, however, revealed no significant difference in comics experience between the two groups ($W = 179, p = 0.579$).

Corsi block dual-task accuracy was determined to be normally distributed in the previous experiments. In contrast, as above, digit span dual-task accuracy was not normally distributed. Levene's test assessed the homogeneity of variance between these two groups and revealed no significant difference in variances $F(1\ 38) = 1.23, p = 0.274$. Because the digit span data was demonstrated to be non-parametric, a Wilcoxon signed ranks test was used to determine whether there was a significant difference between the visuospatial and verbal dual-task accuracy. This indicated that participants in the visuospatial dual-task condition were significantly less accurate than participants in the verbal dual-task accuracy condition. Figure 9 displays the estimated marginal means of the dual-task accuracy of the visuospatial and verbal working memory.

Figure 9 *Comparison of Digit Span and Corsi Block Dual-Task Accuracy*



2.3.4 Discussion

This experiment aimed to investigate the role of verbal working memory in visual narrative processing. The hypothesis posited that since comics do not rely on phonological decoding, verbal working memory would remain unaffected when simultaneously processing a visual narrative. To test this, a titrated dual-task paradigm was employed. If verbal working memory plays a role in visual narrative processing, performing a verbal working memory task alongside processing an image sequence would result in significant interference compared to a single-task condition.

In line with the hypothesis that processing a visual narrative would not affect verbal working memory, this experiment did not return any significant interference effect when pairing comics reading with a verbal working memory task. In contrast to the findings of Experiments 1 and 2, which tested visuospatial components of working memory, processing a comic caused no interference to the maintenance of a verbal working memory stimulus, with participants' ability to perform the dual-task experiment being equal to the participant's ability to perform the single-task condition. This supports the suggestion that processing a sequence of images does not require verbal working memory.

Secondly, this experiment aimed to understand the relative contributions to image sequence processing of visuospatial and verbal working memory. Participants' performance from the Corsi block and digit span dual-task groups was compared to test this. The results of this analysis indicated clearly that the relative contributions made to visual narrative processing of visuospatial and verbal working memory were not equal, with visuospatial working memory seemingly contributing to this kind of processing to a significantly greater degree than verbal working memory. In fact, the findings indicate that, under the conditions tested, that there was no statistically detectable difference in verbal working memory performance between the dual-task and single-task conditions. This lack of significant interference suggests that verbal working memory may not be engaged to the same extent as visuospatial working memory during visual narrative processing; however, this interpretation must be tempered by the study's limited power and sample size.

While the findings of Experiments 1 and 2 align with those of Magliano et al. (2016), the findings of Experiment 3 provided contrary evidence. Magliano and colleagues were similarly interested in the relative contributions of different working memory components. However, their experiment was focused on the role of working memory components activated during inference. Their participants were required to read comic sequences that contained missing panels, meaning they were required to infer missing content to comprehend the comic. Their results indicated the same level of interference caused by a verbal and visuospatial working memory stimulus. However, these results were based on panel viewing times. They infer that the lack of significant difference between viewing times when secondary visuospatial and verbal working memory loads were present indicates that both components are used during visual narrative processing. The current results then provide an alternative perspective that points to the conclusion that verbal working memory is not associated with processing comics.

The contradiction in findings is likely explained by methodological considerations implemented in this study. Firstly, it is argued that the experiments presented in this chapter provide a more direct evaluation of the degree to which each of the working memory components tested contributes to processing a sequence of images. This was achieved by the change in analytical emphasis from the secondary task analysed in Magliano et al. (2016) to the primary task analysed here. This is because, in contrast to measuring viewing times and approximating this as the implication of working memory components in processing sequences of images, the current experiments directly tested the level to which processing a comic interfered with verbal working memory maintenance. This arguably provides a clearer indication of the contribution of each working memory type. Though Magliano and colleagues report similar elements of their findings, they only report having determined that the ability of the participants to complete the tasks was significantly above chance. Interestingly, their findings arguably support those of the current experiment in that participants' accuracy in completing the visuospatial dual-task was 54.6% (assuming percentages) and 66.6% for the verbal working memory dual-task.

This then alludes to the second methodological advantage of the current experiments. Whereas Magliano et al. (2016) applied blanket seven-item length sequences across all participants and both verbal and visuospatial working memory tasks, the current experiments employed a titration procedure prior to the participants completing both the dual-task and control conditions, meaning that the level of working memory load was individualised to each participant. That their participants were only able to recall the working memory concurrent loads with approximately 55% and 67% accuracy arguably speaks to the requirement to titrate working memory concurrent loads (Doherty & Logie, 2016).

While primarily, these findings are of value in demonstrating the cognitive processes associated with comprehension of image sequences; they also contribute to working memory literature more widely by suggesting support for modality-specific functionality. The finding that processing an image sequence did not significantly disrupt the participants' ability to maintain a verbal working memory stimulus indicates that at the very least, storage or maintenance processes of verbal working memory can be employed with some level of modality specificity for different types of information such as verbal and visual data (Cocchini et al., 2002; Jarrold et al., 2011; Thalmann & Oberauer, 2017).

The findings of Experiments 1, 2 and 3 have significant and advantageous implications for the practical application of image sequence-based communication. Considering that cognitive processing capacity is stringently limited, the potential for the medium of comics to distribute cognitive processing costs across components of working memory - as the current findings would suggest - highlights the potential of comics as a medium of information communication that could effectively mitigate some cognitive processing costs. This is complemented by findings demonstrating that conceptual understanding and recognition of information conveyed by images can be achieved as rapidly as 13ms (Potter et al., 2014).

As discussed in the introduction to this experiment, visual narratives are suggested to be governed by language-based processes. This formed part of the initial justification for investigating verbal working memory in visual narrative processing. It was considered that verbal working memory might be responsible for, or associated with, processing visual narratives in line with the suggestion that linguistic principles govern the way that visual narratives are constructed and understood (Cohn, 2013; 2014; Cohn et al., 2016; Magliano et al., 2016). However, the current experiments show that verbal working memory is not involved in visual narrative processing. Instead, this process relies on visuospatial working memory, supporting the idea of a parallel multimodal structure for visual language processing (Cohn, 2014).

Although elements of the methods used in these experiments are argued to represent significant strengths, a number of methodological limitations must also be addressed across the three experiments presented in this chapter. Firstly, the use of a between-groups design where different participants completed visuospatial, visual and verbal working memory tasks, as well as the individualised control tasks, can be seen as a limitation of the current experiments. This design arguably weakens the study's ability to effectively explore the contributions of various working memory components to visual narrative processing. However, this limitation is mitigated reasonably by the cognitive load titration employed across these experiments. In implementing these procedures, the load assigned to each participant during the experiment's dual- and single-task phases were calibrated to account for individual differences in working memory capacities. The effect of this is to reduce the variability in performance that might arise from baseline working memory differences between participants. Worth noting here is that these design choices were made in line with the previous aim of the project. These were important considerations during development and were taken largely for practical reasons relating to the relatively small number of stimuli used in the project. Relatedly, the sample sizes of the experiments presented in this chapter are also relatively low. Although 120 participants were recruited in total, each condition had a relatively small sample size of 20 participants. Consequently, this arguably limits the generalisability of the findings. Where practical, the further experiments of this thesis have subsequently addressed this limitation with repeated measures design.

The finding that visuospatial working memory underpins visual narrative processing provides a useful base for future research into how verbal working memory might be used in comics that use phonological components. Phonological information enriches and contextualises visual narrative processing, whether in dialogue, narration, or sound effects. Given that the base component of visual narratives can arguably be accounted for as visuospatial processing, manipulating phonological components could provide interesting findings in relation to how verbal working memory is used in visual narrative contexts.

Following the conclusion that comics processing is driven by visuospatial working memory, these experiments raise the more specific question of how visuospatial working memory is used in visual narrative processing. Though these experiments were not designed to answer this question, the results suggest that visuospatial working memory is implicated in higher-order cognitive processing. One potential answer to this question is posed by the Scene Perception and Event Comprehension Theory (SPECT: Loschky et al., 2020) that visuospatial working memory provides the cognitive capacity and workspace to process the narrative component of the comic. This converges with the conclusions of Magliano et al. (2016) in their suggestion that visuospatial working memory supports the process of inference generation in comics processing.

A key limitation of the findings that must be acknowledged is the lack of a structured comprehension check administered to the participants immediately after their viewing of each narrative. Comprehension was indirectly assessed through free recall and, as is presented in the first three experiments, indicates that the participants overwhelmingly demonstrated a reasonable understanding of the narrative content. The overall free recall accuracy score of all participants who completed the dual-task conditions of these three experiments was 81%. This arguably indicates that the vast majority of participants encoded and also comprehended the comics sufficiently for the purposes of the experiment and a qualitative review of the responses registered by participants further supports this suggestion.

Related to this, the participants also answered a small set of open-ended questions about each narrative after completing all trials and the free recall task. These were devised to test memory for a specific visual component of each narrative. For example, the question related to the Archer comic asked participants, “What facial hair did the archer have?” Similarly, the Bus question asked participants, “What number was on the side of the bus?” As with the panel manipulation, these questions were devised early on the research process prior to the change in focus to the role of working memory in visual narrative processing. Their initial intention was not to check comprehension but memory, and thus, have not been included in the comprehension check presented in this experiment as they are limited in their utility as a measure of online comprehension. Further to this, their administration at the end of the experiment as opposed to immediately after the presentation of each narrative supports this point. In response to this limitation, future dual-task experiments conducted in this thesis will include a direct comprehension check immediately following the presentation of narratives.

The involvement of visuospatial working memory in comics processing may also encompass situation model construction within visual narratives. Narrative and discourse processing theories (Kintsch & van Dijk, 1988) propose that readers build mental representations—situation models—to understand content. This framework aligns closely with working memory principles, and research has demonstrated a connection between the two. However, most studies in this area have focused on text-based narratives. Comics, with their reliance on sequential images, offer a unique opportunity to examine these processes from a different perspective, which will be explored in the following experiments.

In summary, although the results of this study align with the view that visuospatial working memory plays a significant role in processing image sequences, the absence of a statistically significant effect for the implication of verbal working memory in this process highlights the need for further research. Where previous research has suggested that both verbal and visuospatial working resources support the processing and comprehension of image sequences, this experiment finds direct evidence only for the use of the visuospatial component. However, the absence of a significant interference effect does not indicate that verbal working memory is not implicated in visual narrative processing. If confirmed, these results could have implications for discourse comprehension models as they potentially indicate that the base processes associated with understanding image sequences differ from those of processing text. However, future research should aim to replicate these results with larger samples to ensure their validity. The precise functional role of visuospatial working memory resources used throughout image sequence processing will be explored further.

Chapter 3

3.1. Experiment 4 – Situation Models in Visual Narrative Processing

3.1.1 Introduction

Following the findings of Chapter 2, which indicate that visuospatial working memory underpins visual narrative processing, this experiment was conducted with two aims. Firstly, while research suggests that situation model processing occurs in comics (Cohn & Wittenberg, 2015; Gernsbacher et al., 1990; Brich et al., 2024), the initial aim of this experiment was to verify the application of situation or event model processing to the current stimulus. This was on the basis that to investigate the potential role that working memory plays in situation model processing of visual narratives; it was considered necessary to confirm that the stimuli used in this thesis were being processed under this framework. This was due to a) their relatively short length of only three panels and b) their novelty as stimuli in experiments. The latter point here refers to the idea that authors in this area have typically used consistent stimuli across series of experiments. For example, the A Boy, A Dog and A Frog series (Mayer, 1967) as used in Magliano et al. (2016) and Magliano et al., (2017) and Peanuts (Schulz, 2004) as used by Cohn, (2014) and numerous other studies are well established as visual narrative stimuli. The current comics, on the other hand, are novel as stimuli in experiments of this kind.

The second aim was to investigate the utility of the recognition time paradigm in understanding visual narrative situation model processing. While previous studies have shown that situation model processing is applied to comics using reading times, few have examined narrative processing behaviours through recognition times. This metric offers potential for uncovering new insights into the nature of situation model updating, as explored below. In the context of this research, the terms mental model, situation model, and event model will be used interchangeably to describe the cognitive processes associated with narrative comprehension.

Cognitive processing of narratives has long been theorised to occur as a process of mental model creation (van Dijk & Kintsch, 1978; Johnson-Laird, 1983). In order to comprehend information conveyed to us, particularly in a narrative format, incoming information is organised into a conceptual model rooted in knowledge bases, including previous experiences and our understanding of the physical world. The operational processes of these models have been a significant source of research interest. Two prevailing theories have described the processes associated with understanding narratives from the perspective of mental model creation with different fundamental bases.

Zwaan et al., (1995) propose a model in which the primary focus when comprehending narratives is entities and actions. Their Event Indexing Model suggests that situation models, and thus narrative understanding, are constructed based on five dimensions: spatial, causal, temporal, intention and entity (Zwaan & Radvansky, 1998; Zwaan et al., 1998). As readers progress through narratives, it is suggested that each of these dimensions is consistently tracked for significant changes. Changes detected along any of these dimensions then cause a process of updating to occur, at which point, the new information presented is integrated into the currently held situation model. This iterative process allows the reader to maintain the most up-to-date information about the narrative. Critically, at the point that updating cannot accommodate new information, for example, when there is narrative resolution or in cases when narratives convey discontinuity or incoherence, the current situation model is curtailed, and a new one is created (Gernsbacher et al., 1997; Zwaan & Radvansky, 1998; Hoeben-Mannaert & Dijkstra, 2021).

The theory of Event Segmentation (EST: Zacks & Tversky, 2001; Kurby & Zacks, 2008; Radvansky & Zacks, 2011) sets out a structure of event comprehension which overlaps considerably with the Event Indexing Model. There is an ongoing debate about the definition of an 'event' (Yates et al., 2023), but broadly speaking, events are defined as discrete units of mental experience derived from perception. This definition reveals that whereas the Event Indexing Model was developed as a theory specifically of narrative comprehension, Event Segmentation theory pertains to the perception of general experiences and memory. Because of this, the theory is grounded in more general cognitive assumptions, which allow it to explain a broader set of cognitive phenomena. However, it has been demonstrated to have explanatory power in investigating narrative comprehension, particularly in relation to narratives conveyed through visual means (Huff et al., 2014; Brich et al., 2024; Cohn et al., 2024). To account for the process of updating a situation model, the theory posits the prediction error hypothesis. This hypothesis states that prediction is a fundamental part of perceptual processing and that prediction accuracy is variable. When prediction error is low, perceptual processing is facilitated. However, when prediction error is high, processing is more strenuous.

A gating mechanism is a useful analogy for the Event Segmentation process (Reynolds et al., 2007; Swallow et al., 2009). In this analogy, perceptual processing receives input that initialises an event model. This serves as the basis for the continuous processing of an event. To do this, nodes of information associated with the event model are activated and utilised to generate subconscious predictions of the event and prime the reader for future content. Because cognition of this kind is costly, the base model and relevant activated processing networks are placed behind a gate that is administered by prediction error. As the reader progresses through the narrative, if predictions that have been made inside the gate remain relevant, meaning that prediction error is low, the gate remains closed, and processing is facilitated. When prediction error increases and the information within the gate is no longer facilitative, the gate opens. Processing becomes more intensive at this point, and an event boundary is created. This either results in updating, whereby the same model is maintained but integrated with new information, or an entirely new model is created if the information deviates sufficiently from the old model.

Reading times have provided substantial evidence for processing behaviours during situation model creation. The amount of time spent engaging with a text is widely argued to reflect the depth and extent of cognitive elaboration of the text. This has then been directly linked to the process of situation model construction. The patterns observed in reading times have largely been shown to align consistently with situation model processing, which supports the idea that reading times indicate cognitive effort.

Reading time findings under both the Event Indexing and Event Segmentation frameworks present evidence for the suggestion that constructing a situation model is cognitively demanding (Zacks et al., 2009; Therriault et al., 2006; Rinck & Weber, 2003; Gernsbacher et al., 1990; Gernsbacher & Faust, 1991). Situation model construction involves using existing knowledge to establish a perceptual foundation for processing subsequent narrative elements. Research suggests that multiple situation models are created throughout narrative processing (Pettijohn & Radvansky, 2016). This occurs primarily at two points: during the initial construction at the beginning of a narrative and when the current model becomes inadequate for understanding new information.

Reading times have also demonstrated consistency with situation model processing in comics. In previous research which has manipulated coherence, panel reading times have been shown to decrease to a plateau point in coherent comics (Cohn & Wittenberg, 2015). This is characteristic of the initial heavy processing demand followed by facilitation in reading times as long as the information in the current model remains coherent and relevant. More recently, the concept of creating a new situation model following the completion of a narrative, also termed shifting (Gernsbacher et al., 1990; Gernsbacher & Faust, 1991), has also been demonstrated in comics. In line with situation model processing, when two panel sets conveying independent narratives were placed in succession, participants' reading times were shown to increase when viewing the initial panel of a second event, indicating a re-initialisation or the creation of a second situation model from which to process further narrative components (Brich et al., 2024). Reading times then have clear utility in demonstrating situation model processing in visual narratives.

Another fundamental component of situation and event model processing is model accessibility, which relates to updating (Zwaan & Radvansky, 1998). Updating occurs when new information is coherent with the current situation model but must be integrated as the next component of the narrative to then understand subsequent narrative components (Radvansky & Zacks, 2011). In contrast to basic situation model construction processes, updating is a more complex process, the precise mechanisms of which are subject to ongoing debate (McNamara & Magliano, 2009).

Though reading times can demonstrate updating behaviours (Radvansky & Copeland, 2006), recognition times can arguably be used to greater effect in demonstrating situation model updating behaviours. The premise of using this metric as opposed to reading times relates to the accessibility of information. If information is organised into a situation model framework, then as models are updated, the previous information becomes less accessible as the focus of attention is removed from the older content (Bailey & Zacks, 2015; Pettijohn & Radvansky, 2016). At the point of updating, older information enters a state of decay and, if probed, must then be retrieved from longer-term memory after having been deactivated. This then manifests itself as increased retrieval time, which recognition probes can index for components of a narrative that are no longer active compared to current or most recent model components (Glenberg et al., 1987).

As suggested above, the Event Indexing and the Event Segmentation models have suggested different updating methods. The Event Indexing Model proposes an incremental updating process (Zwaan et al., 1995). Incremental updating suggests that updates occur along the dimensions outlined in the Event Indexing model and only along the dimensions that require updating. For example, if a narrative has established a character in a specific space, and the character then moves to another space, the model would be updated along the spatial dimension, making information about the previous space less accessible. In this regard, updating should only differentially affect the specific dimensions considered to be discontinuous when new information is presented. Alternatively, Event Segmentation theory proposes that, instead of updating incrementally, models are updated globally at event boundaries, which are points during processing where information is naturally segmented into distinct events. This would suggest that the current event model is flushed when new information is encountered that requires readers to update situation models (Zacks et al., 2007; Zacks & Swallow, 2007). Though evidence has accumulated against the view of incremental updating (Radvansky & Copeland, 2010; Pettijohn & Radvansky, 2016) and importantly, recognition time findings have returned evidence in favour of global updating than incremental updating (Bailey et al., 2017; Papenmeier et al., 2019), other evidence suggests that both of these updating types are exhibited by readers (Kurby & Zacks, 2012) with incremental updating perhaps occurring within events, and global updating occurring at boundaries.

Unlike text, the visual nature of comic strips creates interesting questions for the concept of situation model updating. In literature on event models, cognitive load is a primary consideration in situation or event model processing. As suggested above, creating and maintaining situation models is cognitively demanding. However, comics and their visual nature potentially reduce this load significantly. This reduction is made clear when considering the base model of discourse processing. Kintsch's (1988) model of discourse comprehension, from which both the Event Indexing model and, subsequently, Event Segmentation Theory have emerged, proposes that comprehension of narratives is considered to be the product of three levels of processing. In processing a prose-based narrative, the surface level is initially parsed, which comprises low-level processing of the narrative's base form. This results in understanding the basic propositions of the narrative as they appear. The semantic relations between these basic propositions then formulate the second level of processing. Finally, following an understanding of surface and semantic content, the concept communicated by the information is constructed into a representational situation model, providing narrative or discourse comprehension.

These levels of processing represent a logical operational sequence for text comprehension. However, they may be less applicable for directly depictive material such as comics. While clearly, images contain similar semantic relational content that must be understood in order to resolve their representational meaning (Cohn, 2013; 2014; 2016; Cohn et al., 2014), by comparison to text, these processes are likely to occur at a comparatively inordinate speed. Evidence for this suggestion comes from findings which indicate that levels of understanding can be derived from images at rapid speeds (Potter et al., 2014), and even coherence can be maintained when sequential images are viewed at rates of 1 per second, for example (Hagmann & Cohn, 2016). The immediacy of information presented in an image contrasts significantly with the requirement to decode information presented via text. This may mean that processing visual narratives, particularly static visual narratives such as comics, at the very least expedites the initial stages of narrative comprehension as outlined by previous processing theories. How this then manifests in situation model processing, particularly in relation to the concept of updating, is an open question.

This experiment aimed to test various predictions about situation model processing in the visual narrative stimuli used in this thesis. To do this, the coherence of the comics was manipulated. Coherent comics were those that presented a coherent narrative throughout the three panels of the set. Incoherent comics were sets of comics that contained no discernible narrative. Participants were asked to view comic panels in sequence, and their reading times for each panel were recorded. Following this task, the participants then completed a recognition task in which they were asked to indicate whether or not they had seen a given comic panel during the viewing task.

In relation to reading times, it was considered that the coherent comics would be organised into situation models in contrast to the incoherent comics, which would not. Based on findings of situation models providing predictive facilitation, it was hypothesised that 1) overall, coherent sets would be read significantly faster than incoherent sets. Further to this, it was hypothesised that 2) in coherent comics, reading times would decrease with each subsequent panel. Relatedly, on the basis that the first panels should always serve to initiate situation models, it was hypothesised that 3) there would be no difference between initial panels in coherent or incoherent sets. Conversely, it was considered that 4) reading times for incoherent comic panels would not demonstrate any significant difference between each other.

Recognition times have the function of testing the accessibility of information. In the current context, memory for panel contents should be facilitated if it is organised into the structure of a situation model. In relation to this, it was hypothesised that 5) the coherent comic panels would be recognised faster than incoherent comic panels overall. Further to this, it was hypothesised that 6) if situation model updating occurs during the processing of these stimuli, then there will be a similar downward gradient of recognition times across coherent comic panels. Finally, it was hypothesised that 7) there would be no significant difference in panel recognition times between any panel positions of the incoherent comics as, logically, updating in these sets would be an invalid operation.

3.1.2 Method

3.1.2.1 Participants

84 participants (71 Female) were recruited to complete the study in exchange for course credit on an undergraduate Psychology course. Their ages ranged from 18 to 35 years ($M = 20$, $SD = 2.21$). All participants gave informed consent to participate in the study and were naïve to the purposes and intentions of the study. The University Research Ethics Committee approved the experiment.

3.1.2.2 Design

This experiment employed a repeated-measures design with two conditions: coherent comics and incoherent comics. Three dependent variables were measured to assess whether situation model processing characteristics are applied to the current comics stimuli . The first dependent variable was reading time in milliseconds, defined as the duration of time participants spent reading each comic panel. After the reading phase, participants completed a recognition test in which they identified whether they had previously seen specific panels. The second dependent variable was then the proportion of correct responses participants' made during the recognition test. The third dependent variable was recognition time in milliseconds.

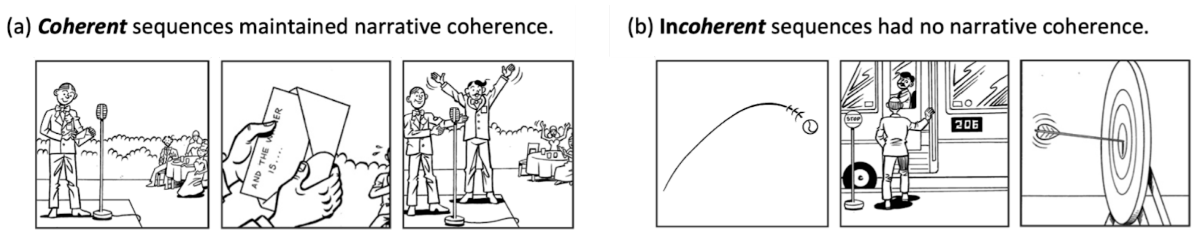
The independent variable for both reading and recognition tasks was the coherence of the comic panel sequences. Participants were presented with two types of panel series: coherent narratives, which followed a logical, connected storyline, and incoherent sequences, where panels were randomly selected from different narratives, resulting in no discernible storyline. This design allowed for the assessment of how coherence impacts reading times, recognition accuracy, and recognition speed.

3.1.2.3 Materials

Each of the 14 unique narrative comics presented in the appendix (A – M)was used in this experiment. However, they were modified to ensure parity between the size of the panels. Given that the panels were intended to be presented to the participants individually in sequence, one of the middle panels of each comic from the four-panel sequences was selected for use as the second panel of each set. This then standardised the size of each panel in each comic. The coherent comics used in each run of this experiment were randomly selected by the program created to conduct the procedure every time it was run. The incoherent comics were created by selecting a number of the coherent comic sets, dividing them into their unique components and then placing these individual, random components together into three-panel comics. As with the unique selection of coherent comics, this procedure was written into the experiment program and was executed every time the experiment was run. Checks were put in place to ensure that no comic contained two panels from the same set and, in generating the presentation, that incoherent comics which contained panels from the same set were not shown in succession.

Four coherent sets were selected, and four incoherent sets were created from the materials. This used 24 of the panels for the narrative-based comics, leaving 18 for the recognition test as the unseen panels. Figure 10 demonstrates the narrative coherence manipulation.

Figure 10 *Examples of Coherent and Incoherent Comic Panel Sets*



3.1.2.4 Procedure

This experiment was created as a Python script by the author and run on a university-managed laptop. Participants completed the experiment in booths designed to be distraction-free. The experiment typically took between 35 and 40 minutes to complete. All participants completed the experiment on a HP Elitebook laptop with a screen resolution of 1920 × 1080

Participants were initially presented with the information sheet and the consent form. They then entered their age, gender details, and email addresses.

In the reading time task, the participants were first presented with on-screen instructions related to the task. These instructions asked the participant simply to 'Read' the panel and, as soon as they had done so, to press the space bar. Each panel presentation started with an interstimulus cross that appeared for 1500ms. This was followed by a comic panel. At the offset of the comic panel, the participants were presented with a question designed to prompt attention. For the first panel in any given sequence, the participant was simply asked to respond yes or no to whether they understood what had happened in the panel. For either of the subsequent panels in any given sequence, the participant was asked whether the panel was related to the one prior. Both coherent and incoherent panel sets were presented in this way. The coherent and incoherent comic set presentation was systematically varied across the participants. This was achieved by setting the presentation of the different comics into four sets of alternating blocks to be presented to the participants. This was implemented to mitigate any sequence effects.

After the participants had completed the reading time task, they were given further instructions about the recognition task. In this phase, the participants were asked to view comic panels and respond using the keyboard keys 'F' or 'J' in response to whether they had seen these panels during the reading task component. The 'F' key indicated 'No', and the 'J' key indicated 'Yes'. This phase began with a series of 30 habituation trials. These trials started by presenting an interstimulus cross for 1500ms. At the offset of the cross, the words 'Yes' or 'No' appeared in the centre of the screen. In response to these stimuli, the participants were required to press the key corresponding to the on-screen prompt as fast as possible. There were equal 'Yes' and 'No' trials. The intention of this task was to habituate the yes and no responses to the relevant keys.

On completing the habituation task, participants were presented with further on-screen instructions for the main component of the task. These instructions stated that the participants were required to view comic panels, decide whether they had seen them in the previous reading task and then respond as quickly as possible using the 'No' and 'Yes' keys ('F' and 'J'). They also stated that the participants should keep their fingers on the 'F' and 'J' keys.

The recognition task trials started with the presentation of an interstimulus cross for 1500ms. A comic panel was presented to the participant at the offset of the interstimulus cross. Once the participant had registered a response, the amount of time that the panel had been on-screen was recorded, and the trial sequence began again with an interstimulus cross. Two 30-second breaks were built in and triggered after the first and second thirds of the stimulus had been responded to. These were implemented to allow the participants to relax their attention.

The recognition task counterbalancing related to the order in which the panels seen by the participants during the reading time task were presented. This measure was taken to reduce the possibility of priming effects during the recognition task, which may have resulted from presenting the coherent panels to the participants in an order that primed their response to subsequent trials. Given the number of recognition stimuli, a Latin square was not appropriate. Consequently, the order of the recognition stimuli was predetermined into six lists, ensuring that each panel set was presented in varying, balanced orders across these lists to avoid any sequential bias. The unseen recognition stimuli were always set into the same positions of this list in order to further break up any order effects.

3.1.3 Results

3.1.3.1 Reading Times Data Preparation

The data was initially screened based on recognition accuracy. Participants who recorded a recognition accuracy of less than 75% were excluded from the analysis, as this was considered a significantly low score on the recognition test. This decision was based on a previous study that demonstrated a high level of recognition for comic panel-like images (Magliano et al., 2017). This removed 10 participants from the data set, leaving 74 participants' data for analysis.

Reading times of the remaining 74 participants were then screened for excessively high or low times. Any time three standard deviations above the mean or registered as less than 0.75

seconds was replaced with the overall mean reading time of the related panel after these outliers were removed. For example, any coherent first position panel reading time outside this range was replaced with the mean of coherent first panel reading times after removing these outliers.

For each of the six-panel types (Coherent/Incoherent * Three positions each), 336 individual recorded reading times were registered, yielding 2,016 total reading time data points. 58 outliers were identified as being outside the range of 0.75 – 3 standard deviations above the mean. These were then imputed with each panel condition's respective mean (i.e. first-position coherent panel reading times above 19.98 seconds were replaced with the average coherent first panel reading time). The imputed outliers represented 2.87% of the total reading time dataset and were approximately evenly spread across each reading time category.

3.1.3.2 Reading Times Analysis

Following preparation, the data was checked to assess ANOVA assumptions. Skewness statistics for reading time residuals (0.85), as well as a Kolmogorov-Smirnov test (D = 0.07, p = .002) and a histogram of the residuals, confirmed this. While this should be considered when interpreting the results of this analysis, repeated measures ANOVA is typically considered robust to violations of the parametric assumptions (Blanca et al., 2017). Similarly, some authors suggest that, practically speaking, 30 participants is sufficient to meet the normality assumption allowing the use of parametric tests (Kwak & Kim, 2017). As such, the analysis proceeded using parametric tests.

Table 4 *Reading Times by Coherence and Panel Position*

Condition	Panel 1	Panel 2	Panel 3
Coherent	4.59 (2.04)	2.94 (1.04)	2.54 (0.877)
Incoherent	4.55 (1.85)	3.73 (1.39)	3.31 (1.15)

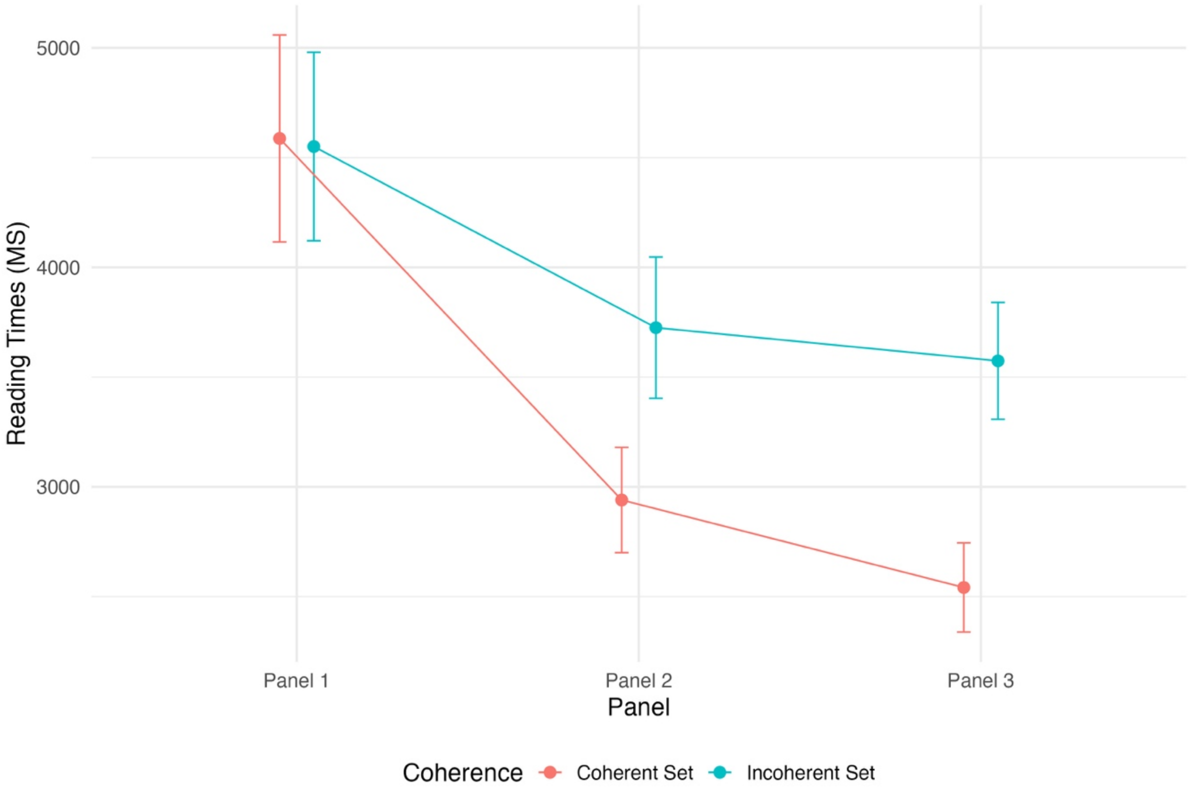
Note. Means and (Standard Deviations) for coherent and incoherent panel set reading times.

A 2 * 3 repeated measures factorial ANOVA was conducted to test the effect of coherence on reading times. Coherence had two levels (Coherent and Incoherent), and the panel position contained three levels (first, second and third). Mauchly's test indicated that sphericity for both the Panel level ($W = 0.77$, $p < .001$) and the interaction effect of Panel by Coherence ($W = 0.78$, p = <.001) was violated. Following this corrected ANOVA results are reported, and the type of correction is indicated when reported.

This analysis revealed that coherence was a statistically significant factor in the amount of time taken to read the comic strips $F(1, 73) = 17.825$, $p < .001$, $\eta^2 = 0.0406$. This was associated with a small to moderate effect size. Overall, this indicates that coherence played a meaningful role in influencing how quickly participants read the comics content.

When controlling for coherence, panel position was also found to have a statistically significant effect on reading times $F(2, 121) = 89.762$, $p < .001$ (Huynh-Feldt corrected), $\eta^2 = 0.1717$. This difference was associated with the largest effect size of the main effect analyses, indicating that in relation to reading times, panel position had the most significant influence on the time taken by the participants to read the comics. Finally, the interaction effect of Coherence by Panel Position was also statistically significant $F(2, 122) = 11.139$, $p < .001$ (Huynh-Feldt corrected), $\eta^2 = 0.0245$. This result indicates significant variation in reading times based on the combination of whether the comic strip was coherent and where the specific panel being read was in the sequence. Holm corrected post hoc tests were subsequently conducted on all of these effects and are reported below in relation to their respective hypotheses. Figure 11 displays the estimated marginal means of reading times by coherence and panel position.

Figure 11 *Effects of Panel Position and Coherence on Reading Times*



Support for the first hypothesis that coherent comics would be read faster than incoherent comics was found in several supporting post hocs. Firstly, the post hoc analysis of overall coherence indicated that participants were, on average, 594ms faster in reading coherent comics than incoherent comics $t(73) = -4.22$, $p < .001$, 95% *CI* [-0.87, -0.31], Cohen's $d = -0.49$, ($MD = -0.594$) and this was statistically significant with a medium associated effect size. Similarly, post hocs of the interaction effect indicated that coherent second $t(73) = -4.86$, $p < .001$, 95% *CI* [-1.28, -0.30], Cohen's *d* = -0.57, (*MD* = -0.785) and coherent third position panels $t(73) = -7.82$, $p < .001$, 95% *CI* [-1.43, -0.63], Cohen's *d* = -0.91, (*MD* = -1.033) were both read significantly faster than their respective incoherent panels positions.

The second hypothesis proposed that there would be a statistically significant downward gradient in coherent panel reading times. Support for this was demonstrated by statistically significant differences between the coherent panels. Coherent first panel reading times were statistically significant by comparison to coherent second panel reading times $t(73) = 8.71$, $p < .001$, 95% *CI* [1.07, 2.22], Cohen's *d* = 1.01, (MD = 1.647). Similarly, coherent second panels took significantly longer to read than coherent third panels $t(73) = 5.02$, $p < .001$, 95% *CI* [0.16, 0.64], Cohen's *d* = 0.58, (*MD* = 0.399). These results indicate clear support for the suggestion that the panels of a sequence were organised into a situation model on the basis that this gradient suggests that each panel was providing facilitative prediction.

Relatedly, the results also revealed support for hypothesis three, that there would be no difference between the initial comic panels of either coherent or incoherent sequences on the basis that the initial panel prompts the formation of a new situation model. As is clear from figure 11, no significant difference was found between coherent and incoherent first-panel positions $t(73) = 0.14$, $p = 0.892$, 95% *CI* [-0.78, 0.85], Cohen's *d* = 0.02, (*MD* = 0.037). When considered in combination with the findings above, which demonstrate differences between second and third-position coherent and incoherent panels, this finding demonstrates that regardless of the coherence of a comic’s subsequent panels, the initial component of a given sequence always caused a slow-down to processing. This could be interpreted as indicating that the first panels always caused the participants to initiate a new situation model to incorporate any subsequent information.

Finally, hypothesis four proposed that, in contrast to the coherent comics, there would be no difference in reading times between incoherent comic panels. Interestingly, this hypothesis found only partial support. In opposition to the hypothesis, the post hoc results indicated that, when reading incoherent comics, participants read panels in the second position significantly faster than they read comics in the first position $t(73) = 4.56$, $p < .001$, 95% *CI* [0.28, 1.37], Cohen's $d = 0.53$, ($MD = 0.825$). However, in support of the hypothesis, incoherent panels in the second and third position were not found to be significantly different from each other $t(73) = 0.97$, $p = 0.674$, 95% *CI* [-0.32, 0.63], Cohen's $d = 0.11$, ($MD = 0.151$). This surprising finding perhaps speaks to the automatic nature of situation model processing. Though this hypothesis only found partial support, the above findings indicate clear differences between incoherent and coherent comics in panel positions two and three. The combination of these findings ultimately provides support for the overarching premise that the current stimulus were processed under the situation model processing framework parameters and this is further explored in the discussion.

3.1.3.3 Recognition Times Data Prep

All recognition time data presented below represents only correct trials. Each participant underwent 42 recognition trials. This yielded 3,528 recognition time data points. Following the removal of 10 participants with a recognition accuracy lower than 75% and the removal of any incorrect recognition trials, a total of 2,787 recognition time data points remained for analysis.

The same screening procedure as applied to reading times was applied to recognition times. The lower limit for this data was set at 0.55, and the upper limit was set at 3 standard deviations above the mean for each respective data set. As with the reading time data, any recognition time data points that fell outside these ranges were imputed with the mean of the data after these had been removed. Of the 2787 recognition time data points used for analysis, 68 outliers (2.45%) were identified that were below or beyond the set range and were imputed with the mean of the related data set. Table 5 displays descriptive statistics for recognition times.

Table 5 *Recognition Time Accuracy Descriptives by Coherence*

Condition	Reaction Time	Accuracy (%)	Overall Accuracy (%)
Coherent	1.043 (0.213)	89.38% (15.21)	89.67% (5.980)
Incoherent	1.096 (0.205)	82.34% (21.15)	

Note. Means (Standard Deviations) for recognition accuracy data split by coherent and incoherent comic conditions. All totals are averages across the data set. Reaction time indicates the average amount of time taken by the participants to recognise either coherent or incoherent panels. Accuracy indicates the overall average accuracy for coherent and incoherent comic panels. The overall average accuracy reflects the participants' average accuracy across both conditions, including the panels that were not shown during the reading time segment of the experiment.

3.1.3.4 Recognition Times Analysis

A 2 * 3 repeated measures factorial ANOVA was conducted to investigate the effect of coherence and panel position on recognition times. This analysis used the same levels as the reading times analysis but with Recognition Times as the dependent variable. Recognition Time residuals indicated a positive skew. This was confirmed by a Kolmogorov-Smirnov test (D = 0.10, p <.001). Skewness for this value was also somewhat problematic at 1.409. However, given that the sample size is relatively large and that a repeated measures design was used to better account for individual variability, it was considered appropriate to proceed with parametric analyses. Transformation was considered, however, it was considered that this would have significantly reduced the meaningfulness of the scales. Consequently, the analysis proceeded using parametric methods. Table 6 displays descriptive statistics for the recognition time data.

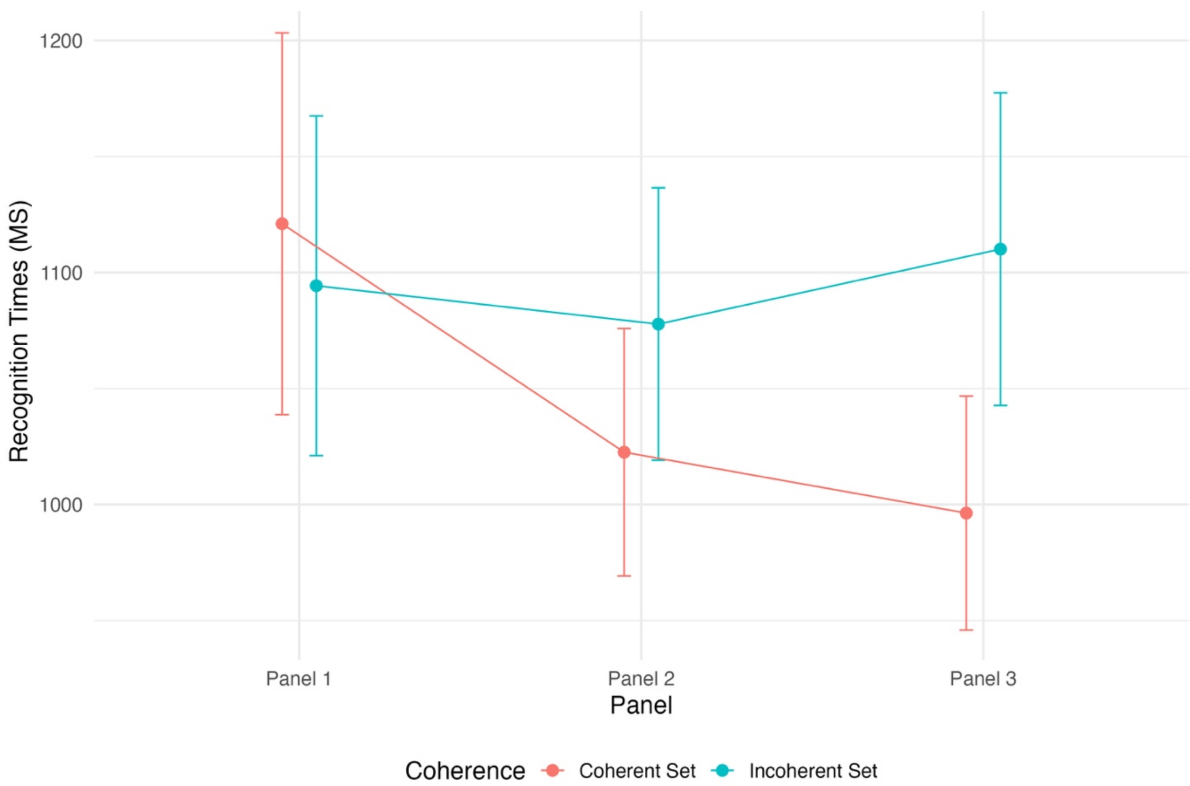
Table 6 *Recognition Time Descriptives by Coherence and Panel*

Condition	Panel 1	Panel 2	Panel 3
Coherent	1.12 (0.355)	1.02 (0.230)	0.996 (0.218)
Incoherent	1.09 (0.316)	1.08 (0.253)	1.11 (0.291)

Note. Means and (Standard Deviations) for coherent and incoherent panel set recognition times.

As with reading times, a 2 * 3 within-factors repeated measures ANOVA was conducted to assess visual narrative situation model processing. In this model, when controlling for panel position, coherence was found to be a statistically significant factor in recognition times, $F(1, 73) = 4.333$, $p = 0.041$, $\eta^2 = 0.0072$. However, this significance was marginal at 0.041 and was associated with a particularly small effect size. Unlike in the reading time data, the effects of panel position, when controlling for coherence, did not significantly influence the participant's recognition times of those panels $F(2, 133) = 2.850$, $p = 0.066$ (Huynh-Feldt corrected), $\eta^2 = 0.0089$. However, in contrast to this, the interaction effect of coherence by panel position was statistically significant $F(2, 146) = 3.807$, $p = 0.024$, $\eta^2 = 0.0105$, indicating that together, these two variables had a combined impact on recognition times. Figure 12 displays the recognition time estimated marginal means of recognition times for coherence by panel position.

Figure 12 *Effects of Coherence and Panel Position on Recognition Times*



Following the significant main effect findings of coherence and the interaction effect, Holm corrected post hoc analyses were conducted to explore these. Partial support was found for hypothesis five. This posited that, based on the situation model processing frameworks' effect on organising information into memorial units, if this was the case, then coherent comics should be recognised faster than incoherent comics. This hypothesis was supported by the overall coherent post hoc test, which indicated that coherent comic panels were recognised faster than incoherent comic panels $t(73) = -2.08$, $p = 0.041$, 95% *CI* [-0.09, -0.00], Cohen's $d = -0.24$, ($MD = -0.047$) and was associated with a small effect size. However, as with the main effect, the significance of this finding was considered to be marginal. Further post hoc comparisons generated following the significant interaction effect finding were less supportive of the hypothesis. The analysis revealed that there was no significant difference in recognition times between the first position coherent and incoherent panels $t(73) = 0.62$, $p = 1.000$, 95% CI [-0.10, 0.16], Cohen's $d = 0.07$, (MD = 0.027). Similarly, no significant difference was demonstrated between second position coherent and incoherent panels $t(73) = -1.59$, $p = 0.929$, 95% CI [-0.16, 0.05], Cohen's $d = -0.18$, (MD = -0.055). There was, however, a statistically significant difference between third (final) position coherent and incoherent panel recognition times $t(73) = -3.41$, $p = 0.016$, 95% CI [-0.21, -0.01], Cohen's $d = -0.40$, (MD = -0.114). Together, these results support the hypothesis that visual narrative information is both stored in a situation model and updated at necessary points.

Hypothesis six posited that, similar to reading times, there should be a downward gradient in recognition times that reflects the diminishing access to deactivated information. Again, only partial support was found for this hypothesis, but the evidence that is presented at least partially supports the suggestion of situation model updating in visual narratives. No significant difference was found between the first and second position coherent panels $t(73) = 2.81$, $p = 0.083$, 95% CI [-0.01, 0.20], Cohen's $d = 0.33$, (MD = 0.098). Similarly, no significant difference in recognition times was found between the second and third position coherent panels $t(73) = 0.93$, $p = 1.000$, 95% CI [-0.06, 0.11], Cohen's $d = 0.11$, (MD = 0.026). However, there was a significant difference between coherent first and coherent third position panels $t(73) = 3.33$, $p = 0.019$, 95% CI [0.01, 0.24], Cohen's $d = 0.39$, (MD = 0.125). This supports the hypothesis that situation model processing is applied to comics in that recognition was significantly improved for the final panel of a coherent comic.

The final hypothesis (7) was also explored using coherence through panel interaction post hoc tests. This hypothesis posited that there would be no significant difference between recognition times for any of the incoherent panels on the basis that these should not have activated situation model processing behaviours and thereby not been organised into a coherent memory structure. This hypothesis was fully supported by comparisons of incoherent panel recognition times. No significant differences were observed between first and second position incoherent panels $t(73) = 0.42$, $p = 1.000$, 95% *CI* [-0.10, 0.13], Cohen's $d = 0.05$, ($MD = 0.017$), second and third position incoherent panels $t(73) = -0.86$, $p = 1.000$, 95% *CI* [-0.15, 0.08], Cohen's $d = -0.10$, ($MD = -0.032$) or first and third incoherent panels $t(73) = -0.35$, $p = 1.000$, 95% *CI* [-0.15, 0.12], Cohen's $d = -0.04$, ($MD = -0.016$). These results clearly suggest that, by comparison to coherent comics (specifically in comparing the coherent and incoherent third panel recognition times), coherent comic panels achieved a privileged status in memory consistent with situation model processing behaviours. Conversely, the incoherent comics demonstrate parity in their accessibility, suggesting that these were not processed under the situation model processing framework.

3.1.5 Discussion

Primarily, this experiment aimed to verify that the short visual narratives used within the thesis adhere to the situation model processing framework. The secondary objective was to explore situation model updating in processing visual narratives and the utility of recognition times in assessing this concept. Reading times have widely been used to demonstrate varying aspects of situation model processing in text. Visual narratives have also previously been demonstrated to adhere to situation model processing using this metric (Cohn & Wittenberg, 2015; Brich et al., 2024; Cohn et al., 2024). On this basis, it was considered that the current stimulus would demonstrate similar processing behaviours. Several hypotheses were posed in relation to reading times: that coherent comics would be read significantly faster than incoherent comics; that there would be a downward trending gradient in coherent comic panel reading times; that incoherent comic panels would demonstrate a pattern indicating no situation model processing, and that processing initial panels of short comics would serve to demonstrate the construction of situation models by demonstrating parity between coherent and incoherent sets.

Updating is a critical feature of narrative processing. In situation model processing literature, recognition times are considered to demonstrate situation model updating behaviours. As readers progress through narratives, they are thought to update situation models, thus deactivating older content that is no longer necessary for the currently held model (Radvansky & Zacks, 2011; O'Brien et al., 1998). In visual narratives, this feature of situation model processing is underexplored. In this experiment, participants first viewed a series of coherent or incoherent comics and the amount of time taken to view each panel was recorded. They then completed a recognition task where they were asked to verify whether or not they had seen a presented comic panel during the reading time task. Given that situation models are an organising structure of memory which should have been generated when the comics presented maintained coherence, it was hypothesised that coherent comic panels would be recognised faster than incoherent comic panels. Secondly, it was considered that because of the sequential deactivation that occurs when situation models update, this would manifest as a similar downward trend in coherent panel recognition times.

The results of the reading time analysis largely confirm that the current short comic stimuli were processed under the framework of situation model processing. Coherent comic panel sets were read significantly faster than incoherent comics, and a downward gradient in reading times was observed, with each subsequent coherent panel being read faster, supporting hypotheses one and two. Additionally, there was no significant difference in reading times between the first panels of coherent and incoherent comics, supporting hypothesis four. However, the analysis revealed unexpected differences within the incoherent comics: the first panels were read significantly slower than subsequent incoherent panels. Interestingly, the second and third panels in incoherent comics showed no significant differences in reading times between themselves, but both were read significantly slower than their respective coherent panels. This indicates only partial support for hypothesis three.

Although the reading time data from this experiment strongly supported the application of situation model processing to short visual narratives, the recognition time findings were less conclusive in demonstrating updating. Coherent comic panels were recognised significantly faster than incoherent panels, supporting hypothesis five and further reinforcing the claim that visual narratives are processed within the situation model framework. This was additionally supported by the lack of a significant effect of panel position on incoherent panel recognition times, with consistent recognition times across the first, second, and third panels. However, unlike the reading time results, no gradient in recognition times was observed for coherent comic panels. Instead, a significant difference emerged between the final position panels and the first position panels in coherent sequences.

As expected, the reading time findings present several clear demonstrations of situation model processing behaviours in visual narrative processing. Firstly, that initial panels of sequences took significantly longer than processing the other panels in a sequence is consistent with much of the research on situation model processing (Gernsbacher et al., 1990; Gernsbacher & Faust, 1997; Cohn & Wittenberg, 2015; Brich et al., 2024; Cohn et al., 2024). This is a clear indication of readers' initial situation model construction, where they establish the context communicated by the first panel and subsequently activate related networks of information. The current findings also support the concept of facilitative prediction in visual narrative processing (Zacks et al., 2009; Reynolds et al., 2007; Swallow & Zacks, 2009). The downward trend observed in reading times supports the idea that readers establish the context of the comic from the initial panel. Once the context is established, activated cognitive networks facilitate prediction. These predictions are then confirmed by each subsequent panel, reducing the time readers need to process each successive panel.

The incoherent comics and their reading times also provided some unexpected findings. Visual language theory suggests that comics utilise structural grammar in communication, making them akin to language (Cohn, 2013; 2015). Specifically, the theory of Visual Narrative Grammar proposes that comics use a hierarchical constituent structure, which relates to the concept that panels, like words in a sentence, have grammatically correct positions in which their comprehensibility is maximised. This is demonstrated by studies using panel placement tasks or manipulating panel ordering (Cohn et al., 2024; Hagmann & Cohn, 2016). Cohn and Wittenberg (2015) report reading time findings that align closely with those observed in the current experiment while also highlighting an interesting contrast. Their results reveal a significant difference in reading times between coherent and incoherent initial panels, with incoherent panels being read significantly faster. They interpret this as evidence supporting visual narrative grammar, as it aligns with the concept of positional comprehensibility, thereby reinforcing the framework of visual narrative grammar.

In contrast, the current results indicate no significant difference between incoherent and coherent initial panels. In this experiment, this is taken to indicate that situation models are constructed from any initial panel regardless of coherence. This appears to be incongruous with the suggestion of visual narrative grammar. However, as in language, visual narrative grammar further suggests that comics are better understood by those with more experience or fluency in comics, making them more efficient visual narrative processors (Cohn, 2020). This likely explains the discrepancy in findings here. Cohn & Wittenberg's (2015) participants, who were specifically selected for their comics reading ability, had an average VLFI score of around 12, which is considered average overall. Contrastingly, the participants in this study had an average VLFI score of 7, which is considered low. Arguably, the difference in findings between these results and those of Cohn & Wittenberg (2015) actually supports the theory of visual narrative grammar on the basis that higher comic book reading fluency seems likely to be related to positional comprehensibility.

The recognition time findings reported in this experiment, while providing a less conclusive picture of updating during visual narrative comprehension, present some evidence that visual narratives adhere to situation model processing. Firstly, the results evidence the suggestion of situation model theorists that throughout a narrative, representational situation models are updated with new information, causing old information to become deactivated (Zacks & Bailey, 2015; Pettijohn & Radvansky, 2016). Support for this comes from the finding that the final panels were recognised significantly faster in coherent comics than the first panels. This demonstrates that as the narrative progresses and new information is introduced, readers are more likely to deactivate older information in favour of maintaining the most recent events.

The findings then are suggested to align with the principles of situation model processing where comprehension processes dynamically affect memory and the current processing model to reflect the contemporary state of the narrative. However, this interpretation should be tempered with some caution on the basis that the coherent second panels were not found to be significantly different in recognition time by comparison to the final panels. The lack of significant difference suggests that updating of situation models may not occur in a strictly linear fashion across all components of the narrative. This again would suggest support for the theory of visual narrative grammar in that certain hierarchical components of a visual narrative may cause updating and, thus, deactivation of content more potently than others. Overall, the results support the idea that visual narrative comprehension follows the framework of situation model processing. However, the findings of this study suggest that the specific dynamics of visual narrative comprehension may require adjustments to this theory. These adjustments would account for the unique updating and maintenance processes involved in interpreting the visual nature of comics.

While the current stimuli were not designed to test the concepts of incremental or global updating in visual narrative processing, the exploratory recognition time findings provide some perspective on these concepts. Arguably, the recognition time results support the suggestion of Kurby & Zacks (2007) that both incremental and global updating occurred throughout comprehension, and, interestingly, both of these updating procedures are demonstrated even in these short visual narratives. That there was an overall difference between coherent and incoherent panel recognition indicates that updating did occur throughout processing. Had significant differences been found for recognition times between each coherent panel, this would have demonstrated clear support for the concept of panel-based global updating. Alternatively, had there been a significant difference between coherent and incoherent panels overall but no difference between any coherent panels, this pattern of results would suggest more definitive support for incremental updating. The actual pattern of recognition time results then demonstrates support for both updating types. The partial lack of a significant difference between coherent comics could then be interpreted as showing incremental updates between panels one and two, as well as two and three, but an overall global update in content between panels one and three.

These findings contribute to a growing body of evidence that suggests that situation or event model processing behaviours are robust across different modalities. The reading time findings demonstrate this clearly and align with previous experiments using this metric with static visual narrative stimulus to demonstrate narrative comprehension processes (Cohn & Wittenberg, 2015; Cohn et al., 2024; Cohn & Paczynski, 2013). Particularly striking is the finding that these processing behaviours are exhibited when the narratives they are derived from are only three panels long. Furthermore, the reading time findings also contribute to the suggestion of Cohn and Paczynski (2013) that visual narrative comprehension is likely to be facilitated, as in other narrative comprehension formats, by narrative prediction.

Moreover, the experiment contributes to the general finding that recognition times have significant utility in explicating processes of narrative comprehension in static visual narratives. This is particularly noteworthy because, while recognition times are typically associated with basic visual processing, these findings suggest they can also capture complex dynamic processes involved in visual narrative comprehension. This is surprising given that visual narrative processing has been demonstrated to be performed incredibly quickly (Potter et al., 2014; Hagmann & Cohn, 2016) and thus may not have been considered sensitive enough to capture higher-order cognitive processes related to visual narrative comprehension.

Some limitations of the current experiment should also be considered. As with any laboratory-based experiment, the ecological validity of the methods used in this study should be considered when interpreting the results. When reading visual narrative-based content in a normal environment, readers would not be presented with each panel singularly but as full strips or combinations of panels presented as pages in comic book materials. Previous work that analyses gaze behaviours has demonstrated, for example, that comics readers make numerous regressions or backward eye movements to previous panels and that, typically, static visual narrative sequences are not naturally processed in a linear sequence (Foulsham et al., 2016). While this method of presentation was necessary to the current experiment in investigating explicit situation model processing and narrative updating behaviours, it seems likely that this somewhat unnatural presentation would have affected the participants' processing of the narratives. To this end, this has been altered in subsequent experiments which explore the use of working memory in visual narrative situation model processing.

Secondly, and regarding future research, recognition times have been presented here as an approximate measure of the updating process, and the rationale for this is grounded in metrics commonly employed in previous studies. Arguably, they have successfully demonstrated that these behaviours occur in visual narrative situation model processing to some degree. However, this simple behavioural metric cannot capture the full complexity and nuance of updating behaviours. Future research could expand on this basis by incorporating recognition times with additional measures such as eye tracking (Swets & Kurby, 2016) or think-aloud protocols (Kurby & Zacks, 2011) to capture more of the online process of updating during visual narrative comprehension.

In summary, this experiment has produced several findings related to situation model processing behaviours in visual narrative comprehension. Firstly, they show that prediction, guided by situation model processing, is important in understanding static visual narratives. The experiment similarly contributes to a growing body of literature, which suggests that panel reading times can be used in visual narrative processing research to explore processes associated with comprehension. Moreover, it contributes findings related to updating behaviours, suggesting that incremental and global updating types are utilised in processing visual narratives. With the confirmation that the current stimuli are processed under the situation model processing framework, the thesis now turns towards the use of working memory components in the situation model processing of visual narratives.

Chapter 4

4.1 Experiment 5 – Visuospatial Working Memory in Situation Model Processing

4.1.1 Introduction

Experiments 1 to 3 of this thesis identified that visuospatial working memory underpins visual narrative processing. The fourth experiment then determined that comics, and specifically the current comics stimuli, are processed under the situation model processing framework. However, the extent to which visuospatial working memory is responsible for processing visual narratives under the situation model comprehension framework needs to be clarified. It is widely suggested in the literature on narrative comprehension that working memory plays a crucial role in creating, maintaining and updating situation models. However, literature that has explored the use of working memory in situation models has predominantly focused on text-based narratives. In contrast, the functional role of working memory in visual narratives and its role in comprehension of this medium remains largely underexplored. Consequently, this experiment aims to investigate how working memory contributes to visual narrative situation model processing.

As is outlined by the previous experiment, comprehension of narratives is theorised to be achieved by the creation of situation models (Kintsch & van Dijk, 1988; Zwaan & Radvansky, 1998; Kurby & Zacks, 2008; Radvansky & Zacks, 2014). Based on the initial Construction Integration model (Kintsch & van Dijk, 1988), which informs the Event Indexing model (Zwaan

& Radvanksy, 1995), Event Segmentation Theory (Zacks et al., 2007; Zacks & Swallow, 2007) and the Scene Perception and Event Comprehension theory (SPECT: Loschky et al., 2020), situation models are outlined as the representational level of narrative comprehension. Creating a situation model requires integrating the initial surface form levels, where the basic units of information are perceived, and the semantic levels, where the meaning of the basic surface level is decoded using associatively activated networks in long-term memory. It has been widely demonstrated and is widely agreed upon, that information presented in a narrative is remembered as the representational situation model and not the surface-level perceptual features of the narrative (Morrow et al., 1987; Rinck et al., 1997; Haenggi et al., 1995; Bower & Morrow, 1990; Glenberg et al., 1987; Taylor & Tversky, 1992).

Another widely accepted feature of situation model processing is that working memory plays a critical role in the representational model constructed by the process of narrative comprehension (Kintsch & van Dijk, 1978; Gernsbacher et al., 1990; Zwaan et al., 1995; Radvansky, 2012; Zacks et al., 2007; Zacks & Swallow, 2007; Cohn, 2020; Loschky et al., 2020). As working memory is essentially considered to be a framework for conscious processing, it is both intuitive and logical to suggest that this activated portion of memory in which information can be maintained and manipulated, as well as communicated bidirectionally with long-term memory, is responsible for the maintenance of situation models derived from narratives. However, while there is agreement on the importance of the working memory framework to situation model processing, currently available empirical evidence provides a somewhat mixed picture in relation to its precise utility in situation model processing.

Multiple studies on situation model processing reference the concept of working memory and interpret their findings in relation to its role. However, many studies rely on theoretical frameworks or indirect evidence to support their interpretations instead of empirically testing this construct through span tasks, dual-task paradigms, or other measures specifically designed to assess working memory. While this is not an unreasonable approach, studies that do not directly measure working memory generally have plausible grounds for relating their findings to this construct.

For instance, theoretical frameworks such as the eye-mind hypothesis (Hutson et al., 2018) suggest that the current contents of vision align with what is active in working memory, while research on age-related declines in working memory (Zacks & Bailey, 2015) is frequently cited to explain associations between age and situation model processing behaviours. Indirect evidence from these studies indicates that working memory content influences visual attention selection. Furthermore, age-related declines in working memory capacity are linked to a tendency among older adults to rely more on global updating strategies rather than incremental ones. Additionally, through meta-analysis, working memory has been shown to play a critical role in language comprehension processes (Daneman & Merikle, 1996; Daneman & Carpenter, 1980). Since situation model processing provides a framework for understanding narratives, assuming a connection between working memory and situation model processing is reasonable. However, if these constructs are as closely related as suggested, empirical methods specifically designed to test the role of working memory in situation model processing are necessary to validate this assumption.

Few studies have demonstrated a direct link between working memory and situation model processing. Sargent et al. (2013) provide the most compelling evidence for a direct link. These authors were interested in whether there was a causal relationship between memory for events and event segmentation behaviour. Their participants viewed videos depicting an actor performing an everyday task and were asked to segment the actions into broad and narrow events across two viewings. Participants were then administered a battery of cognitive tests, including working memory measures. The results of a structural equation model subsequently demonstrated that working memory capacity positively predicted event segmentation ability and thus was found to affect memory for events indirectly. Age was also demonstrated to negatively predict working memory, which supports the suggestion that age-related declines in working memory affect situation model processing.

Further direct support for the use of working memory in situation model processing comes from a body of evidence which explores the creation of spatial situation models. This relates to the concept that we use mental representations to map and organise spatial communication. Research in this area has reported that working memory components, delineated by the multi-component theory into visuospatial and verbal components, demonstrate selective interference effects (De Beni et al., 2005). Where articulatory suppression, a task designed to disrupt verbal processing, has been shown to affect both spatial and non-spatial text-based narratives, a spatial tapping task was shown only to affect spatial texts. Similar findings suggest that where verbal working memory is critical to the verbal or phonological input, visuospatial working memory is utilised to develop a conceptual mental map (Brunye & Taylor, 2008; Pazzaglia et al., 2010). In turn, this line of research suggests that the situation model level of representation is differentially affected by the format of information (Picucci et al., 2013) and utilises working memory accordingly.

Though there is a reasonable body of evidence for the suggestion that working memory underpins situation model-based processing, other research investigating span measures in relation to the use of situation model processing has produced null findings. Radvansky and Copeland (2001) examined the formation of situation models and the role of working memory in reading tasks by utilising working memory span measures to investigate its influence on situation model updating. Their findings revealed that participants were able to successfully create and update situation models, with objects deemed currently activated and associated with a narrative protagonist being more readily accessible than those that had been deactivated, a finding that mirrors classic accessibility findings in situation model processing (Haenggi et al., 1995). However, the results also indicated that this ability to create and update situation models was not directly related to working memory span. Similar methods then produced results that demonstrated that reasoning and situation model processing were not related to working memory span capacity (Radvansky & Copeland, 2004). Additionally, Thompson and Radvansky (2012) found that a working memory distractor task did not affect the accessibility of current situation models. Taken together, these findings suggest that in order to countenance literature suggesting that working memory is critical to language comprehension, Radvansky and colleagues interpret these findings as an indication that working memory is more likely to be associated only with the initial surface level of situation model processing (Radvansky & Copeland, 2006).

As suggested above, theories that account for narrative and event model processing widely assume a critical role of working memory. The account of situation model processing provided by the Scene Perception and Event Comprehension Theory (SPECT: Loschky et al., 2020) most appropriately characterises this concept for visual narratives situation model processing. This theory applies visual cognition and narrative processing principles to static and dynamic visual narratives. Broadly speaking, SPECT describes front-end processes, which are used to register input through attentional selection, leading to information extraction. These processes represent the basic input surface-level processing described in the Construction-Integration (CI) theory of narrative processing (Kintsch & Van Dijk, 1978). The information extracted by these Front-End processes is then fed backwards into working memory where, it is suggested, the current situation model is constructed, maintained, updated and then shifted from, as necessary, throughout narrative processing. These latter processes are described as back-end and are bidirectional communication mechanisms with long-term memory and associatively activated semantic networks that resolve information passed from front-end processes into a meaningful representation. These then effectively describe the semantic layer of processing of CI and resultant situation model levels.

The combination of SPECT's comprehensive framework (Loschky et al., 2020) and Radvansky & Copeland's (2001; 2004; 2006) work on the utility of working memory in situation model processing emphasises the nuance and complexity in understanding the functional cognitive mechanisms that underpin narrative processing. However, together, they indicate a potential discrepancy regarding the level of processing at which working memory is most involved. Where SPECT, as well as other narrative or event processing theories (Zwaan & Radvansky, 1995; Zacks et al., 2007; Zacks & Swallow, 2007), suggest that working memory is primarily responsible for the constructed representation level, the available empirical evidence on the direct role of working memory in situation model processing suggests that this memory system is perhaps more strongly associated with textbase, Front-End, or perceptual input processes than the representational construction.

Currently, only one paper has explored the role of working memory in visual narrative processing. Magliano et al. (2016) explore this in relation to the use of inferences during visual narrative processing and demonstrate that working memory does appear to be implicated in this process. However, as discussed, the method of this paper contains limitations related to a lack of titration and the use of reading times to approximate working memory component involvement. Secondly, as outlined above, while SPECT explains the use of working memory in visual narrative situation model processing, the available evidence is mixed in relation to the role of working memory in situation model processing, with some research suggesting it may be less centrally related to the representational level. As such, further research is needed to clarify the contribution of working memory to situation model processing, particularly in the context of visual narratives. This study then aimed to address these gaps by investigating the use of working memory in visual narrative processing with a direct measure of working memory in simple visual narratives.

The methodology used in this experiment combined the methods of the experiments presented in chapters 2 and 3. In this experiment, participants were exposed to both coherent and incoherent comics as part of a dual-task experiment. They were required to maintain a visuospatial working memory load while processing a comic. Situation model processing is considered to be a naturally automatic process (Zwaan & Radvansky, 1998; Zacks et al., 2007; Zacks & Swallow, 2007), meaning that regardless of whether a coherent situation model can be created, a reader will naturally attempt to form a situation model from any given narrative content. Situation model processing is also considered to have specific cognitive costs. This is established by studies which demonstrate that construction and shifting require significant cognitive effort on the basis that reading times are consistently demonstrated to increase at these narrative processing points. This forms the basis of the hypotheses for the current experiment.

Experiment 4 has demonstrated both the automatic application and the cognitive costs that characterise situation model processing in the visual narrative stimuli used in this thesis. Similarly, Experiment 4 has indicated that coherence facilitates reading times, which can logically be interpreted as these sequences representing a lower cognitive load. However, the greater the discontinuity presented in a narrative, the more burden is placed on cognitive processing as resolution of the narrative sequence is attempted. This is demonstrated by the finding of Experiment 4 that incoherence in sequences increases reading times. Following this, the overarching hypothesis of this experiment was that if working memory is responsible for situation model processing and creation, then narrative incoherence should significantly increase visuospatial working memory load. If this increased cognitive load is related to working memory, then incoherent narratives should increase working memory load as the reader attempts to process the narrative into a mental representation.

As in the first set of experiments, the dual-task paradigm applied here was used under the framework that if processes share resources, then the performance of a simultaneous maintenance task will be degraded due to the diversion of cognitive resources towards having to process the narrative. As such, it was hypothesised that, in line with the findings of Chapter 2, processing a comic, coherent or otherwise, should increase visuospatial working memory load, resulting in a detriment to performance in a concurrent visuospatial dual-task. Furthermore, it was hypothesised that if visuospatial working memory serves the situation model level of processing, then incoherent comics will cause significantly more visuospatial working memory dual-task performance than coherent comics on the basis that incoherent comics disrupt the ability to create a situation model.

4.1.2 Method

4.1.2.1 Participants

41 undergraduate psychology students (34 female) at Sheffield Hallam University participated in this study in exchange for course credit. Their average age was 19 (*SD* = 3.5). All participants had normal-to-corrected vision. All participants gave informed consent to participate in the study and were naïve to the purposes and intentions of the study. The Sheffield Hallam University Research Ethics Committee approved the experiment.

4.1.2.2 Materials

This experiment used the standardised square panel comic stimuli as described in the previous experiment. These were also selected and generated in the same way as described in Experiment 4. The coherent comics were then, for all intents and purposes, randomly selected, and the incoherent comics were randomly generated. Because of the requirement to create random comics which used comic panels from unique narratives, as well as the necessity for practice trials, a total of 4 coherent and 4 incoherent comics were used. These were matched by 4 control task trials.

The python-programmed Corsi block task, as described in Experiment 1, was used again in this experiment, with all timings remaining the same. This presented blocks in randomised places on the screen, which then flashed in a specific order that the participants were required to remember.

This iteration of the test implemented some minor usability upgrades. In this version, the participants could unclick the buttons during the response phase, undoing accidental block clicks. The blocks also displayed a number indicating the order in which the participant had clicked them, allowing them to check their responses.

4.1.2.3 Measures

As in Chapter 2, the VLFI was used to check the parity of comics experience between the participants, and this is presented in the following experiment. This was particularly relevant in this study, given that previous findings have demonstrated that people with greater exposure to comics tend to have a higher tolerance for narrative incoherence within comic panels (Hagmann & Cohn, 2016).

4.1.2.4 Design

This experiment used a repeated measures design with three conditions: coherent comics, incoherent comics, and a control condition. Accuracy on the Corsi block test served as the dependent variable. All participants completed the working memory titration procedure to establish individualized visuospatial interference levels, and the same positional accuracy scoring procedure from Experiments 1–3 was applied.

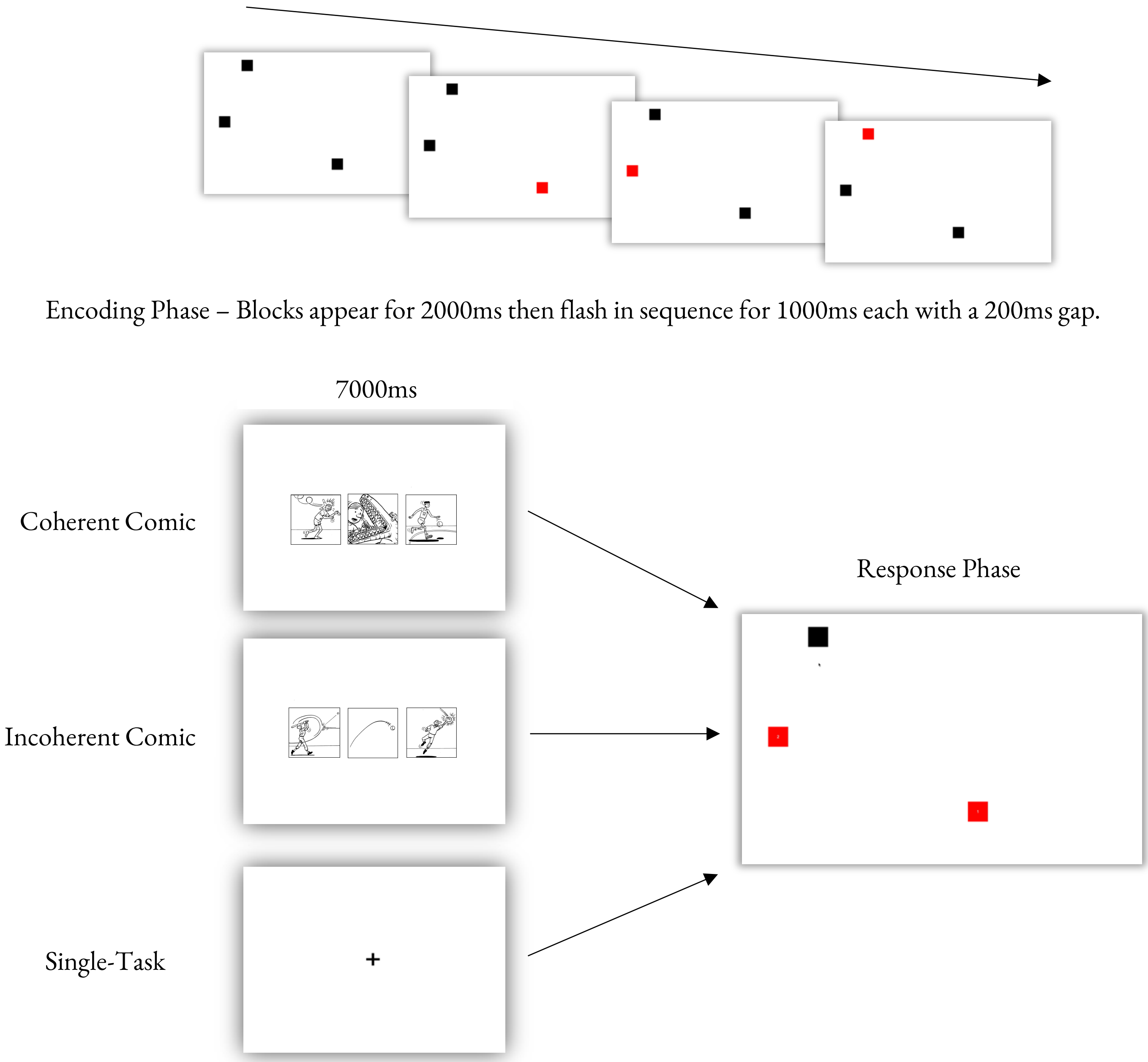
4.1.2.5 Procedure

Participants completed the experiment in booths designed to reduce distraction. After reading the information sheet and providing consent to participate in the study, they were asked to provide their age, gender description, and email address. As with the other experiments, this information was only used to provide debriefing information and was not stored in accordance with university data protection regulations.

Participants then completed the titration procedure of the Corsi block task. In brief, during this phase, the participants were required to complete increasingly difficult trials of the Corsi block test. If a participant failed two trials at any level, the titration phase ended, and their interference level was set to the previous level, at which they scored 100% on at least two trials. At the offset of the titration phase, the participants were presented with the Visual Language Fluency

Index in a computerised format. After the participants had completed the VLFI, the dual-task phase began. Figure 13 depicts the procedure of Experiment 5.

Figure 13 *Graphical Representation of Procedure in Experiment 5*



This phase began with practice trials, demonstrating the experiment's dual-task procedure under the Coherent, Incoherent, and Control conditions in that order. After the participants completed the practice trials, they completed four blocks of the three trial types: coherent comic, incoherent comic, and control. Each block consisted of one dual-task trial of each condition. The trial presentation order within the blocks was randomised programmatically so that participants completed a coherent, incoherent, and control trial task in a randomised order.

Coherent comic trials were dual-task trials in which the participants completed the Corsi block task with a coherent three-panel comic presented for 7000ms between seeing and repeating the pattern. Incoherent comic trials followed the same procedure using incoherent comics. Control trials were single-task trials in which the participants were simply required to maintain the Corsi block pattern for 7000ms.

Given the limitations outlined in relation to Experiments 1 – 3 and the lack of a structured online comprehension check, following any dual-task trial which contained a comic stimulus, either coherent or incoherent, the participant was asked to rate the coherence of the comic on a one to seven Likert-style response. These data were used to assess the participant's engagement with the material and to screen any participants who did not appropriately engage with the comic material while maintaining the visuospatial working memory sequence. Additionally, following any coherent comic strip dual-task trial, the participants were asked a true-or-false question about the comic strip to ensure engagement. This approach aimed to provide direct evidence of participants consistent attention and narrative understanding.

4.1.3 Results

4.1.3.1 Data Screening and Manipulation Check

Initially, the data was screened by summing the number of correct answers the participants registered following the coherent dual-task trials. One participant scored less than 50% on the coherent comic questions, and this was taken to indicate significant inattention to the dual-task stimulus. Their data was subsequently removed from the data set, leaving 40 participants' data for analysis.

A non-parametric Paired Wilcoxon signed rank test was conducted to assess the participants' responses to the coherence attention check question. These responses were gathered after the participants had completed the dual-task trials and related to the perception of how coherent the comic stimulus they had just seen was. Participants responded by rating the stimulus on a Likert-style scale ranging from 1 (Completely Incoherent) to 7 (Completely Coherent). The non-parametric variant of this test was selected because of both the ordinal nature of the responses and highly skewed distributions.

The results revealed that the coherence manipulation had been successful. The median rating of coherent comics was 27 (out of a possible 28), while the median incoherent comic rating was 8 (out of a possible 4). The difference between these two rating averages was found to be statistically significant $W = 741$, $p < .001$, and was associated with a large effect size: $r = 0.87$. This indicates that the coherence manipulation was successful, supporting the suggestion that the participants engaged with both concurrent tasks to a reasonable degree.

4.1.3.2 Dual-task Accuracy Assumptions

The participants' accuracy on the coherent, incoherent and control trials was averaged to produce mean scores for each condition. A one-way repeated measures ANOVA was then used to test for differences in dual-task accuracy caused by the variation in coherence. The residuals of this model were generated and assessed against the parametric assumptions. Residual skewness was found to be well within the acceptable range (-.37). Further to this, the Kolmogorv-Smirnov test of normality demonstrated that the distribution of the residuals did not significantly differ from that of a normal distribution ($D = 0.09$, $p = 0.18$). Consequently, the analysis proceeded using parametric tests. The assumption of sphericity was also satisfied with Mauchly's test, indicating no significant violation, ($W = 0.906$, $p = 0.153$).

4.1.3.3 Dual-task Accuracy Analysis

The main effect of the ANOVA model revealed a statistically significant effect of comic type on dual-task accuracy $F(2, 78) = 6.633$, $p = 0.002$, and this was associated with a medium effect size $\eta^2 = 0.0670$. Post hoc tests were conducted to assess more precisely which conditions were significantly different from each other. Table 7 displays the means and standard deviations for accuracy in each condition.

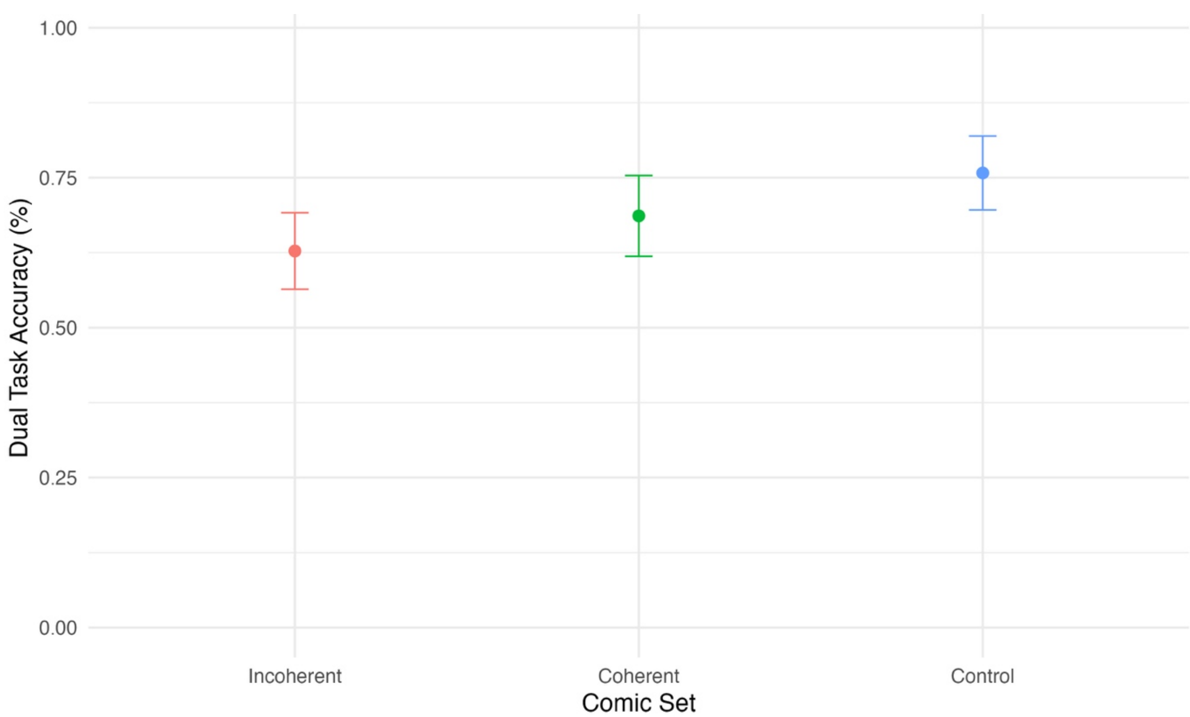
Table 7 *Corsi Block Accuracy Descriptives by Coherence*

Condition	Mean	SD
Control	0.758	0.193
Coherent	0.686	0.211
Incoherent	0.628	0.199

Note. Means and Standard Deviations (SD) for visuospatial working memory performance across conditions. Means are expressed as percentages, I.E .757 = 75.7%.

Three, Holm-corrected post hoc tests were conducted to assess for differences in visuospatial dual-task accuracy between the incoherent, coherent and control task conditions. Figure 14 displays the estimated marginal means for dual-task accuracy of each condition.

Figure 14 *Effects of Coherence/Control Condition on Corsi Block Accuracy*



Accuracy following the control condition was found to be approximately 7% higher in comparison to the coherent comic condition (*MD* = -0.072), and this difference was determined to be statistically significant $t(39) = -2.38$, $p = 0.044$, 95% CI [-0.15, 0.00], and associated with a small to moderate effect size: Cohen's $d = -0.38$. Further to this, incoherent comics also significantly affected the participants' ability to maintain a visuospatial working memory sequence $t(39) = -3.29$, $p = 0.006$, 95% CI [-0.23, -0.03] with the control condition being performed with approximately 13% ($MD = -0.130$) greater accuracy than the incoherent comic condition. This effect was determined to be moderate Cohen's $d = -0.52$. These findings then support the first hypothesis that visuospatial dual-task maintenance would be significantly affected by the requirement to process a comic. Consequently, this supports the suggestion that visuospatial working memory is implicated in visual narrative processing. Further to this, these findings support those of Experiment 1 in suggesting that visuospatial working memory contributes significantly to visual narrative processing.

However, the secondary hypothesis did not find support. While there was an overall average reduction in accuracy following the incoherent and coherent conditions of approximately 6% ($MD = -0.059$) that was similar to the reduction between control and coherent comics, this difference was not found to be statistically significant $t(39) = 1.58$, $p = 0.122$, 95% CI [-0.03, 0.15], Cohen's $d = 0.25$. In plain terms, this result indicates that visuospatial working memory was no more affected by a coherent comic than an incoherent comic.

4.1.3.4 VLFI Scores and Corsi Block Correlations by Coherence Condition

Pearson correlations were conducted to examine the relationship between visuospatial working memory accuracy and VLFI scores across the three task conditions: coherent comics, incoherent comics, and single-task control. The mean VLFI score of the participants was 8.66 (*SD* = 6.23). For the coherent comics condition, the correlation between accuracy and VLFI was small and not statistically significant, $r(38) = .09$, $p = .591$, 95% CI [−.23, .39]. For performance in the incoherent comics condition, the correlation was near zero and also non-significant, $r(38) = .02$, $p = .901$, 95% CI [−.29, .33]. In the single-task control condition, the correlation was somewhat larger but remained non-significant, $r(38) = .18$, $p = .258$, 95% CI [−.14, .47]. In contrast to the findings of experiment 1, the results indicated no clear relationship between VLFI and visuospatial working memory task accuracy under any condition.

4.1.4 Discussion

This experiment aimed to investigate the contribution of visuospatial working memory to visual narrative situation model processing. The experiment posed two hypotheses. Firstly, based on the findings presented in Chapter 2, it was hypothesised that the requirement to process a comic would significantly impact visuospatial working memory maintenance. Secondly, it was hypothesised that if visuospatial working memory serves the situation model level of processing, then incoherent comics, which disrupt situation model creation, should generate greater interference to visuospatial working memory than coherent comics.

The major findings of this experiment can be summarised succinctly as follows: First, coherent comics significantly affected visuospatial working memory maintenance compared to a single-task control. Second, incoherent comics also significantly affected visuospatial working memory maintenance compared to a single-task control. Third, coherent and incoherent comics were not significantly different in the degree to which they affected visuospatial working memory.

The results of this experiment show that both coherent and incoherent comics significantly interfered with the participants' ability to maintain a visuospatial working memory sequence. This suggests that visuospatial working memory was actively engaged by both tasks, as evidenced by the decline in performance on the visuospatial dual-task by comparison to the single-task control. When participants processed a comic while simultaneously maintaining a Corsi block sequence, the demands on visuospatial working memory were heightened, leading to diminished task performance. These findings, consistent with those from experiments one to three, support the hypothesis that visuospatial working memory plays a key role in processing visual narratives. However, further findings of this experiment demonstrate that current models which describe the specific utility of working memory in this process may require some reconsideration.

On the other hand, the second hypothesis posed in this experiment was not supported. The third finding outlined above - that coherent and incoherent comics caused no significantly different level of interference to the participants' ability to maintain a visuospatial working memory sequence - indicates that the manipulation of coherence did not notably affect the visuospatial working memory resources required in maintaining the dual-task sequence. This finding would suggest that working memory may not be as critical to deeper levels of visual narrative processing as indicated in current theoretical frameworks outlining processes of visual narrative comprehension. In contrast to the suggestions of SPECT (Loschky et al., 2020), which outlines that the use of working memory as critical to the generation and subsequent maintenance of current visual narrative event models, these results more closely align with the interpretations of Radvansky & Copeland (2001; 2004; 2006) which suggest that the primary use of working memory is perhaps more related to surface level, perceptual feature registration.

This gives rise to an alternative explanation for the current findings concerning the degree to which surface-level processing in visual narratives may interact with situation model construction. Given that comics use directly depictive symbols to convey narratives, they are distinct from text-based narratives in the degree to which the perceptual surface may promote access to narrative meaning. One possibility in relation to this is that the iconic nature of comics reduces the cognitive demands typically required to build a situation model from more abstract inputs, such as text. This visual immediacy could reduce the reliance on visuospatial working memory for deeper narrative integration, thereby limiting the need for extended manipulation or maintenance of information over time.

In this sense, the results do not imply that visuospatial working memory directly contributes to constructing the situation model itself, but rather that it is primarily engaged in surface-level, perceptual processes consistent with the 'front-end' component outlined in the SPECT model (Loschky et al., 2020). Consequently, the lack of a differential effect between coherent and incoherent comics on dual-task performance may be explained by the relative ease with which narrative meaning can be extracted from iconic visual sequences, without extensive use of working memory resources for deeper inferencing.

Under this interpretation, the current findings refine rather than reject the assumption that situation model construction operates across narrative modalities. While the principle of equivalence (Cohn, 2020) suggests that similar comprehension structures may apply across formats, the process by which they are accessed or constructed may differ. Specifically, in comics, the surface-level representations may be sufficiently rich and informative to support narrative understanding with minimal involvement from working memory beyond initial perceptual registration.

A weakness of the current experiment that should be considered is the lack of account for the inference required to process the current comics stimulus. While all of the comics took place over ambiguous amounts of time, some of the events evidently took longer than others. Sailing a ship to an island, for example (Appendix M), clearly takes longer than firing an arrow into a target (Appendix A). Bridging inferences are a critical element of static visual narrative sequences as comics, unlike film or dynamic visual narratives, cannot depict all, or even the majority of a given scene or sequence of events. This means that readers must piece the narrative together themselves. On the basis of the account above, inference seems to be a likely use of visuospatial working memory (Magliano et al., 2016; Brich et al., 2024), particularly given the role of working memory in manipulating as well as processing information. The findings of this experiment indicate that visuospatial working memory significantly contributes to visual narrative processing. While the coherence manipulation did not cause a differential impact to visuospatial working memory maintenance, implying that visuospatial working memory is not related to deeper semantic levels of processing, the method of this study is arguably not conclusive in the suggestion that inference during visuall narrative processing is not also underpinned by visuospatial working memory. Future research should combine visuospatial working memory measures with comics that are more systematically varied in coherence or manipulate specific narrative components to explore this concept further.

In conclusion, this experiment supports the position that visuospatial working memory plays a significant role in processing visual narratives. However, its contribution to the representational situation model processing level may be less central than previously thought. Though both coherent and incoherent comics caused significant interference to visuospatial working memory maintenance, the lack of a differential effect between the two conditions suggests that in visual narratives, working memory may have greater involvement in processing and registering the surface-level perceptual features than the degree to which it is implicated in the situation model representation. The findings suggest that in visual narratives, and particularly comics, where the medium directly depicts events, the basic units of comics potentially act as both the surface perceptual features and situation models simultaneously. This potentially causes a reduction in working memory for higher-level narrative processing. Overall, the findings indicate that for situation models in visual narratives, medium-specific processes should be considered, which may have further ramifications for comprehension of visual narrative information.

4.2 Experiment 6 – Verbal Working Memory in Situation Model Processing

4.2.1 Introduction

The findings of Experiment 5 suggest that visuospatial working memory does not appear to be implicated in visual narrative situation model processing. This was on the basis that there was no significant difference in participants' ability to maintain a visuospatial working memory sequence while simultaneously processing coherent or incoherent comics. This leads to the conclusion that working memory may play a more limited role in visual narrative situation model processing than current theories suggest (Loschky et al., 2020). In the current experiment, the contribution of verbal working memory in visual narrative situation model processing is considered using the same paradigm as experiment 5. The results are then compared to the use of visuospatial working memory.

As discussed in the introduction to experiment 5, working memory is considered to be critical to situation model processing (Gernsbacher et al., 1990; Zwaan & Radvansky, 1995; Zacks et al., 2007; Zacks & Swallow, 2007; Loshcky et al., 2020). However, several debates are still yet to be resolved in relation to the specific utility of this cognitive framework in narrative processing. Despite the most prevalent model of working memory suggesting a multi-component architecture which divides working memory into visuospatial and verbal components, theories such as SPECT (Loschky et al., 2020) do not account for this consideration. However, despite the widespread acceptance of the multi-component architecture of working memory, SPECT (Loschky et al., 2020) does not explicitly differentiate between verbal and visuospatial components in its account of narrative processing. This omission limits its ability to incorporate evidence that highlights distinct cognitive mechanisms supporting these separate modalities. In light of the substantial body of research demonstrating clear functional and structural distinctions between verbal and visuospatial working memory (Izmalkova et al., 2022, Kruley et al., 1994; Gyselinck et al., 2002; Gyselinck et al., 2007; Shah & Miyake, 1996; Cockcroft, 2022; Vernucci et al., 2021) the lack of such differentiation in SPECT presents a something of a theoretical gap. Specifically, the model does not adequately address how these separable components might independently or interactively contribute to different aspects of narrative comprehension. Addressing this limitation could yield a more nuanced framework that better aligns with empirical evidence from studies emphasising the specialised roles of verbal and visuospatial working memory in various cognitive tasks. The findings of this thesis point towards the conclusion that, in visual narrative processing, visuospatial working memory has a specialised role that relates to 'front end' (Loschky et al., 2020) or perceptual processes and thus far, verbal working memory has not been shown to contribute to visual narrative processing.

Inference is a crucial aspect of all narrative processing. A distinction made in the literature, and one that warrants clarification here, is between mapping and inference. Mapping typically refers to the alignment of new narrative information with an existing situation model, supporting the continuity of events, actions, or characters. In contrast, inference is generally defined as the active process of filling in gaps in a narrative by drawing on prior knowledge, context, or expectations. Importantly, inference is often considered to be triggered by explicit disruptions or omissions. While these processes are described as distinct, this separation may reflect a methodological necessity as opposed to a meaningful cognitive division. In practice, these processes likely exist along a continuum, with the degree of narrative coherence influencing the extent to which inferential effort is required. Even in well-structured sequences, mapping seems likely to require inferential processes, even if they might be considered low level, to establish temporal, causal, or spatial links between panels.

These processes are critical to visual narrative comprehensiona and this has been evidenced by event-related potential studies (Cohn, 2021), which identified electrophysiological markers of semantic incongruence during comic reading, particularly when critical panels were substituted with action stars. These findings demonstrate inference generation in visual narratives because readers actively work to resolve the disruption caused by missing or incongruent information. When a critical panel is replaced, the coherence of the narrative is disrupted, creating a gap in the situation model that the reader must fill using prior knowledge, contextual cues, and predictive reasoning. The observed electrophysiological markers, such as increased neural activity associated with semantic processing, reflect the cognitive effort involved in generating inferences to integrate the altered or missing content into a coherent narrative framework. This process highlights the active and dynamic nature of visual narrative comprehension, where inference serves to maintain narrative continuity despite disruptions. Similarly, visual narrative readers have been shown to engage more actively in visual search when information is missing. This is argued to further indicate efforts to inferentially resolve narrative coherence (Hutson et al., 2018).

This relates to the concepts of both updating and event boundaries in that when prediction error occurs, and the current model is no longer useful, models are revised, or new base models are created which have more utility in predicting the impending narrative events (Zacks et al., 2007; Zacks & Swallow, 2007; Reynolds et al., 2007). Another way to approach this is to consider that as the inference required to connect two panels increases in complexity or difficulty, the cognitive load on the reader should also increase. This is because, as demonstrated in Experiment 4, initiating situation models is a mentally demanding process. This is particularly significant in static visual narratives, where readers must actively inferentially connect each subsequent panel with the previous one. Though inference is critical to narrative processing, particularly in visual narratives, much less attention has been paid to the relationship between this process and its operation under the working memory framework.

Research has shown that verbal and visuospatial working memory play distinct roles in narrative comprehension. Friedman and Miyake (2000) found that individuals with higher verbal working memory capacity were better at generating inferences essential for understanding text-based narratives. In contrast, visuospatial working memory was linked to comprehending the spatial aspects of a narrative. Importantly, there was a clear dissociation between these two functions: spatial working memory did not contribute to inferential or causal understanding, and verbal working memory did not support spatial comprehension. These authors then suggest that verbal working memory is responsible for causal information and, by extension, inference, while visuospatial working memory was suggested to be responsible for spatial details in situation models.

Currently, the findings of this thesis can be interpreted as suggesting that, though situation model processing is applicable across different types of media, the process by which these models are arrived at may be specific to the modality in which the narrative is presented. The working memory-related findings of Experiments 1 to 3 clearly indicate that while visuospatial working memory is related to processing visual narratives, verbal working memory appears unrelated to this construct. Similarly, the findings of Experiment 5 suggest that while visuospatial working memory plays a role in processing visual narrative sequences, its role may be limited to surface-level processing.

The methodology used by Friedman and Miyake (2000) may have unintentionally influenced the working memory components required to respond to their probes. Specifically, spatial probes were presented as maps, requiring participants to answer yes/no questions about spatial content, while causal probes were presented as text-based questions, potentially reinforcing the association between visuospatial working memory and spatial processing, as well as verbal working memory and causal inference. Furthermore, as with the majority of situation model or narrative processing research, Friedman & Miyake's work solely used prose texts and does not consider image-based or multimodal narratives. Because of this, the generalisability of their findings to other narrative formats is limited. This methodological design highlights the need for further research examining how verbal working memory functions specifically in the context of static visual narratives, which may rely on different cognitive processes than those observed in text-based narratives.

Research directly examining the role of working memory in processing visual narratives is limited. Magliano et al. (2016) is one of the few studies to explore the contributions of visuospatial and verbal working memory to inference generation in visual narratives. As discussed, their research previously identified roles of both working memory components. However, as suggested in Experiment 3, this possibly over-extends the role of verbal working memory in linguistic-based processes. This then raises the question of whether the verbal working memory is specifically necessary for inferential connections in visual narratives or whether its involvement is more general and dependent on the modality of the narrative.

Given that the results of this thesis currently suggest that visual narratives primarily rely on visuospatial working memory, it seems unlikely that verbal working memory would play a significant role in inference generation. If verbal working memory does play a role in processing visual narratives, as suggested by previous research, disruptions to situation model processing caused by incoherence should interfere with maintaining a concurrent verbal working memory load. This is because generating inferences to reconcile incoherent panels would compete for the same cognitive resources.

The dual-task paradigm used in Experiment 5 was applied to investigate the role of verbal working memory in visual narrative processing. This approach tests whether maintaining a pre-loaded verbal working memory stimulus is hindered when a secondary task competes for the same cognitive resources. By employing this paradigm, the study aimed to clarify the contribution of verbal working memory to visual narrative comprehension, with potential implications for cognitive theories of narrative processing and the structure of working memory architecture.

Based on the previous experiments and theoretical framework outlined, the following hypotheses regarding verbal working memory use in situation model processing were proposed. Firstly, it was hypothesised that processing coherent comic strips would not significantly interfere with maintaining a verbal working memory sequence compared to a control condition. This hypothesis was based on the previous experiments in the thesis (1 to 3), which demonstrate no significant activation of verbal working memory in processing visual narratives.

Second, it was hypothesised that processing incoherent comic strip narratives would not significantly interfere with participants' ability to maintain a verbal working memory task compared to the control condition. While previous research (Friedman & Miyake, 2000; Magliano et al., 2016) suggests that verbal working memory is involved in the inferential component of narrative processing, the findings of this thesis challenge and highlight the need to re-examine the role of verbal working memory in situation model processing and the potential for modality-specific contributions. Experiment 4 demonstrated that incoherent comics disrupt participants' ability to form inferential connections, leading to increased cognitive load due to expectation violations. If verbal working memory were essential for making these inferential connections, it would be expected that incoherent comics would impair the maintenance of a verbal working memory stimulus. However, the findings of the current thesis suggest otherwise, indicating that verbal working memory may not play a critical role in processing inferential connections in visual narratives, thus justifying the hypothesis.

Finally, it was hypothesised that visuospatial working memory, compared to verbal working memory, would show significant interference when processing coherent and incoherent comic strips. This aligns with modality specificity in working memory contributions to situation model processing. This hypothesis is supported by Experiment 5, which showed visuospatial activation during visual narrative processing, and Experiment 3, which showed no significant activation of verbal working memory.

4.2.2 Method

The general method of this experiment, including the procedure, was the same as the previous experiment. The titration, visual language fluency, and dual-task phases were all the same as in Experiment 5. The only difference was that the digit span task was used instead of the Corsi block test. Readers are referred to the previous procedure section for further details.

4.2.2.1. Participants

39 participants (29 female) completed this experiment study in exchange for course credit on an undergraduate Psychology course. All participants had normal to corrected vision. All participants gave informed consent to participate in the study and were naïve to the purposes and intentions of the study. The University Research Ethics Committee approved the experiment.

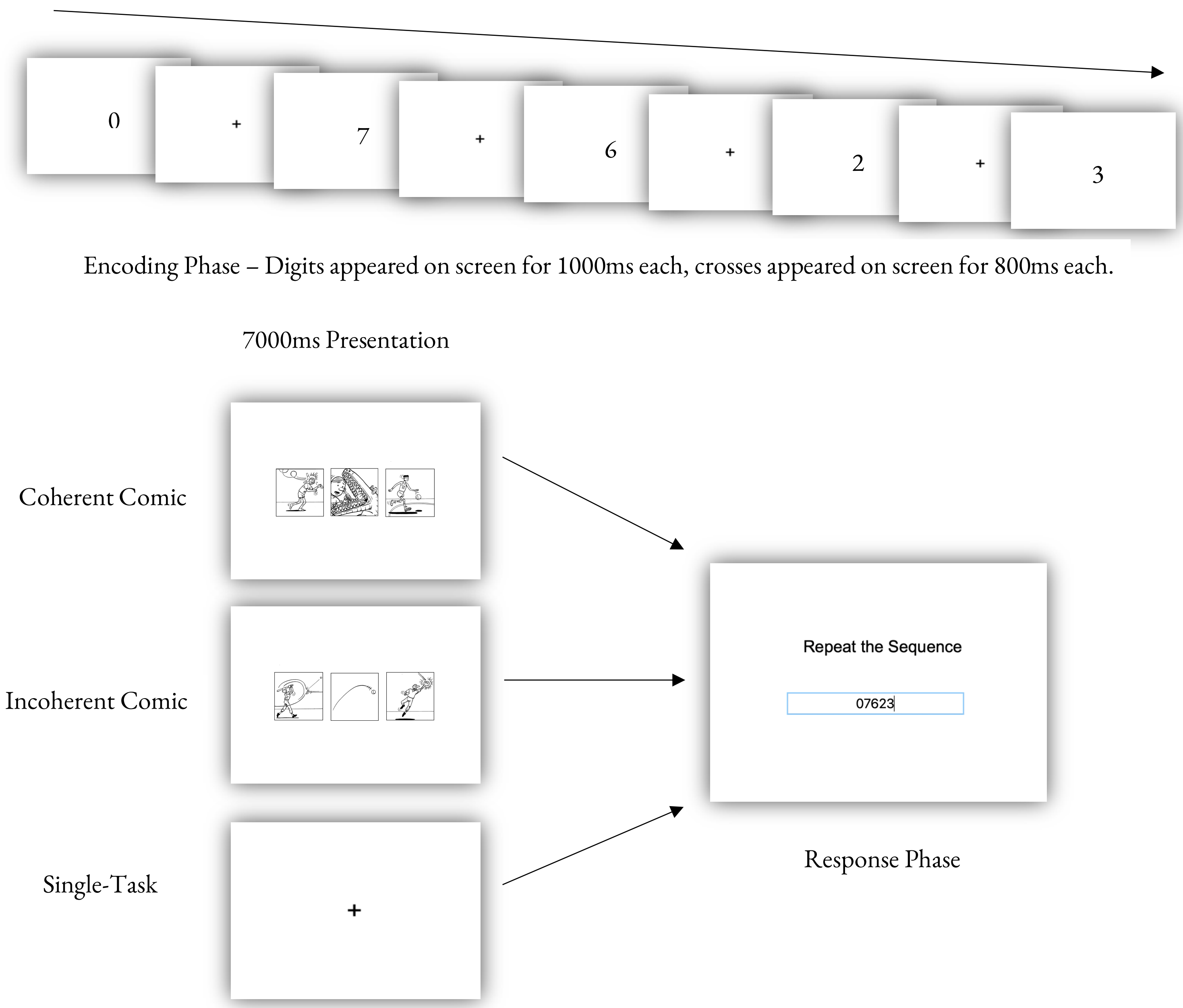
4.2.2.2 Materials

The comics selection and generation method used in this experiment were the same as in Experiment 5. The digit span task was functionally equivalent to that used in Experiment 3, with similar timings, but included modifications to address procedural disparities with the Corsi block task. Figure 15 depicts the procedure of experiment 6.

In the Corsi procedure, participants were presented with all blocks for a sequence at the start, making it clear when the sequence was about to begin and immediately revealing its length. This also signalled when the sequence length increased during the titration phase. By contrast, in the original digit span task, the fixation cross preceding the sequence was visually very similar to the digits, potentially reducing participants' attention. Additionally, participants were unaware when sequence lengths increased, possibly leading to relaxed attention.

Two modifications were introduced in the digit span task to address these issues and make the tasks more comparable. First, participants were informed of the number of digits in each upcoming sequence (e.g., "5 Digits" displayed on-screen). Second, a traffic light-style colour transition was added to the pre-fixation cross to better capture participants' focus. These changes aimed to reduce cognitive load related to attention, allowing participants to concentrate more on storing and maintaining verbal working memory. The modifications appeared to improve task performance, as indicated by an increase in participants' average verbal working memory span.

***Figure 15** Graphical Representation of Procedure in Experiment 6*



4.2.3 Results

The current data and the data from the previous experiment are presented here as a combined analysis in a mixed methods ANOVA. This approach was taken as it allows for both a comparison of repeated measures factors (coherence) and independent measures factors (dual-task type) while also providing answers to the hypotheses posed in the current experiment.

4.2.3.1 Data Preparation, Screening and Manipulation Checks

The data from this experiment were prepared using the same procedure as the previous one. Participants' accuracy across each of the conditions was averaged to create incoherent, coherent (both dual-task) and control (single-task) accuracy scores. All participants scored above the requisite 50% on the comics questions, meaning that no data based on this criterion needed to be removed.

To confirm that participants adequately engaged with the visual narratives of the dual-task procedures, their coherence ratings for coherent and incoherent comics were averaged and compared. As the coherence ratings were not normally distributed and used an ordinal Likert scale (1: Not Coherent At All – 7: Completely Coherent), a non-parametric test was conducted. A Wilcoxon signed-rank test revealed that the manipulation was successful, with participants giving significantly higher coherence ratings to coherent comics than to incoherent ones ($W = 666$, $p < .001$). The effect size was large ($r = 0.872$).

The analysis provided later in this section compares visuospatial and verbal dual-task performance, so it was important to confirm that the different groups of participants who completed each experiment equally recognised the coherence manipulation. To this end, a 2 * 2 mixed measures ANOVA (Verbal Dual-task/Visuospatial Dual-task * Coherent/Incoherent ratings) was used to make sure that there was no difference in comics engagement. ANOVA was chosen here due to a lack of a non-parametric alternative for mixed measures ANOVA. This analysis revealed that there was parity between the participant groups, with the only significant difference in coherence ratings being registered as a main effect of Coherence $F(1, 74) = 476.440$, $p < .001$, $\eta^2 = 0.8216$. Neither the main effect of Dual-task $F(1, 74) = 0.003$, $p = 0.955$ nor interaction effect of Dual-task by Coherence $F(1, 74) = 0.777$, $p = 0.381$ yielded a significant difference.

Finally, the participants' VLFI scores were tested to check for parity in their comics reading experience as this has the potential to affect the results. A Wilcoxon signed ranks test, used here on the basis that the VLFI scores were not normally distributed, revealed that there was no significant difference between the visual language fluency index scores of participants in the verbal and visuospatial working memory conditions ($W = 714$, $p = .408$). This then supports the conclusion that the results of this experiment were not affected by the participants' prior experience with comics.

4.3.2.2 Verbal Dual-Task Accuracy Assumptions.

The normality of the residuals for both visuospatial and verbal-dual-task accuracy was assessed using the Kolmogorov-Smirnov test. This result indicated a significant deviation from normality ($D = 0.09$, $p = 0.026$), suggesting that the residuals were somewhat negatively skewed. This contrasts with the residual normality observed in Experiment 5 and seems likely to be due to a ceiling effect present in the verbal working memory data. However, the skewness statistic of the combined model residuals was not found to be within a problematic range (-0.44). Levene's test was then conducted to assess the homogeneity of variance between the verbal and visuospatial dual-task accuracy. This revealed significant differences in accuracy variability between participants who completed the verbal and visuospatial working memory versions of the experiment $F(1, 226) = 16.821$, $p < .001$.

While the assumption tests indicate some issues with normality and homogeneity of variance, parametric tests were proceeded with for two reasons. First, as noted in previous results sections, F-tests and ANOVA are considered to be robust to violations of these parametric assumptions (Blanca et al., 2017; Tabachnik & Fidell, 2018). Second, there is no non-parametric equivalent for mixed ANOVA, therefore necessitating the use of parametric methods. Table 8 displays descriptive statistics for all conditions by dual task type.

4.3.2.3 Combined Analysis Main Effects

A 2 (Dual-task type: visuospatial, verbal) * 3 (Comic Coherence: incoherent, coherent, control) mixed-measures ANOVA was conducted to examine the differences in working memory activation across visuospatial and verbal working memory tasks during visual narrative processing. Initially, the sphericity assumption was tested using Mauchly's, and no violation was found, (W = 0.971, p = 0.344).

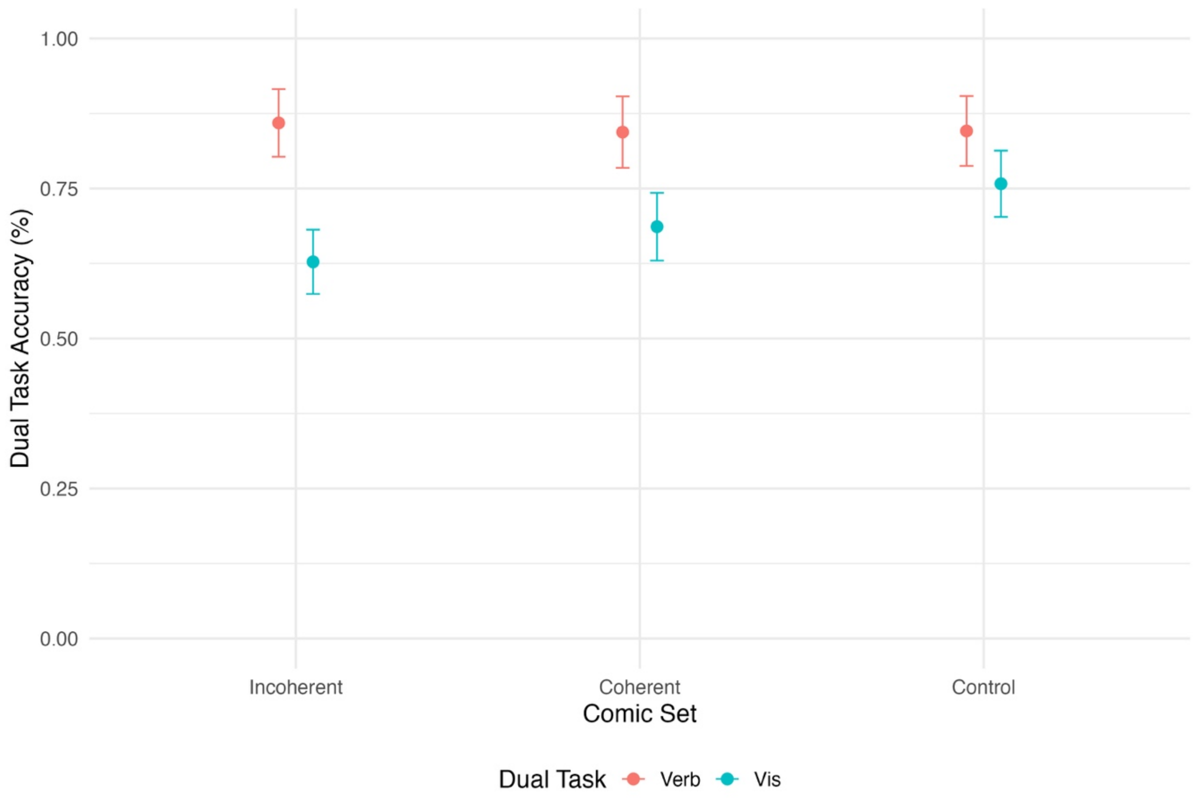
Table 8 *Corsi Block and Digit Span Accuracy Descriptives by Coherence*

Condition	Coherent	Incoherent	Single-Task
Visuospatial	0.686 (0.211)	0.628 (0.199)	0.758 (0.193)
Verbal	0.844 (0.136)	0.859 (0.129)	0.846 (0.154)

Note. Means and (Standard Deviations) for the visuospatial and verbal working memory by comics condition.

Both of the main effects of this ANOVA were found to be statistically significant. However, given the inclusion of control conditions in both the experimental design and analysis procedure, the interaction effect is the most meaningful analysis. This is because the control conditions provide a baseline for performance, which enhances the interpretation of the interaction effect but complicates the interpretation of non-interaction effects. The main effects are presented here for completion. Briefly, the dual-task type had a significant main effect on accuracy $F(1, 74) = 26.633$, $p < .001$, with a large effect size: $\eta^2 = 0.1749$. There was also a main effect of condition (incoherent dual-task, coherent dual-task or single-task control) $F(2, 148) = 3.500$, $p = 0.033$, with a small effect size, $\eta^2 = 0.0191$. However, most importantly, the main analysis also revealed that there was a statistically significant interaction between dual-task type by condition on accuracy $F(2, 148) = 5.177$, $p = 0.007$. Though this finding was associated with a similarly small effect size to the coherence effect: $\eta^2 = 0.0280$, it indicates that the influence of condition on accuracy differed depending on the type of dual task being performed. In other words, the interaction suggests that the effect of task condition on accuracy is not uniform across dual-task types, highlighting the varying demands placed on visuospatial and verbal working memory in these conditions. Figure 16 displays the estimated marginal means for each of the conditions. Several Holm-corrected post hoc t-tests were conducted to explore the significant interaction effect revealed by the ANOVA.

Figure 16 *Comparitive Effects of Coherence on Visuospatial and Verbal Working Memory Performance*



4.3.2.4 Verbal Dual-Task Accuracy Comparisons

Initially, repeated measures post-hoc tests examining the effect of condition on verbal working memory dual-task accuracy were conducted. As is indicated in figure 16, there was a distinct lack of difference between accuracy recorded between the coherent comic, incoherent comic and single-task control conditions. In the verbal working memory condition, accuracy following the coherent comics was found to be equivalent to accuracy following an incoherent comic $t(74) = -0.48$, $p = >.99$, 95% CI [-0.09, 0.06], Cohen's $d = -0.05$, ($MD = -0.015$). Even more strikingly, accuracy following the coherent comic condition was not found to be statistically significant by comparison to the control condition with an approximate difference of 0.2% $t(74) = -0.06$, $p = >.99$, 95% CI [-0.08, 0.07], Cohen's $d = -0.01$, ($MD = -0.002$). Finally, incoherent comics were also found to cause no significant level of interference by comparison to the control condition in which participants were simply required to hold the verbal working memory sequence in mind $t(74) = 0.38$, $p = >.99$, 95% CI [-0.07, 0.10], Cohen's $d = 0.04$, ($MD = 0.013$). These findings contrast substantially with the findings of the previous experiment and provide further support for the position that verbal working memory does not make any significant contribution to visual narrative processing.

To explore the significant interaction effect between coherence and dual-task type, a series of between-groups post hoc tests were conducted. These tests revealed several interesting findings. First, a comparison of accuracy between the control conditions in the verbal and visuospatial working memory tasks showed no statistically significant difference ($t(74) = 2.19$, $p = 0.192$, 95% CI [-0.03, 0.21], Cohen's $d = 0.50$, $MD = 0.088$). This result suggests a reasonable equivalence between the dual-task types in the absence of secondary task demands and highlights the efficacy of titrating working memory tasks to individual differences.

In contrast, a statistically significant difference was observed between visuospatial and verbal working memory accuracy following the coherent comic condition ($t(74) = 3.83$, $p = 0.002$, 95% *CI* [0.03, 0.28], Cohen's $d = 0.88$, $MD = 0.158$). This result strongly supports the idea that visuospatial working memory primarily supports visual narrative processing, whereas verbal working memory is less involved. Finally, for the incoherent comic condition, accuracy between verbal and visuospatial working memory tasks was found to differ significantly ($t(74) = 5.94$, $p < 0.001$, 95% *CI* [0.11, 0.35], with a very large effect size, Cohen's $d = 1.36$). This finding reinforces the conclusion that verbal working memory does not play a significant role in visual narrative processing. Across both coherence conditions, participants maintained verbal working memory sequences without significant disruption, whereas visuospatial working memory was more heavily engaged and affected.

4.3.2.5 VLFI Scores and Digit Span Accuracy Correlations by Coherence

As in experiment 5, Pearson correlations were conducted to assess the relationship between verbal working memory accuracy and VLFI scores across the three conditions of the experiment: coherent comics, incoherent comics, and single-task control. Across all conditions, the mean VLFI score of participants who completed this experiment was 8.19 ($SD = 6.37$). In the coherent condition, the correlation between accuracy and VLFI was negative and not statistically significant, $r(34) = -.15$, $p = .380$, 95% CI [−.46, .19]. Similarly, in the incoherent condition, the correlation was nearly identical, $r(34) = -.15$, $p = .377$, 95% CI [−.46, .19]. In the control condition, the correlation was also negative and non-significant, though slightly larger, $r(34) = -.20$, $p = .247$, 95% CI [−.49, .14]. As such, no reliable linear association was found between VLFI and verbal working memory dual-task accuracy.

4.2.4 Discussion

The objective of the current experiment was to determine whether verbal working memory resources support visual narrative processing. Three hypotheses were tested. It was first proposed that processing coherent comic strips would not significantly interfere with maintaining a verbal working memory sequence. Second, that processing incoherent comics would not have any significant effect on verbal working memory. Finally, visuospatial working memory would show significant interference when compared to verbal working memory when processing both coherent and incoherent comics. These hypotheses, contrary to the currently available evidence (Friedman & Miyake, 2000; Magliano et al., 2016), were formulated on the basis of the previous experiments in this thesis.

The results provided support for all three hypotheses. Across all conditions, coherent, incoherent and single-task control, verbal working memory performance remained equivalent meaning that participants were able to maintain verbal working memory stimulus with no interference caused by the requirement to simultaneously process a coherent or incoherent comic. As in Experiment 3, this finding again contrasts with previous research, such as Magliano et al. (2016), in which verbal working memory is suggested to contribute to visual narrative processing. Instead, these results indicate that verbal working memory is not significantly involved in visual narrative processing at the level of the situation model.

In comparison, visuospatial working memory showed significant dual-task costs when participants were required to simultaneously process a comic. This then provides further evidence to support the conclusion that visuospatial working memory underpins visual narrative processing. Similarly, the findings also support the idea that modality-specific processes play a critical role in narrative processing and should be considered further in theories such as SPECT (Loschky et al., 2020).

The results of this experiment provide additional evidence that visual narrative processing relies on modality-specific resources. The absence of any impact on maintaining a verbal working memory stimulus indicates that these resources are neither engaged nor required during visual narrative processing. This finding not only reinforces the conclusion that verbal working memory is unnecessary for visual narrative situation model processing but also challenges the idea that verbal working memory is essential for inference generation. Although earlier research (Friedman & Miyake, 2000) has suggested a role for verbal working memory in generating inferences, the current findings suggest otherwise. Given that inference generation is a cognitively demanding process, the manipulation in this experiment was expected to impose greater strain on working memory resources when processing incoherent comics. However, this expectation was not supported, further questioning the involvement of verbal working memory in this context.

Instead, these findings suggest that inference generation may depend on other narrative processing mechanisms. Specifically, it appears likely that inference generation during narrative processing relies more heavily on deeper semantic levels of processing. This supports the conclusion that the role of working memory in visual narrative situation model processing should be reconsidered, emphasising its function in handling perceptual front-end processes rather than the broader role currently proposed in SPECT (Loschky et al., 2020). Furthermore, since the stimuli used in this experiment included relatively simple narratives, future research that conducts similar experiments using semantically impoverished narrative stimuli to further explore this hypothesis would likely provide significant value.

The results of this study further contrast with the conclusions of Magliano et al. (2016), who proposed that verbal working memory contributes to inference generation during visual narrative processing based on reading times and the impact of a dual-task memory load. While reading times can provide valuable insights, they may not fully capture the cognitive load or resource allocation involved in visual narrative processing. Using the dual-task paradigm, the current experiment offers a more direct assessment of working memory activation. The findings show that participants' ability to maintain verbal working memory sequences remained consistent across conditions, indicating that verbal working memory resources were not significantly taxed despite the additional processing demands imposed by visual narratives.

One potential limitation of this experiment is the possibility that the participants disengaged their attention when processing incoherent comics, perhaps realising quickly that they were incoherent. This could have led to lower cognitive effort being employed in this condition and potentially skewing the results. However, the findings of the manipulation check provide a level of evidence against this. These analyses demonstrated that in rating the coherence of the comic they had just seen, the coherent comics were consistently rated highly, and the incoherent comics were consistently rated low in coherence. Furthermore, significant differences were found between these ratings in both versions of the experiment. This indicates that the participants had registered the content of the panels and also registered whether that content was congruous with one another. Arguably, the process of arriving at the conclusion of coherence or incoherence can be suggested to reflect a process of attempting to generate inferential connections between the panels, with a decision about this only being possible by assessing this. Consequently, the finding that processing either coherent or incoherent comics was not significantly different in the visuospatial working memory condition and not significantly different from control in the verbal working memory condition suggests that these components of working memory are not directly linked to connecting panel information by inference.

While the comics are argued to have naturally elicited the process of inference generation, given that much of comics processing requires this, there is still a question of the point of disengagement in relation to the incoherent comics. As discussed several times, the rapid nature of visual processing likely means that the participants were able to reach the conclusion that the contents of the comics were incoherent very quickly and this may have led them to disengage from any further inference process. To this end, future research which explores the use of working memory in visual narrative processing would benefit from the use of comics with graded coherence as opposed to the binary coherent/incoherent manipulation used in this experiment.

It is also important to note that while the current findings suggest that verbal working memory does not contribute significantly to visual narrative processing, the stimuli used could be argued to not have explicitly manipulated inference in the way that other studies have (e.g., Magliano et al., 2016; Cohn & Wittenberg, 2015). Inference in narrative processing often involves the active bridging of narrative gaps, such as when a critical panel is missing or replaced. In contrast, the current manipulation involved coherent versus incoherent sequences that may disrupt narrative mapping but do not necessarily elicit the same kind of inference demands. As such, while the results indicate a limited role for verbal working memory in mapping-based processing, they could be argued not to directly speak to the role of working memory in inference-driven situation model construction. This then remains an important direction for future research.

Despite these limitations, the experiments presented in this chapter represent one of the first attempts to directly investigate and measure working memory activation in visual narrative processing by combining the dual-task paradigm with a situation model-based coherence paradigm. While the results shed light on the role of working memory in visual narrative processing, the most robust and justified conclusion that can be drawn is consistent with the findings of Experiments 1 to 3: visuospatial working memory appears to be primarily responsible for online visual narrative processing. In contrast, verbal working memory does not appear to make any significant contribution to visual narrative processing. This conclusion is supported by the significant reduction in visuospatial dual-task accuracy observed when participants were required to simultaneously process a comic strip and the lack of interference caused to verbal working memory dual-task accuracy.

In contrast, and opposing previous findings that implicate verbal working memory in causal connections and inference generation, the results replicate and confirm that verbal working memory is not involved in visual narrative processing. Furthermore, the findings from this experiment, along with the previous one, suggest that the role of working memory in visual narrative processing may be limited to perceptual, 'front-end' processes and does not extend to deeper visual narrative situation model processing. While this hypothesis requires further research to be substantiated, the results highlight the importance of considering modality-specific influences in visual narrative processing. These findings underscore the need for theoretical frameworks addressing visual narrative processing to give greater attention to the role of modality specificity.

Overall, these findings are perhaps unexpected on the basis that working memory relates to both storage and manipulation of information within it, yet does not appear to be engaged by the process of inferentially connecting information between panels or in visual narrative processing past front-end processes. In light of the suggestion that modality-specific processes should be considered when expanding visual narrative processing theories, the final experiment will focus on the role of mental imagery in processing visual narratives. Mental imagery is highly specific to both visual and spatial information and reflects the ability to both generate and manipulate visual mental representations. The final experiment in this thesis will explore this cognitive ability to better understand its influence on visual narrative processing.

4.3 Experiment 7 – Mental Imagery in Visual Narrative Processing

4.3.1 Introduction

Experiment 6 explored the role of verbal working memory in visual narrative processing. Though this cognitive component has previously been associated with both comics (Magliano et al., 2016) and tracking causality in narratives (Friedman & Miyake, 2000), no evidence was found to indicate that verbal working memory working has any role in visual narrative processing of textless comics. This finding is congruous with the suggestion that while visual narrative processing behaviours are adequately captured by situation model processing theory, it seems likely that the generation of these models requires modality-specific processes. Given the apparent demonstration of these modality-specific influences, the final experiment considers the role of visual mental imagery (MI) ability in visual narrative processing.

Visual mental imagery is the ability to create and manipulate images in the mind in the absence of external perceptual stimuli. This ability is considered to play an important role in several cognitive functions, such as memory (Marre et al., 2021) and reasoning (Knauff & May, 2006). The distinction of *visual* mental imagery is important in its relation to visual narrative processing, as mental imagery is not limited to the visual modality. Mental imagery abilities have been demonstrated in auditory (Willander & Baraldi, 2010), gustatory (May et al., 2004) and also tactile (Kosslyn et al., 1995) modalities. A common demonstration of the way that mental imagery affects perception and experience is to imagine sipping a glass of freshly squeezed lemon juice and notice the astringency effect (the coagulation of saliva, which creates a feeling of puckering and tightness) that simply imagining this process causes. Visual mental imagery has also been demonstrated to exert an effect on experience and perception in that this cognitive function has been significantly associated with anxiety (Hackmann et al., 1998), OCD (Speckens et al., 2007) and depression (Blackwell et al., 2015) where overactive visual mental imagery plays a role in re-experiencing in PTSD, or intrusive imaginings in OCD. Thus, understanding this cognitive phenomenon is arguably critical to understanding much of cognition.

In relation to visual narrative processing, it is proposed that visual mental imagery specifically relates to both visual and spatial components of visual narrative processing. The visual nature, or the idea that mental images are directly depictive of their referents, has been debated. One school of thought was that mental imagery is propositional as opposed to depictive (Pylyshyn, 1973; 2002). This theory argued that instead of mentally pictured scenes, processed information would be encoded as a set of linguistic components with relations between objects. However, neuroscientific methods have since demonstrated that visualisation, or forming a mental image, recruits the primary visual cortex, which proponents of the depictive theory of imagery suggest demonstrates clear evidence of the visual nature of mental imagery (Pearson & Kosslyn, 2015; Albers et al., 2013; Slotnick et al., 2005; Dijkstra et al., 2019). What might be considered typical imagery then, is argued to directly relate to visual, depictive representations.

Evidence supporting the idea that mental imagery involves spatial and visual components comes from studies highlighting the spatial aspect of situation model processing. Specifically, findings indicate a distinct spatial element in the use of situation models and this can arguably be interpreted as reflecting mental imagery generation. For instance, studies on foregrounding information show that representational situation models incorporate spatial relationships. One example of this is foregrounding of certain situation model components, meaning that currently activated elements of situation models are more accessible than deactivated ones (Haenggi et al., 1995). Moreover, experiments using the map learning paradigm further support the spatial nature of mental imagery. In this paradigm participants first memorise a map which includes a number of items in each space. Participants then read a narrative in which a protagonist moves through the learned map environment. Both probe times and anaphoric sentence reading times have demonstrated a consistent effect of spatial distance between the protagonist and the item in question (Rinck et al., 1997; Bower & Morrow, 1990; Bower & Rinck, 2001; Kosslyn, 1973; Kosslyn et al., 1978). In other words, the further away from the current protagonist a probed item is, the longer it takes to retrieve information about it which indicates that visual mental imagery representations contain information that relates directly to space.

Given the parallels between them, visual mental imagery has an interesting relationship with visuospatial working memory (VSWM) and the connection between the two is a topic of ongoing debate. With the significant overlap in definitions between the two: where visuospatial working memory is suggested to be a cognitive capacity which processes, manipulates, retrieves and maintains visuospatially based information, and visual mental imagery is proposed as the ability to generate and manipulate mental imagery in the absence of external stimuli, intuitively, it feels clear that these two concepts are directly linked. Indeed, the link between these two cognitive faculties is so apparent that some authors have suggested they are actually the same faculty (Tong, 2013).

However, while research has not definitively separated VSWM and MI, the balance of evidence suggests at least some level of dissociation. The most convincing evidence currently available for this comes from studies investigating Aphantasia (Zeman et al., 2015), a condition which describes the inability to generate mental imagery. If these two cognitive components were functionally indistinguishable, aphantasia, which has been shown variously to affect different kinds of processing (Wicken et al., 2021; Keogh & Pearson, 2018; Kay et al., 2024) should significantly affect visuospatial working memory tasks, however, several studies report no deficit in aphantasics visuospatial working memory performance (Dawes et al., 2022; Bates & Farran, 2021; Bates et al., 2024).

As suggested, there is significant variation between individuals in their capacity for mental imagery. Aphantasics report very low to zero mental imagery ability and at the opposing end of the spectrum, hyperphantasics report overly vivid mental representations (Zeman et al., 2020). The link between mental imagery and situation models is similar to that outlined in Experiment 4 between working memory and situation model processing in that there is an implicit assumption of connection between the two. The situation model level of narrative processing is for example, described in theories explicitly as the 'representational level' (Zwaan, 2016; van Dijk & Kintsch, 1983), yet, the concept of mental imagery has only recently begun to be associated with situation model processing literature. These two concepts are indirectly tied together through mental imagery research and the effect of individual differences on processing. Though the focus of this experiment was not situation models, one particularly stark illustration of the effect that greater mental imagery plays in processing are the findings of Wicken et al., (2021). Their study assessed fear responses in aphantasics and typical visual imagers with galvanic skin response (GSR). When asked to read a fear inducing story, sentence by sentence, in comparison to typical imagers, the aphantasics in the study registered strikingly little GSR. Their participants were subsequently asked to view a series of fear inducing images and in this, aphantasic and typical imagers produced equivalent GSR responses. This demonstrates two important findings. First, it indicates that mental imagery plays a critical role in some instances of conjuring emotion. Second, it demonstrates very plainly that mental imagery is linked to narrative processing in that typical imagers formed vivid enough representational models of the text based content to produce a measurable fear response.

Mental imagery then has an interesting relationship with reading. Given the variability that exists between individuals in mental imagery, it seems likely that this significant effects on reading-related abilities. This is further evidenced by several recent studies which outline effects related to both comprehension and reading engagement. Suggate & Lenhard (2022) report findings that mental imagery ability significantly predicts both reading comprehension and reading speed. Related to this, research has also begun to demonstrate a profile of lower engagement in reading related to lower visual mental imagery ability (Speed et al., 2024; Williams & Suggate, 2024). This is intuitively logical, in that, the inability to build the worlds conveyed by great fiction in our minds eye perhaps seems likely to leave one with a cold feeling about reading. However, beyond simple engagement, Suggate & Lenhard also demonstrate an association of comprehension with mental imagery. This then suggests that the functional role of MI in reading processes goes beyond the generation of a superficial representation.

One possibility is that a function of mental imagery is to support the ability to generate inferences. Several branches of research support this suggestion. As outlined above, the primary visual pathway has been associated with both top-down visual working memory input and bottom-up visual mental imagery production (Albers et al., 2013). Further research in this area has similarly demonstrated that the primary visual pathway (V1 - V3) is also activated during trajectory-specific processing and critically, is activated even when the trajectory of a given object is occluded (Zbären et al., 2023). Objectively, this demonstrates that when a visual stimulus is presented as moving, but is then obscured, pathways associated with mental imagery are activated. Further highlighting the parallel between visual mental imagery and visuospatial working memory, research has shown visuospatial working memory is significantly impacted when intuitive physics expectations are not met. Balaban et al. (2023) used contralateral delay activity, which refers to an event-related potential which is considered to provide a measure of the amount of information being stored in visuospatial working memory (Ikkai et al., 2010). When participants in Balaban et al. (2023) were shown animated videos which depicted stimuli that violated their expectations, their participants' contralateral delay activity was shown to drop sharply. This is taken to indicate that the object was removed from working memory. This, in turn, indicates support for the concept that working memory provides a site of mental simulation. This could then be interpreted as evidence for the naturalistic use of mental imagery in generating inference.

Further evidence for this suggestion comes from findings which indicate that mental imagery is capable of priming perception. In binocular rivalry studies, two different images are presented separately to each of a participant's eyes leading to perceptual competition. This competition results in only one of the two competing stimuli reaching awareness. Experiments using this technique where participants are asked to imagine a specific stimulus before viewing stimuli pairs have shown that this has the effect of biasing perception towards the mentally imagined stimulus increasing the likelihood that the imagined stimuli reaches awareness (Pearson et al., 2008). Further to this, subsequent studies have demonstrated that the strength of bias towards the imagined stimulus is predicted by self-reported vividness of mental imagery (Pearson et al., 2011) using the VVIQ (Marks, 1973). Together the findings returned by binocular rivalry, as well as intuitive physics studies strongly support the suggestion that mental imagery biases basic visual processes. However, necessarily these studies have focused on the use of reductive, non naturalistic stimuli.

At present, only one study has investigated mental imagery in relation to visual narrative processing. Boerma et al. (2016) reported that children with higher levels of mental imagery ability demonstrated greater comprehension of picture-only texts compared to text-only and picture-plus text conditions. This indicates a significant effect of mental imagery, specifically in visual narrative processing and arguably supports the suggestion that the process of inferentially connecting images in visual narrative sequences makes use of visual mental imagery ability.

Supporting this, the reading time data from the fourth experiment presented in this thesis revealed a downward trend in reading times for coherent comics. As each subsequent panel met the expectations set by the previous one, readers required less time to view them implying that they were mentally simulating the narrative. This would be in line with the suggestion of Zacks (2007) and the prediction error hypothesis proposed to account for the process of updating event models. Given the influence that mental imagery has on basic perceptual processes, as well as the findings by Boerma et al., (2016) which specifically demonstrate a relationship between mental imagery ability and visual narrative processing, it seems likely that mental imagery ability plays a role in visual narrative processing. The purpose of the final experiment then was to explore the relationship between mental imagery and visual narrative processing through reading times.

This experiment used the same coherence and incoherence framework as Experiments 5 and 6. Participants completed the VVIQ (Marks, 1973) as a measure of mental imagery ability and subsequently viewed coherent and incoherent sets of comic panels in sequence. Their panel reading times were recorded. Overall, it was hypothesised that mental imagery ability would show facilitative effects on visual narrative processing. In line with this, it was hypothesised that high mental imagery participants would demonstrate faster comics reading times overall. It was also hypothesised that the same would be true for incoherent comics, as if mental imagery ability is facilitative to processing, then while incoherent comics should slow reading times down, mental imagery ability seems likely to compensate for incoherence through mental simlation. Further to this, if mental imagery ability facilitates predictive visual narrative processing, then it seems likely that there would be an interaction effect between panel position and mental imagery ability with higher mental imagery participants showing faster reading times for later panels.

4.3.2 Method

For transparency, it is important to note here that the reading time data used in this experiment is the same as that presented in Experiment 4. VVIQ data was collected as part of this procedure and is presented in this experiment for analysis.

4.3.2.1. Participants

84 participants (71 Female) completed the study. Their age ranged from 18 to 35 years (M = 20, SD = 2.21). Participants were recruited through a research participation scheme in which undergraduates receive course credits for participating in research. All participants gave informed consent to participate in the study and were naïve to the purposes and intentions of the study. The experiment was approved by the University Research Ethics Committee.

4.3.2.2 Design

This experiment employed a repeated measures design and reused several variables from Experiment 4 for analysis. Specifically, reading times from comics with manipulated coherence collected in the procedure of Experiment 4 were used as dependent variables.

4.3.2.3 Materials

Mental imagery ability was measured using the Vividness of Visual Imagery Questionnaire (VVIQ; Marks, 1973). This self-report measure has been used consistently in research on mental imagery. Participants are asked to picture a general scenario (I.E. "Think of a horizon") and then picture details about the scenario ("Picture the sun is rising above the horizon into a hazy sky."). Participants then rate how vividly they have been able to picture the detailed scenario. The VVIQ consists of four general prompts, each with four detail prompts meaning that participants answer 16 questions in total. Each prompt is rated on a scale of 1 (No image at all) to 5 (Perfectly clear and vivid). This produces scores from 16 to 80. This measure is both widely used in mental imagery research and as has been validated through reliability and construct validity assessments (Campos & Perez-Fabello, 2009). The internal consistency and reliability of this scale has been demonstrated to be high with Cronbach's Alpha values of 0.91. The VVIQ is presented in Appendix N.

4.3.2.4 Procedure

Participants in this study were initially presented with the information sheet followed by the consent checks. They then provided age, gender details and an email address. Broadly, the procedure of this experiment included completing the VVIQ and completing the comics reading task. The presentation of the VVIQ component of the experiment was counterbalanced to avoid order effects, meaning that half of the participants completed the survey prior to the reading task and half completed it after.

In the reading time task, the participants first recieved on-screen instructions to read each comic panel and press the space bar when finished. Each comic set presentation began with a 1500ms interstimulus cross, which as followed by a comic panel. After viewing the panel, participants answered a question: for the first panel, they confirmed whether they understood it;

for the subsequent panels, they indicated if the panel related to the previous one. Both coherent and incoherent panel sets were presented in a counterbalanced order.

For the VVIQ, the broad prompts were presented first. These were then followed by the detail prompts that the participants were asked to picture and rate. For each scenario envisaged, the initial broad prompt (For example "Think of a relative or friend who you see frequently") remained at the top of the screen as a reminder. The subsequent detail prompts ("Picture the exact contours of their face, head, shoulders and body") were presented underneath this text along with the rating scale.

4.3.3 Results

4.3.3.1 Aphantasia Screening

Initially, the data was screened for participants who had scored low enough on the mental imagery ability scale to be suggested to be aphantasic. This was considered important for two reasons. First, Aphantasia has only recently been considered a condition (Zeman et al., 2015) and as such, its effects are not well understood. Second, the research question posed by this experiment relates to how mental imagery ability affects visual narrative processing and on this basis it was considered appropriate to remove participants who are considered to lack this ability.

Zeman et al., (2015) reports that the highest recorded VVIQ score from people who then self reported as aphantasic was 32. It should be noted that other prominent mental imagery authors consider scores of 17 - 32 on the VVIQ simply as representing low mental imagery (Dawes et al., 2020). This reflects the issue of metacognitive awareness with self-report of mental imagery scores, in that participants who report low mental imagery ability may simply have low metacognitive awareness of their mental imagery. As the experiment did not include a check of metacognitive awareness to rule this possibility out, the more expansive cut off suggested by Zeman et al., (2015) was used to mitigate this issue. This procedure excluded 3 participants from the analysis with scores of 19, 16 and 21. Subsequently the lowest VVIQ score in the data was 33. This left 71 participants data for analysis.

4.3.3.2 VLFI Scores and Mental Imagery Ability Correlation

A Pearson correlation was conducted to explore the relationship between mental imagery ability and VLFI scores. Intuitively, mental imagery appears likely to be a component of visual narrative processing. This is related to the requirement for connection and mapping between the static images of a panel that would appear to require the participant to generate content that links the depicted information. Given that the primary medium of communication in comics is visual, it follows logically that the mechanism to generate comprehension of visual sequences would be related to visual mental imagery. However, the analysis indicated that there was no significant linear relationship between VLFI scores and VVIQ data in the data collected $r(69) = 0.067$, $p = 0.57$, 95% CI [-0.169, 0.296].

4.3.3.3 Mental Imagery Ability Split

Based on their responses and the overall score recorded on the VVIQ, the participants were split into high and low imagery groups. This preparation was taken on the basis of the ordinal and self reported nature of the VVIQ. Forming imagery groups based on VVIQ score is supported by several previous studies which have employed this method (Zeman et al. 2020; Takahashi & Yasunaga 2012; McKelvie et al., 1994). Participants were then designated as either "High" or "Low" imagery ability based on a median split of their scores. The median of the aphantasic screened data was 61. This reflects a typical average as reported by other studies using the VVIQ (Zeman et al., 2020; Tabi et al., 2022; Keogh et al., 2021). 34 participants were determined to have high imagery ability and 37 were designated as having low imagery ability.

4.3.3.4 Assumption Checks

Prior to the main analyses, the assumptions of normality and homogeneity of variance were assessed. The Kolmogorov-Smirnov test indicated that the residuals of the ANOVA model were not normally distributed $D = 0.085$, $p = 0.004$. Levene's test also indicated that there was a significant violation of the homogeneity of variance assumption, $F(1, 424) = 7.26$, $p = 0.007$. However, despite these violations, the analysis proceeded with parametric testing methods on the basis that, as discussed in previous results sections, F-tests are considered to be robust against violations of the parametric assumptions (Tabachnik & Fidell, 2016; Blanca et al., 2017).

4.3.3.5 Main Effects

A 2 * 3 * 2 mixed design ANOVA was conducted to examine the effects of coherence (coherent vs incoherent), panel position (positions one, two and three), and imagery ability (high vs low) on panel reading times. This analysis revealed several significant main effects, but only one significant interaction effect. As in Experiment 4, significant main effects of both coherence condition $F(1, 69) = 19.131, p < .001$, $\eta^2 = 0.0484$, panel position $F(2, 115) = 87.197, p < .001$ (Huynh-Feldt corrected), $\eta^2 = 0.1842$ and the interaction between the two $F(2, 117) = 10.356, p < .001$ (Huynh-Feldt corrected), $\eta^2 = 0.0253$ were found. Although these data, and thus, the findings are the same as those presented in Experiment 4, meaning that these results were expected, they are presented here again with the inclusion of mental imagery ability as an additional factor in the ANOVA. The inclusion of this factor, and the fact that the primary findings remained the same, indicate that the effects of coherence and panel position on reading times remain robust and valid even when account for differences in mental imagery ability. This suggests that mental imagery ability does not influence or interact with processing either coherent or incoherent comics across different panel positions.

The main effect of mental imagery ability on reading times was found to be statistically significant $F(1, 69) = 7.569, p = 0.008$, and this was associated with a small to medium effect size $\eta^2 = 0.0486$. This indicates that the participants' mental imagery ability significantly affected the amount of time that the participants took to read the panel sets. However, this was the only statistically significant effect of mental imagery ability. No interaction effect was found between imagery ability and panel position $F(2, 115) = 2.801, p = 0.074$ (Huynh-Feldt corrected), $\eta^2 =$

0.0072, nor was there any effect of imagery ability on coherence $F(1, 69) = 0.007$, $p = 0.936$. Finally, there was also no statistically significant effect of the three way interaction of imagery ability, coherence and panel position on reading times $F(2, 115) = 0.670$, $p = 0.488$ (Greenhouse-Geisser corrected), $\eta^2 = 0.0017$.

4.3.3.6 Post Hocs - High vs Low Mental Imagery Ability Comparison

Although significant main effects of coherence and panel position were found, post hoc analyses are not reported for these on the basis that they present no new information beyond the findings covered in Experiment 4. However, the main effect of imagery represents a novel finding and as such, a between-groups, Holm-corrected post hoc t-test revealed, somewhat surprisingly, that lower mental imagery participants were actually significantly faster to read the comics than higher mental imagery participants. On average, participants with lower mental imagery were 631ms faster ($MD = 0.631$) in reading the comics overall $t(69) = 2.75$, $p = 0.008$, 95% CI [0.17, 1.09], and this was associated with a medium effect size Cohen's $d = 0.65$.

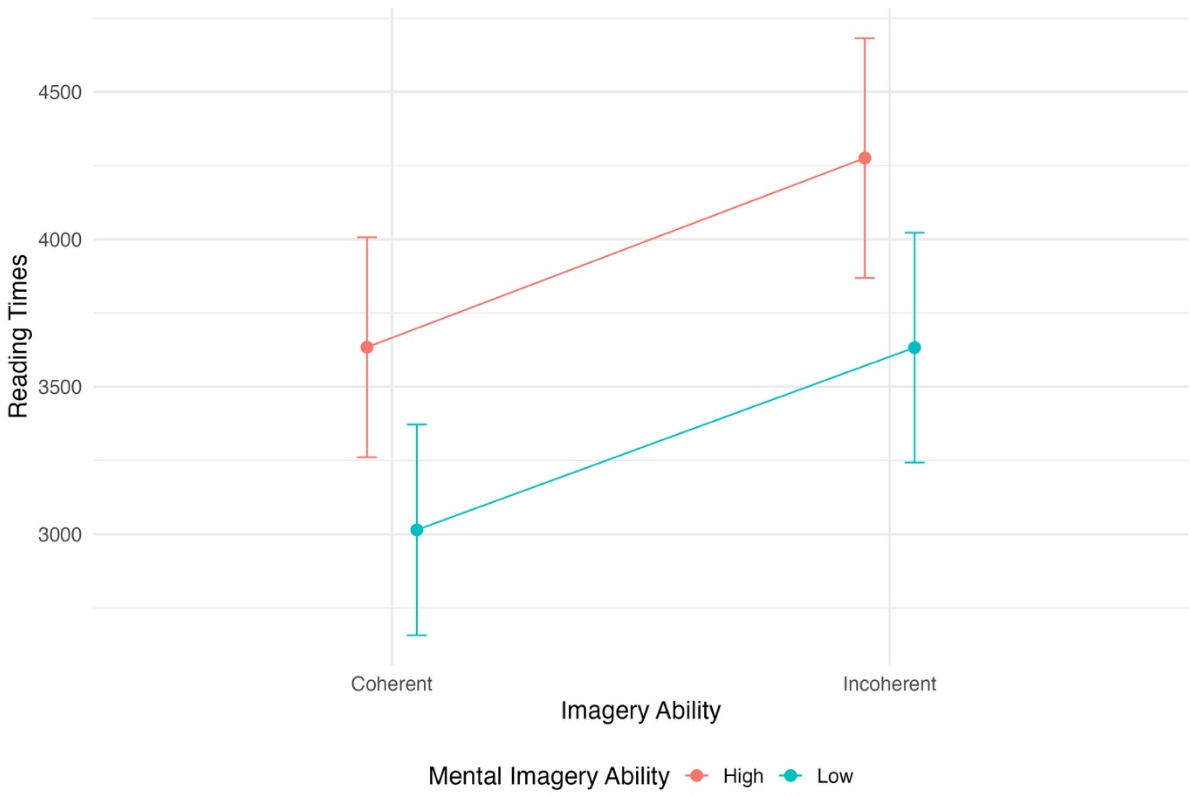
Table 9 *Reading Times by Coherence and Mental Imagery Ability*

Imagery	Coherent Comics	Incoherent Comics
High	3.63 (1.93)	4.28 (1.62)
Low	3.01 (1.24)	3.63 (1.46)

Note. Means and (Standard Deviations) for coherent and incoherent panel set reading times by imagery ability.

Figure 17 displays reading time estimated marginal means by coherence and imagery ability and this serves to demonstrate the consistency of the main effect revealed in the analysis. High imagery participants consistently exhibit longer reading times, regardless of coherence, in comparison to low imagery participants. Overall, the results indicate that mental imagery ability does impact visual narrative processing through reading times however, the results of this experiment were entirely contrary to the hypotheses proposed. High imagery ability participants read comics significantly more slowly than low imagery ability participants.

Figure 17 *Effects of Coherence on Reading Times by Mental Imagery Ability*



4.3.4 Discussion

The aim of the current experiment was to explore the potential role of mental imagery in visual narrative processing. Previous research indicates that mental imagery is linked to narrative processing with mental imagery generated from text based processing having been shown to elicit emotional responses from text based narratives (Wicken et al., 2021). Similarly, mental imagery is suggested to support children's comprehension of picture stories (Boerma et al., 2016). Additionally, findings suggest that faster reading speeds are predicted by greater mental imagery (Suggate & Lenhard, 2022). This experiment then aimed to probe the relationship between mental imagery and visual narrative processing. The most likely role of mental imagery in visual narrative processing relates to the concept of closure, or connecting the information presented in the panels of a comic. It was hypothesised that higher mental imagery ability would be related to faster comic panel reading speeds. While the findings of the study contradicted the hypothesis and no effect of mental imagery was found in relation to processing and reading times, mental imagery was found to affect visual narrative processing in an alternative way.

To test this participants panel by panel reading times for both coherent and incoherent three panel comics were recorded. Their self reported mental imagery ability was also recorded using the VVIQ. As suggested, the findings did not provide support for the idea that mental imagery plays a role in connecting the content depicted in comic panels. To support the hypothesis it was expected that there would be significant interaction effects between imagery and coherence, or imagery and panel position revealing a negative relationship between imagery and reading times, with higher mental imagery ability being associated with faster panel reading speeds. However, these interaction effects were not observed. In fact, in contrast to the hypothesis, the results actually indicated a positive relationship between mental imagery and visual narrative processing where slower (or higher) reading speeds were associated with higher mental imagery ability and faster (or lower) reading speeds were associated with lower mental imagery. This is a significant finding that reflects interestingly on the association between mental imagery and visual narrative processing.

This positive relationship between mental imagery and reading speed was somewhat unexpected. Importantly, there were no specific instructions conveyed to the participants in this study regarding the use of mental imagery meaning that this finding reflects a relatively naturalistic use of mental imagery during visual narrative processing. This finding contrasts significantly with previous research which has demonstrated a link between higher mental imagery and faster reading speeds in text based narratives (Suggate & Lenhard, 2022) and arguably demonstrates a significant divergence in the functionality of mental imagery in processing different types of narratives.

The current finding indicates that participants with higher mental imagery were perhaps engaging in a greater degree of mental simulation than participants with lower mental imagery ability, however, this did not appear to be related to processing behaviours. Despite this, the finding aligns with research that suggests that lower mental imagery is related to lower reading engagement (Speed et al., 2024; Williams & Suggate, 2024). This finding is interesting given that it extends this concept to static visual narratives and adds a novel dimension to understanding visual narrative processing.

This interpretation is also interesting to consider in relation to the concept often advanced in arguments against the use of materials like comics in educational settings, that static visual narrative are commonly considered to be lower quality reading materials (Aleixo et al., 2020). This argument is often centred around the idea that because visual narratives use depictive content and are understood quickly (Hagmann & Cohn, 2016; Potter et al., 2014) they are less taxing (Coderre, 2020) and or perhaps have some role in reducing imaginative capacity. However, these findings - showing that higher mental imagery is linked to greater engagement in visual narrative processing - challenge that notion. That the finding mirrors those of studies which directly assess the relationship between mental imagery and text based narratives indicates that the modality of communication is not necessarily the determining factor in engagement. Instead, cognitive abilities related to reading and processing narratives overall play an important role in engagement. This finding then arguably goes some way towards dispelling the concept that static visual narratives provoke less engagement than text , and suggests that, similarly to text based narratives, mental imagery enhances engagement with static visual narratives in the same way as text.

Although mental imagery was found to influence the overall reading times of comics, no interaction effects were observed between imagery ability and either coherence or panel position. This suggests that while individuals with higher imagery ability may engage more deeply with visual narrative sequences, it also suggests that mental imagery ability does not appear to modulate processing of coherent and incoherent static visual narratives. This is in contrast to the suggestion of Suggate & Lenhard (2022) that mental imagery plays a critical role, and predictive role in reading processes. This finding arguably also presents some challenge to the concept that mental imagery plays a role in closure (McCloud, 1993).

Given the concept of closure relates to connecting static pieces of visual information across panels, it seems intuitively logical that increasing the level of incoherence in a comic would interact with the readers mental imagery ability. As such, if the closure process relied on mental imagery, then an interaction between imagery ability and incoherence would have been expected as the reader would need to mentally bridge unconnected scenes. The results of experiment four, which demonstrated apparent modality specific processes in visual narrative situation model generation, seemed to support this idea. The congruence between the visual modality of comics and mental imagery should have influenced processing, particularly with incoherent comics. However, the results of the study show that although participants with higher mental imagery seemed more engaged by the comics, this doesn't appear to have been affected by incoherence. Therefore, while further research should affirm this, the process of closure may not be closely related to mental imagery.

The findings pose a similar contrast to those of Boerma et al., (2016). The hypothesis that faster reading speeds would be associated with mental imagery holds within it the suggestion that mental imagery is essentially critical to comprehension of visual narrative processing. However, the lack of interaction effects between imagery and coherence suggest that mental imagery may not be fundamentally connected to visual narrative processing as first thought.

The somewhat unexpected findings of this experiment are subject to at least two limitations that should be considered in future research which explores the relationship between mental imagery and visual narrative processing. The first limitation, which simultaneously conveys a strength, relates to the use of simple visual narratives. On the one hand, the simplicity of the narratives can be considered a strength, as the experiment has demonstrated an effect of mental imagery even with simple, naturalistic comic stimuli. This serves to highlight the potential for further research into this relationship.

However, the simplicity of the narratives also introduces a limitation. The use of basic uncomplicated narratives may have reduced the cognitive demands placed on the participants, but consequently this may have limited the ability to detect more nuanced or complex effects of mental imagery. While these simple stimuli have effectively outlined the impact of mental imagery on engagement in comics, deeper reasoning may not have been activated. This is perhaps more likely to be the case in relation to the incoherent comics. As in experiment four, after participants had realised that the comic sequence was incoherent, they might have disengaged from efforts to create a coherent sequence from the comics and similarly to experiments five and six, further experiments in this area would arguably benefit from graded coherence. Consequently, while the findings remain valid, the simplicity of the stimuli may have obscured interactions between mental imagery and processing visual narratives.

A further weakness of the current study is the use of self-report mental imagery measures. Though, as noted in the method section, the VVIQ is both widely used and has been validated (Campos & Perez-Fabello, 2009; Jankowska & Karwowski, 2022), self report measures of mental imagery are considered to be less optimal than behavioural metrics. Along with their findings that mental imagery ability predicts reading performance, Suggate & Lenhard (2022) also demonstrate a lack of findings related to self-report measures. However, though this arguably is a limitation, this experiment adds to a growing body of research which, contrary to their suggestion, indicates that these measures can be deployed with reasonable precision in experimental paradigms (Tabi et al., 2020; Zeman et al., 2015). However, further research in this field would arguably benefit from a greater battery of mental imagery ability testing.

As this represents one of the first experiments to explore how the concepts of mental imagery and visual narrative processing are related from a cognitive perspective, several research directions could follow this work. Previous work which explores the role of mental imagery in text-based narratives suggests that this cognitive function plays a critical role in processing narratives (Suggate & Lenhard, 2022) where other research suggests that mental plays a significant role in emotional processing. This may then indicate that mental imagery has narrative-modality based specific functions. Further experiments designed to explore these functions would provide findings that would expand understanding of visual narrative processing, mental imagery, and related functionality. This then highlights the potential of comics as experimental stimuli in further investigations of mental imagery.

Overall, the findings of the current experiment indicate that mental imagery functions as a supplementary process in visual narrative comprehension. Contrary to the initial hypothesis that higher mental imagery ability would be associated with faster panel viewing times due to enhanced predictive processing, it was demonstrated that better mental imagery ability was actually related to increased panel viewing times. This indicates that individuals with stronger mental imagery ability potentially engage more deeply with visual narratives, leading to longer processing times. Given the role of mental imagery in visual narrative processing is a highly under-explored area, further research is necessary to confirm these findings and better understand its contribution to comics processing. This study then highlights the untapped potential of comics as experimental stimuli to understand mental imagery processes and demonstrates that this area is fertile ground for additional investigation.

Chapter 5 - General Discussion

The study of narrative comprehension has primarily been conducted from the perspective of text-based narratives, with this kind of processing now being well documented (Cohn & Wittenberg, 2015; Brich et al,. 2024; Cohn et al., 2024). However, until recently, similar work on visual narratives had yet to be considered. This thesis aimed to investigate the roles of working memory and mental imagery in visual narrative situation model processing. Specifically, the research aimed to determine how verbal and visuospatial components of working memory contribute to visual narrative processing alongside mental imagery. The findings discussed offer methodological advancements and new insights into how individuals process image sequences as narratives and the processes underpinning the construction of meaning from them.

5.1 Summary of Key Findings

A number of interesting findings arose from the experiments conducted in this thesis which used a series of behavioural metrics to explore the use of working memory and situation model processing in visual narratives. Experiments 1 - 3 found that when individual differences in working memory are controlled, verbal working memory does not appear to be implicated in processing a visual narrative sequence. The main finding from these experiments is the suggestion that visuospatial working memory provides the primary processing capacity for visual narrative processing whereas no evidence was found for the activation of verbal working memory during visual narrative processing.

Experiment 4 confirmed that the short visual narratives used in these experiments are processed under the situation model processing framework (Zwaan & Radvansky, 1998; van Dijk & Kintsch, 1988). This was demonstrated using reading times that showed coherent comics were read faster than incoherent comics. Additionally, a recognition task based on the comics seen during the reading time task suggested that participants engaged in situation model updating behaviours, although this had certain limitations. These findings support the suggestion that visual narrative processing employs processes parallel to text-based linguistic processing in that visual narratives invoke a dynamic, cognitive representation of the narrative that people build to comprehend the narrative.

Experiments 5 and 6 explored the use of working memory in situation model processing of visual narratives. The results of these experiments indicate that neither visuospatial working memory nor verbal working memory are directly associated with situation model processing. This suggests that while working memory supports aspects of visual narrative comprehension, based on the current results, it appears to be outside the domain of forming or updating visual narrative situation models during narrative processing.

Finally, Experiment 7 tested the relationship between mental imagery and visual narrative processing and found that contrary to the hypotheses proposed, better mental imagery ability was associated with longer reading times. This is interpreted as indicating greater engagement with the narrative. Though this was not tested, it could suggest that participants with more vivid mental imagery ability may have constructed more detailed representations of the visual sequences. This unexpected result implies that mental imagery ability has the potential to influence the depth with which people process visual narratives, even if it does not necessarily make visual narrative processing more efficient.

5.2 Visuospatial and Verbal Working Memory in Visual Narrative Processing

The hypotheses of Experiments 1-3 were based on the findings of Magliano et al. (2016), which indicate that both verbal and visuospatial working memory are involved in visual narrative processing. In the case of Magliano et al. (2016) and Zhao et al. (2024), understanding visual narrative processing and working memory has relied on reading times of visual narrative sequences as the primary metric. In contrast, the current experiments assessed participants' ability to accurately maintain and recall stimuli that specifically required the use of different components of working memory. The main hypothesis was that participants' ability to maintain visuospatial working memory stimuli would be affected to a significantly greater degree by the requirement to process a comic than verbal working memory stimuli.

Magliano et al. (2016) argued that both visuospatial and verbal working memory are activated during visual narrative processing, as evidenced by changes in reading times when participants were required to retain a verbal working memory stimulus. However, the findings of the current experiments challenge the notion of verbal working memory activation in visual narrative processing. While the results confirm that visuospatial working memory is indeed activated during the reading of visual narrative sequences, aligning with Magliano et al. (2016), they diverge from their findings in significant ways. Not only was visuospatial working memory found to be activated to a greater degree than verbal working memory, as evidenced by the significant interference effects observed, but verbal working memory did not appear to be activated at all. These results suggest that verbal working memory resources are not required for visual narrative processing.

Based on these findings, Experiments 5 and 6 posed related hypotheses. It was expected that visuospatial working memory resources would be implicated in creating and developing a visual narrative situation model. However, it was hypothesised that verbal working memory resources would not contribute. These experiments confirmed that verbal working memory does not contribute to visual narrative processing. However, unexpectedly, it was also demonstrated that though visuospatial working memory does not appear to contribute to visual narrative processing at the situation model level. While these hypotheses were not supported, the findings did corroborate those of Experiments 1 to 3. Visuospatial working memory resources were again found to be significantly affected by processing comics as a dual-task stimulus. In contrast, verbal working memory dual-task performance was similar to a control condition in which participants had no secondary task.

5.3 Theoretical Implications for Visuospatial and Verbal Working Memory

The findings from Experiments 1-3, 5, and 6 present a third perspective on the cognitive processes that underpin visual narrative processing. Whereas Magliano et al. (2016) implicated both visuospatial and verbal working memory, and Zhao et al. (2024) contend that visuospatial working memory is not involved in visual narrative processing, the current findings suggest that visual narrative processing is predominantly supported by visuospatial working memory. This represents an intuitive and parsimonious perspective on visual narrative processing. Previous research on working memory has outlined distinct differences in the use cases of visuospatial and verbal working memory resources (Logie, 1995; Baddeley, 2000). Although it is important to acknowledge that some theorists disagree with the idea of fractionable resources (Morey & Cowan, 2004, Morey, 2018; Vergauwe et al., 2021), the differential use cases are often described along the lines of visual versus verbal tasks (Baddeley et al., 2021). Therefore, it seems more logical that visual narrative processing would rely primarily on visuospatial rather than a combination of visuospatial and verbal working memory.

The evidence presented here, which suggests that verbal working memory is not necessary for visual narrative processing, has important implications for theories of visual narrative comprehension. Notably, the absence of verbal working memory involvement does not undermine the idea that visual narratives are processed using the same general frameworks as those applied to language comprehension. Magliano et al. (2016) argue that verbal working memory contributes evidence for a shared narrative processing framework; the current findings challenge only this specific aspect of their work and do not refute the broader applicability of the processing framework. On the weight of the evidence, visual narratives seem highly likely to adhere to the same cognitive processing frameworks as verbal narratives, where information is processed sequentially and integrated into a dynamic content model. The absence of verbal working memory activation suggests that different working memory resources, specifically visuospatial resources, are utilised to achieve narrative comprehension.

The finding that verbal working memory is not significantly activated in visual narrative processing arguably supports visual language theory (Cohn, 2013) to a greater extent than if verbal working memory were involved. The reliance on visuospatial working memory, evidenced by the significant interference in visuospatial dual-task conditions shown in these experiments, aligns with the theory of visual narrative grammar. This theory suggests that visual narratives are guided by similar principles that influence processing, just as verbal language follows a grammatical structure. The reliance on visuospatial working memory supports an independent but parallel visual language system, which operates similarly to verbal language but relies on distinct cognitive mechanisms. Consequently, the findings support the suggestion of a parallel processing architecture for multimodal, or visual narratives (Cohn, 2014).

5.4 Situation Model Processing in Comics

The findings strongly support the idea of a unified narrative processing framework applicable across modalities. Several hypotheses were posed in Experiment 4, aiming to investigate situation model processing in the visual narratives used in the experiments. It was expected that coherence would facilitate reading speed. In line with this, it was hypothesised that coherent comics would be read faster than incoherent comics. It was also expected that the panels of a coherent comic would be read with increasing speed across the sequence. There was also expected to be no significant difference in reading times for initial panels. Concerning the incoherent comics, it was hypothesised that there would be no significant downward gradient of reading times as in the coherent comics. Recognition times of panels viewed were also tested. For these times, it was hypothesised that coherent panels would be recognised faster than incoherent panels, that there would be a similar downward gradient in coherent panel recognition times related to updating, and that there would be no significant difference between the incoherent panel recognition times.

The results of Experiment 4 supported the majority of these hypotheses in demonstrating that the visual narratives used across the experiments of this thesis produced behavioural outcomes that are characteristic of situation model processing. This strongly supports the idea that situation model processing can be applied to text and visual narratives and aligns with previous research indicating situation model processing facets in visual narratives (Cohn & Wittenberg, 2015; Brich et al., 2024; Cohn et al., 2024). The clear, significant downward trend in reading times across the coherent panels and the significant differences demonstrated between coherent and incoherent comic sets strongly indicate the effect of facilitative semantic processing that parallels text-based processing. The latter effect, in line with the prediction error hypothesis of event segmentation theory (Zacks et al., 2007; Zacks & Swallow, 2007) demonstrates that as readers progress through the narrative, semantic congruence evokes associatively activated networks of information, meaning that predictions can be made about the direction or general trajectory of the narrative. Semantic congruence in the coherent comics results in those predictions being fulfilled, thus decreasing reading times. This is then supported by an overall significant difference between coherent and incoherent comic sets, which indicates that processing requires greater effort and, thus, more time when predictions are not fulfilled. This influence of semantic processing potentially explains the seemingly idiosyncratic findings of Zhao et al. (2024) which indicate no association of visuospatial working memory in visual narrative processing.

As demonstrated in Experiment 4, panel reading times are useful in identifying semantic processing during visual narrative comprehension. However, the significant influence of semantic processing on reading times raises questions about their reliability as a metric for assessing the activation of working memory resources. In particular, the facilitative prediction observed in Experiment 4, as supported by Zhao et al. (2024), highlights how semantic cues can drive reading efficiency independently of working memory demands. Furthermore, reading times in visual narratives are also influenced by image processing efficiency (Potter et al., 2014; Hagmann & Cohn, 2016), which likely introduces additional variability unrelated to working memory. These compounding factors undermine the assumption that reading times can directly or accurately reflect the activation of verbal or visuospatial working memory.

5.5 Working Memory in Visual Narrative Situation Model Processing

The findings of the thesis also have interesting ramifications for the theoretical use of working memory in visual narrative situation models. As outlined in the introduction to Experiment 5, the relationship between working memory and situation model processing is one that is rooted more often in assumption than on an evidential basis. Some experiments have demonstrated evidence in favour of the association (Sargent et al., 2013), whereas others have demonstrated a lack of relationship (Radvansky & Copeland, 2001; 2004; 2006). Regardless, working memory is consistently cited in prominent narrative, discourse or event processing models as a critical component of the processing framework (Zwaan & Radvansky, 1998; van Dijk & Kintsch, 1988; Zacks et al., 2007; Zacks & Swallow, 2007).

Understanding of working memory use in visual narrative situation model processing is currently limited. As referenced throughout the thesis, Magliano et al., (2016) provides primary evidence of the use of working memory in visual narrative situation model processing. However, the assumption of working memory use in situation model processing is also prevalent here. This and other research led to the development of SPECT (Loschky et al., 2020), which proposes a generalised framework for processing visual narratives. This theory outlines the use of both 'front end' or perceptual processes and 'back end' or what could be described as semantic processes. The theory posits that the function of working memory in visual narrative processing is related to the back end, suggesting that this framework is primarily responsible for constructing and maintaining the situation model.

The first three experiments in this thesis demonstrate the use of working memory in visual narrative situation model processing. The fifth and sixth, using minor methodological improvements and an improved design, then further corroborate the use of visuospatial working memory, but not verbal working memory in this kind of processing. However, the primary hypothesis of these experiments was to test the concept that situation model processing recruits working memory processes. Following the demonstration that incoherence disrupts situation model processing in Experiment 4, Experiments 5 and 6 posed the hypotheses that if situation model processing recruits verbal or visuospatial working memory, then incoherent comics should have caused significantly more interference to working memory stimulus maintenance than coherent comics. Given that the incoherent comics lacked a logical narrative, this should have caused disruption to situation model processing. This is evidenced by the apparent processing facilitation related to coherent narratives demonstrated in Experiment 4, in conjunction with the significant difference found in reading times between coherent and incoherent comics. However, support for these hypotheses was not provided by the results of this experiment.

The lack of support for the hypotheses in Experiments 5 and 6 suggests that the role of working memory in visual narrative situation model processing may be more nuanced than previously considered. Though the findings indicate that visuospatial working memory is involved in visual narrative processing, the lack of any significant effect to working memory of the coherence manipulation suggests that working memory does not appear to be affected by narrative coherence. Consequently, these findings suggest that the involvement of working memory in visual narrative processing appears to be limited to perceptual processing. This then suggests, in contrast to the position of SPECT (Loschky et al., 2016), that visuospatial working memory would appear to support front-end or perceptual processes to a greater degree than back-end processes.

As is described in the discussion of Experiment 5, these findings could also be interpreted as indicating a need to revisit the concepts posed by situation model processing in relation to visual narratives. The requirement for working memory involvement in situation models derived from narratives communicated through text or verbal language is arguably much clearer than it is for visual narratives. This is because, in typical text reading, processing text requires the reader to essentially process the surface level of the text and subsequently maintain that model as a given event state to provide a basis for further information to be integrated while simultaneously processing new information. The product of this is that the surface level of text-based narratives is lost, with numerous studies (Bower & Morrow, 1990; Haenggi et al., 1995; Glenberg et al., 1997) having demonstrated that what is stored in memory is the dynamic, representational situation model and not a representation of the surface level of text. From this perspective, it could be argued that situation models derived from comics have a significantly lower profile of cognitive processing requirements. Perhaps the reason that working memory does not appear to be involved in the narrative processing component of visual narratives, as demonstrated in Experiments 5 and 6, is because the situation model of visual narratives is essentially provided as an immediate product of the surface level. Further research is clearly necessary to support this position more fully; however, this could potentially indicate that while situation model processing theory is applicable across media types as a framework for narrative processing, the resultant models perhaps reflect some modality specificity. In turn, this would support the position of visual language theory (Cohn, 2012) and the concept of a parallel processing architecture for visual narratives (Cohn, 2014).

5.6 The Role of Mental Imagery in Visual Narrative Processing

Arguably, the most significant contributions of this thesis are the findings related to visual narrative processing and mental imagery in the seventh experiment. Simultaneously, these findings were also the most unexpected. The relationship between mental imagery and reading is often seemingly an inferred assumption (Sadoski & Paivio, 2004; Speed et al., 2024; Suggate & Lenhard, 2022; Williams & Suggate, 2024). Based on recent findings which have demonstrated that better mental imagery is associated with greater reading speed and improved comprehension of a text (Suggate & Lenhard, 2022), and findings which indicate that mental imagery enhances picture story comprehension amongst children (Boerma et al., 2016), it was hypothesised that increased mental imagery ability would be related to faster reading times. The results indicated that the reverse was true: more vivid mental imagery ability was associated with slower reading times.

The findings of this study hold several significant theoretical implications, not just for visual narrative processing but also for the broader understanding of mental imagery and narrative comprehension as a whole. Notably, the observation that increasingly vivid mental imagery is associated with slower processing suggests that generating mental images during visual narrative processing demands cognitive resources, which vary across individuals. An interesting aspect of these findings is that in Experiment 7, participants were not explicitly instructed to employ mental imagery while reading the comics. This implies that the observed use of mental imagery likely reflects natural utilisation of this cognitive capacity. Consequently, these results reinforce the idea that when individuals possess greater mental imagery capacity, it can enhance their engagement with and, perhaps, comprehension of visual narratives. This is arguably, another area in which visual narrative processing parallels linguistic processing given that previous research has demonstrated that comprehension is enhanced by mental imagery ability (Suggate & Lenhard, 2022). Although this experiment does not definitively establish the additive effect of mental imagery on visual narrative processing, the findings indicate that stronger mental imagery may promote a deeper level of narrative engagement. This is also reflected in mental imagery literature which indicates similar findings in text-based processing (Speed et al., 2024). However, further research is necessary to provide more robust evidence and to more fully support the conclusions drawn in this thesis concerning mental imagery.

These findings also align with the concept of situation model processing, which, as discussed, is often implicitly connected to mental imagery but is rarely explicitly addressed. As outlined in Chapter 1, situation models represent the highest level of narrative understanding (Zwaan & Radvansky, 1998; van Dijk & Kintsch, 1988; Johnson-Laird, 1983). This theory provides a structural framework under which narratives are comprehended by integrating narrative elements into a coherent mental representation. The observed relationship between mental imagery and engagement in visual narratives suggests that individuals with the ability to generate vivid mental imagery are potentially better equipped to construct and maintain situation models. This is on the basis that enhanced mental imagery promotes deeper engagement (Speed et al., 2024). This then seems likely to generate a more immersive and detailed understanding of the narrative, which in turn contributes to a richer mental simulation. Additionally, the slower reading times observed correspond with the increased cognitive requirements associated with constructing and updating situation models.

5.7 Strengths and Contributions of the Thesis

This thesis provides novel empirical insights into the role of working memory in visual narrative processing. Specifically, the findings affirm that visuospatial working memory plays a critical role in visual narrative comprehension and challenge previous research that suggested that verbal working memory also plays a significant role. These results refine our understanding of the cognitive resources required for processing visual narratives and suggest that the role of working memory might need to be reconsidered in relation to how working memory interacts with situation model processing. By clarifying the division of labour between visuospatial and verbal working memory, the thesis advances theoretical models of narrative comprehension.

The research also introduces two new innovative experimental design features that enhance the validity of these findings. First, the use of a titration procedure to investigate working memory contributions (Doherty & Logie, 2016) allowed for the dual task loads to be tailored to each participant's working memory capacity. This personalised approach ensured that participants were neither over- nor underloaded, minimising variability due to the individual differences in cognitive capacity. In using these procedures, the reliability of the results is substantially increased, meaning that the observed effects more accurately reflect differences in cognitive processing rather than differences in task difficulty, task demands or individual variation.

Second, the thesis contributes an enhanced methodological approach by employing a dual-task paradigm that prioritises accuracy over reading times, in contrast to previous research. This modification offers significant advantages, notably that it provides a more direct assessment of the cognitive resources involved in visual narrative processing with the focus on accuracy allowing for a clearer understanding of the allocation of cognitive resources. Together, these methodological advancements present a robust experimental framework that can be utilised in future research to further explore the cognitive mechanisms underlying visual narrative processing.

Another significant contribution of the thesis is its exploration of mental imagery's role in visual narrative processing. The findings demonstrate that individuals with more vivid mental imagery tend to spend longer viewing the panels of a comic. This suggests that mental imagery is an enhancing feature of visual narrative processing, which is reflected in the idea that longer viewing times are equivalent to greater engagement and possibly greater cognitive load during processing. Beyond this, the use of comics as experimental stimuli highlights their untapped potential as a tool for advancing our understanding of mental imagery processes. In demonstrating comics' effectiveness in engaging cognitive processes related to mental imagery, this thesis underscores their value in both cognitive research and practical applications.

5.8 Limitations of the Study

While this study has successfully demonstrated several interesting and novel findings in relation to visual narrative processing, a number of limitations must be acknowledged. Foremost are the results returned from using the visual pattern test (Della Sala et al., 1999). As is discussed in Experiment 2 (Section 2.2), this experiment intended to explore any potential differentiation between the visual and spatial components of working memory activation in visual narrative processing. As in the other experiments in Chapter 2, this experiment saw participants complete a titration phase of the visual pattern test, followed by a dual-task experimental phase where the secondary task was to process a comic. The results of this experiment could have been interpreted as suggesting that visual working memory is activated to a lesser degree than spatial working memory, as visual pattern dual task accuracy was found to be significantly higher than Corsi block dual task accuracy, meaning that spatial working memory is taxed to a greater degree than visual working memory in comics processing. However, the results also revealed that the control conditions of these tasks appeared not to be equivalent, with the Corsi block control task being performed with significantly lower accuracy than the visual pattern test.

Consequently, the differing baselines between the control conditions of the visual pattern test and the Corsi block task present an issue to the extent that conclusions can be drawn about visual and spatial activation. Though these tasks have been directly compared in previous research (Della Sala et al., 1999), the inherent difficulty levels of the two tasks in this experiment arguably make them somewhat incomparable. As such, while the data points to the idea that spatial working memory might be involved in visual narrative processing to a greater extent than visual working memory, this interpretation should be considered cautiously due to the unequal baselines.

A further limitation of this project is related to the results of Experiment 7. The finding that greater mental imagery is related to longer reading times in visual narratives is a significant and novel contribution to the field of visual narrative processing in that it demonstrates a link between areas that have previously not been considered together. However, this experiment's reliance on self-report measures in determining mental imagery ability represents a limitation that should be addressed in future research. Mental imagery is a peculiar concept in that most people intuitively understand what that term means and have experience with the phenomena. However, they are inevitably limited to their own experiences of mental imagery. Because of this, self-report-based measures of mental imagery are arguably limited in their scope as, fundamentally, these measures require a comparison against an external reference point. In the case of mental imagery, this presents a significant issue to the validity of self-report metrics as the responder arguably has no reference point against which to judge their mental imagery ability. In the case of Experiment 7, this factor formed a significant part of the reasoning behind using a median split with the VVIQ. However, while this method arguably mitigates the validity issue to a certain degree, it cannot overcome the inherent subjectivity of self-report data. This limitation means that while the results suggest a link between mental imagery and visual narrative reading times, this link's exact nature and reliability remains somewhat uncertain.

Finally, an overarching limitation of the current project is its reliance on behavioural outcomes as the primary measures of cognitive processing. While the behavioural methods used in these experiments have provided valuable data on participants' cognitive interactions with visual narratives, they can only offer indirect inferences about the underlying cognitive mechanisms. This reliance on observable behaviour then necessitates interpretative analysis, which has the potential to lead to over or misinterpretation. Reading times, for example, can indicate deeper cognitive engagement or processing difficulty, but further evidence, such as EEG data (Cohn & Kutas, 2015; 2017), would have provided a complementary framework which could have informed interpretation and refined the results.

One enhancement that could have strengthened the results and is suggested as a follow-up to the thesis' findings concerning visuospatial working memory is an investigation of eye movements. Oculomotor movement has been linked to the use of visuospatial working memory in studies showing that making voluntary eye movements can disrupt the contents of, specifically, spatial working memory (Postle et al., 2006). This has the potential to explain some of the variation and use of visuospatial activation in working memory during visual narrative processing. However, the relationship between visuospatial working memory and eye movements is complex, with recent research demonstrating that visuospatial working memory contents influence eye movements towards primed content (Bahle et al., 2017). This would be interesting to combine with the current accuracy plus coherence manipulation setup to explore semantic processing in comics further.

5.9 Future Research Directions

One of the major contributions of this thesis is the development of an innovative experimental method aimed at testing working memory and situation model processing within the context of comics reading. By adapting the methodology from Magliano et al. (2016), wherein the working memory task is positioned as the primary task and visual narrative processing as the secondary, this work has laid a foundation for a broader exploration of the interplay between working memory activation and comprehension processes in comics.

A primary direction for future research would involve replicating Magliano et al.'s (2016) study with the adjusted methodology used in the experiments of this thesis. This would arguably provide a more explicit examination of how inference-making is influenced by working memory. One challenge associated with replicating this type relates to the control of panel presentation time during dual-task conditions. In employing reading times, the original study by Magliano et al., (2016) bypasses this but, as discussed above, arguably confounds the investigation of working memory processes because of this metrics inextricable link to semantic processing when applied to stimulus containing narrative. As such, a study like this would need to consider the length of time the panels were shown to participants and ensure control of this, given that time decay is a significant factor in working memory (Barouillet et al., 2004; Vergauwe et al., 2021).

Two potential solutions are proposed to mitigate this issue. First, conducting a panel viewing time norming study, which could run in essentially the same manner as the fourth experiment in this thesis, could be used to establish average panel viewing times. These could then be used to standardise the presentation time of panels across all participants. Alternatively, an individualised approach involving a programmed procedure could be implemented in which panel presentation times during the dual task portion of the experiment could be programmatically calibrated in an initial phase, similar to the titration procedure used in the current experiments. The participants' viewing times of an initial series of panel sets could be recorded, and these personalised viewing times could then be applied in the main experimental phase, thereby using person-specific system pacing to control presentation times and, thus, time-based influence on working memory. With the application of these methods, the visual narrative manipulation of Magliano et al. (2016), in which bridging events are removed, forcing the participants to generate significant inference, could be deployed to assess the role of working memory in inference-making during visual narrative processing.

A further significant contribution of this thesis is the finding of an association between more vivid mental imagery and slower reading speeds in comic panels. This relationship suggests that readers who experience stronger mental imagery take more time to process each panel, possibly because vivid imagery requires greater cognitive resources. Future research could focus on how mental imagery affects the processing of different visual narrative features. Concepts such as the depiction of time, motion, and speed have the potential to provide both highly interesting and informative findings.

One clear potential line of research here relates to panel transitions. Static visual narratives like comics communicate information by juxtaposing narratively linked images in sequence. The narratives conveyed through this kind of visual means are naturally impoverished by necessity of the medium, meaning that relatively little of the narrative is shown. This is, of course, variable in relation to what the sequence is trying to convey, and while not a universal principle, it is broadly how comics operate. Though it is now reasonably well established that comics narratives are processed under the situation model processing framework (Cohn et al., 2024; Cohn & Wittenberg, 2015; Brich et al., 2024), the cognitive operations related to how panels are connected and processed as occurring in a sequence should be studied further. Mental imagery is intuitively appealing as a function which has the potential to explain the ability to explain panel processing behaviours.

Mental imagery ability and the findings presented in this thesis then offer some initial findings related to how these processes may function. The association between vivid mental imagery and longer reading times suggests that readers with greater mental imagery ability may engage more deeply with the cognitive process of filling in the gaps between panels. Future studies in this area could then explore how individual differences in mental imagery ability affect panel connection by manipulating the size of the inferential connection required.

5.10 Conclusion

This research has explored the cognitive mechanisms underpinning visual narrative comprehension, focusing on the roles of working memory and mental imagery. The findings have demonstrated that visuospatial working memory primarily supports visual narrative processing. In contrast to previous research, the results have also demonstrated that verbal working memory is not significantly engaged during image sequence processing. The research provides further support for the concept that visual narrative processing aligns significantly with how verbal narratives are known to be processed under the situation model processing framework. More vivid mental imagery ability was also associated with longer reading times, which suggests deeper narrative engagement rather than greater cognitive efficiency. These findings refine theoretical models of visual narrative processing and support the notion of visual language processing as an independent but parallel system to verbal language, underpinned by visuospatial working memory. Methodologically, the study's innovative dual-task paradigms and combination with coherence while maintaining a focus on dual-task accuracy and participant-specific titration procedures enhance the precision of cognitive resource measurement and contribute valuable tools for future research.

Despite its strengths and contributions, the thesis also acknowledges limitations, including a reliance on behavioural metrics and self-report-based measures of mental imagery. These issues highlight the need for further research to complement and support the findings using methods like EEG and other increasingly objective measures of mental imagery. Future research could explore how mental imagery bridges inferential gaps in panel transitions and investigate the nuanced relationship between working memory and narrative coherence. By challenging existing assumptions and offering methodological advancements, this research provides a foundation for a deeper understanding of visual narrative processing and its cognitive underpinnings.

References

Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C., & de Lange, F. P. (2013). Shared Representations for Working Memory and Mental Imagery in Early Visual Cortex. *Current Biology*, *23*(15), 1427–1431. https://doi.org/10.1016/j.cub.2013.05.065

Aleixo, P., Matkin, D., & Kilby, L. (2020). What do teachers think about the educational role of comic books?: A qualitative analysis. *Studies in Comics*, *11*, 387–404. https://doi.org/10.1386/stic_00037_1

Allen, K., Higgins, S., & Adams, J. (2019). The Relationship between Visuospatial Working Memory and Mathematical Performance in School-Aged Children: A Systematic Review. *Educational Psychology Review*, *31*(3), 509–531. https://doi.org/10.1007/s10648-019-09470-8

Alloway, T. P., Gathercole, S. E., Kirkwood, H., & Elliott, J. (2009). The Cognitive and Behavioral Characteristics of Children With Low Working Memory: *Child Development*, *80*(2), Article 2. https://doi.org/10.1111/j.1467-8624.2009.01282.x

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*(11), Article 11. https://doi.org/10.1016/S1364-6613(00)01538-2

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, *4*, 829–839. https://doi.org/10.1038/nrn1201

Baddeley, A. (2012). Working Memory: Theories, Models, and Controversies. *Annual Review of Psychology*, *63*(1), Article 1. https://doi.org/10.1146/annurev-psych-120710-100422

Baddeley, A. D. (2007). *Working memory, thought, and action*. Oxford University Press.

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

Baddeley, A., Hitch, G., & Allen, R. (2021). A Multicomponent Model of Working Memory. In *Working Memory: State of the Science* (First). Oxford University Press.

Bahle, B., Beck, V. M., & Hollingworth, A. (2018). The architecture of interaction between visual working memory and visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(7), 992–1011. https://doi.org/10.1037/xhp0000509

Bailey, H. R., Kurby, C. A., Sargent, J. Q., & Zacks, J. M. (2017). Attentional focus affects how events are segmented and updated in narrative reading. *Memory & Cognition*, *45*(6), 940–955. https://doi.org/10.3758/s13421-017-0707-2

Bailey, H. R., & Zacks, J. M. (2015). Situation model updating in young and older adults: Global versus incremental mechanisms. *Psychology and Aging*, *30*(2), 232–244. https://doi.org/10.1037/a0039081

Balaban, H., Smith, K., Tenenbaum, J., & Ullman, T. (2023). Intuitive physics guides visual tracking and working memory: The dynamics of neural processing in expectation violation. *Journal of Vision*, *23*(9), 4903. https://doi.org/10.1167/jov.23.9.4903

Barrett, L. F., Tugade, M. M., & Engle, R. W. (2004). Individual Differences in Working Memory Capacity and Dual-Process Theories of the Mind. *Psychological Bulletin*, *130*(4), 553–573. https://doi.org/10.1037/0033-2909.130.4.553

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time Constraints and Resource Sharing in Adults' Working Memory Spans. *Journal of Experimental Psychology: General*, *133*(1), 83–100. https://doi.org/10.1037/0096-3445.133.1.83

Bates, K. E., & Farran, E. K. (2021). Mental imagery and visual working memory abilities appear to be unrelated in childhood: Evidence for individual differences in strategy use. *Cognitive Development*, *60*, 101120. https://doi.org/10.1016/j.cogdev.2021.101120

Bates, K. E., Smith, M. L., Farran, E. K., & Machizawa, M. G. (2024). Behavioral and Neural Correlates of Visual Working Memory Reveal Metacognitive Aspects of Mental Imagery. *Journal of Cognitive Neuroscience*, *36*(2), 272–289. https://doi.org/10.1162/jocn_a_02085

Bestgen, Y., & Vonk, W. (2000). Temporal Adverbials as Segmentation Markers in Discourse Comprehension. *Journal of Memory and Language*, *42*(1), 74–87. https://doi.org/10.1006/jmla.1999.2670

Blackwell, S. E., Browning, M., Mathews, A., Pictet, A., Welch, J., Davies, J., Watson, P., Geddes, J. R., & Holmes, E. A. (2015). Positive Imagery-Based Cognitive Bias Modification as a Web-Based Treatment Tool for Depressed Adults: A Randomized Controlled Trial. *Clinical Psychological Science*, *3*(1), 91–111. https://doi.org/10.1177/2167702614560746

Blanca, M. J., Alarcón, R., & Arnau, J. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema*, *29.4*, 552–557. https://doi.org/10.7334/psicothema2016.383

Bobrowicz, K., Weber, A., & Greiff, S. (2024). The successful use of a search strategy improves with visuospatial working memory in 2- to 4.5-year-olds. *Journal of Experimental Child Psychology*, *238*, 105786. https://doi.org/10.1016/j.jecp.2023.105786

Boerma, I. E., Mol, S. E., & Jolles, J. (2016). Reading Pictures for Story Comprehension Requires Mental Imagery Skills. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01630

Bopp, K. L., & Verhaeghen, P. (2005). Aging and Verbal Memory Span: A Meta-Analysis. *The Journals of Gerontology: Series B*, *60*(5), P223–P233. https://doi.org/10.1093/geronb/60.5.P223

Bower, G. H., & Morrow, D. G. (1990). Mental Models in Narrative Comprehension. *Science*, *247*(4938), 44–48. https://doi.org/10.1126/science.2403694

Bower, G. H., & Rinck, M. (2001). Selecting one among many referents in spatial situation models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(1), 81–98. https://doi.org/10.1037/0278-7393.27.1.81

Brich, I. R., Papenmeier, F., Huff, M., & Merkt, M. (2024). Construction or updating? Event model processes during visual narrative comprehension. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-023-02424-w

Brockmole, J. R., & Logie, R. H. (2013). Age-Related Change in Visual Working Memory: A Study of 55,753 Participants Aged 8–75. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00012

Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about Semantic Illusions: Rethinking the functional role of the P600 in language comprehension. *Brain Research*, *1446*, 127–143. https://doi.org/10.1016/j.brainres.2012.01.055

Brown, L. A., Forbes, D., & McConnell, J. (2006). Short Article: Limiting the use of Verbal Coding in the Visual Patterns Test. *Quarterly Journal of Experimental Psychology*, *59*(7), 1169–1176. https://doi.org/10.1080/17470210600665954

Brunetti, R., Del Gatto, C., & Delogu, F. (2014). eCorsi: Implementation and testing of the Corsi block-tapping task for digital tablets. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00939

Brunyé, T. T., & Taylor, H. A. (2008). Working memory in developing and applying mental models from spatial descriptions. *Journal of Memory and Language*, *58*(3), 701–729.https://doi.org/10.1016/j.jml.2007.08.003

Campos, A., & Pérez-Fabello, M. J. (2009). Psychometric Quality of a Revised Version Vividness of Visual Imagery Questionnaire. *Perceptual and Motor Skills*, *108*(3), 798–802. https://doi.org/10.2466/pms.108.3.798-802

Cansino, S., Hernández-Ramos, E., Estrada-Manilla, C., Torres-Trejo, F., Martínez-Galindo, J. G., Ayala-Hernández, M., Gómez-Fernández, T., Osorio, D., Cedillo-Tinoco, M., Garcés-Flores, L., Beltrán-Palacios, K., García-Lázaro, H. G., García-Gutiérrez, F., Cadena-Arenas, Y., Fernández-Apan, L., Bärtschi, A., & Rodríguez-Ortiz, M. D. (2013). The decline of verbal and visuospatial working memory across the adult life span. *AGE*, *35*(6), 2283–2302. https://doi.org/10.1007/s11357-013-9531-1

Cansino, S., Torres-Trejo, F., Estrada-Manilla, C., Pérez-Loyda, M., Ramírez-Barajas, L., Hernández-Ladrón-deGuevara, M., Nava-Chaparro, A., & Ruiz-Velasco, S. (2020). Predictors of Working Memory Maintenance and Decline in Older Adults. *Archives of Gerontology and Geriatrics*, *89*, 104074. https://doi.org/10.1016/j.archger.2020.104074

Claessen, M. H. G., van der Ham, I. J. M., & van Zandvoort, M. J. E. (2015). Computerization of the Standard Corsi Block-Tapping Task Affects Its Underlying Cognitive Concepts: A Pilot Study. *Applied Neuropsychology: Adult*, *22*(3), 180–188. https://doi.org/10.1080/23279095.2014.892488

Cocchini, G., Logie, R. H., Sala, S. D., MacPherson, S. E., & Baddeley, A. D. (2002). Concurrent performance of two memory tasks: Evidence for domain-specific working memory systems. *Memory & Cognition*, *30*(7), 1086–1095. https://doi.org/10.3758/BF03194326

Coderre, E. L. (2020). Dismantling the "Visual Ease Assumption:" A Review of Visual Narrative Processing in Clinical Populations. *Topics in Cognitive Science*, *12*(1), Article 1. https://doi.org/10.1111/tops.12446

Coderre, E. L., & Cohn, N. (2024). Individual differences in the neural dynamics of visual narrative comprehension: The effects of proficiency and age of acquisition. *Psychonomic Bulletin & Review*, *31*(1), 89–103. https://doi.org/10.3758/s13423-023-02334-x

Cohn, N. (2013). *The Visual Language of Comics: Introduction to the Structure and Cognition of Sequential Images.* Bloomsbury.

Cohn, N. (2015). Narrative conjunction's junction function: The interface of narrative grammar and semantics in sequential images. *Journal of Pragmatics*, *88*, 105–132. https://doi.org/10.1016/j.pragma.2015.09.001

Cohn, N. (2020). *Who understands comics? Questioning the universality of visual language comprehension*. Bloomsbury Academic.

Cohn, N. (2021). A starring role for inference in the neurocognition of visual narratives. *Cognitive Research: Principles and Implications*, *6*(1), 8. https://doi.org/10.1186/s41235-021-00270-9

Cohn, N., & Bender, P. (2017). Drawing the line between constituent structure and coherence relations in visual narratives. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(2), Article 2. https://doi.org/10.1037/xlm0000290

Cohn, N., & Foulsham, T. (2022). Meaning above (and in) the head: Combinatorial visual morphology from comics and emoji. *Memory and Cognition*, *50*(7), 1381–1398.https://doi.org/10.3758/s13421-022-01294-2

Cohn, N., Jackendoff, R., Holcomb, P. J., & Kuperberg, G. R. (2014). The grammar of visual narrative: Neural evidence for constituent structure in sequential image comprehension. *Neuropsychologia*, *64*, 63–70. https://doi.org/10.1016/j.neuropsychologia.2014.09.018

Cohn, N., & Kutas, M. (2015). Getting a cue before getting a clue: Event-related potentials to inference in visual narrative comprehension. *Neuropsychologia*, *77*, 267–278. https://doi.org/10.1016/j.neuropsychologia.2015.08.026

Cohn, N., & Kutas, M. (2017). What's your neural function, visual narrative conjunction? Grammar, meaning, and fluency in sequential image processing. *Cognitive Research: Principles and Implications*, *2*(1), 27. https://doi.org/10.1186/s41235-017-0064-5

Cohn, N., Murthy, B., & Foulsham, T. (2016). Meaning above the head: Combinatorial constraints on the visual vocabulary of comics. *Journal of Cognitive Psychology*, *28*(5), 559–574. https://doi.org/10.1080/20445911.2016.1179314

Cohn, N., & Paczynski, M. (2013). Prediction, events, and the advantage of Agents: The processing of semantic roles in visual narrative. *Cognitive Psychology*, *67*(3), 73–97. https://doi.org/10.1016/j.cogpsych.2013.07.002

Cohn, N., Paczynski, M., Jackendoff, R., Holcomb, P. J., & Kuperberg, G. R. (2012). (Pea)nuts and bolts of visual narrative: Structure and meaning in sequential image comprehension. *Cognitive Psychology*, *65*(1), Article 1. https://doi.org/10.1016/j.cogpsych.2012.01.003

Cohn, N., Paczynski, M., & Kutas, M. (2017). Not so secret agents: Event-related potentials to semantic roles in visual event comprehension. *Brain and Cognition*, *119*, 1–9. https://doi.org/10.1016/j.bandc.2017.09.001

Cohn, N., Van Middelaar, L., Foulsham, T., & Schilperoord, J. (2024). Anaphoric distance dependencies in visual narrative structure and processing. *Cognitive Psychology*, *149*, 101639. https://doi.org/10.1016/j.cogpsych.2024.101639

Cohn, N., & Wittenberg, E. (2015). Action starring narratives and events: Structure and inference in visual narrative comprehension. *Journal of Cognitive Psychology*, *27*(7), 812–828. https://doi.org/10.1080/20445911.2015.1051535

Cuder, A., Živković, M., Doz, E., Pellizzoni, S., & Passolunghi, M. C. (2023). The relationship between math anxiety and math performance: The moderating role of visuospatial working memory. *Journal of Experimental Child Psychology*, *233*, 105688. https://doi.org/10.1016/j.jecp.2023.105688

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, *3*(4), 422–433. https://doi.org/10.3758/BF03214546

Darling, S., & Havelka, J. (2010). Visuospatial bootstrapping: Evidence for binding of verbal and spatial information in working memory. *Quarterly Journal of Experimental Psychology*, *63*(2), Article 2. https://doi.org/10.1080/17470210903348605

Darling, S., Havelka, J., Allen, R. J., Bunyan, E., & Flornes, L. (2020). Visuospatial bootstrapping: Spatialized displays enhance digit and nonword sequence learning. *Annals of the New York Academy of Sciences*, *1477*(1), 100–112. https://doi.org/10.1111/nyas.14429

Dawes, A. J., Keogh, R., Andrillon, T., & Pearson, J. (2020). A cognitive profile of multi-sensory imagery, memory and dreaming in aphantasia. *Scientific Reports*, *10*(1), 10022. https://doi.org/10.1038/s41598-020-65705-7

De Beni, R., Pazzaglia, F., Gyselinck, V., & Meneghetti, C. (2005). Visuospatial working memory and mental representation of spatial descriptions. *European Journal of Cognitive Psychology*, *17*(1), 77–95. https://doi.org/10.1080/09541440340000529

Della Sala, S., Gray, C., Baddeley, A., Allamano, N., & Wilson, L. (1999). Pattern span: A tool for unwelding visuo–spatial memory. *Neuropsychologia*, *37*(10), 1189–1199. https://doi.org/10.1016/S0028-3932(98)00159-6

Delogu, F., Drenhaus, H., & Crocker, M. W. (2018). On the predictability of event boundaries in discourse: An ERP investigation. *Memory & Cognition*, *46*(2), 315–325. https://doi.org/10.3758/s13421-017-0766-4

Dijkstra, N., Hinne, M., Bosch, S. E., & Van Gerven, M. A. J. (2019). Between-subject variability in the influence of mental imagery on conscious perception. *Scientific Reports*, *9*(1), 15658. https://doi.org/10.1038/s41598-019-52072-1

Doherty, J. M., & Logie, R. H. (2016). Resource-sharing in multiple-component working memory. *Memory & Cognition*, *44*(8), 1157–1167. https://doi.org/10.3758/s13421-016-0626-7

Duff, S., & Logie, R. (2001). Processing and Storage in Working Memory Span. *Quarterly Journal of Experimental Psychology*, *54*, 31–48. https://doi.org/10.1080/02724980042000011

Foulsham, T., & Cohn, N. (2021). Zooming in on visual narrative comprehension. *Memory & Cognition*, *49*(3), 451–466. https://doi.org/10.3758/s13421-020-01101-w

Foulsham, T., Wybrow, D., & Cohn, N. (2016). Reading Without Words: Eye Movements in the Comprehension of Comic Strips: Eye movements in comic narratives. *Applied Cognitive Psychology*, *30*(4), Article 4. https://doi.org/10.1002/acp.3229

Friedman, N. P., & Miyake, A. (2000). Differential roles for visuospatial and verbal working memory in situation model construction. *Journal of Experimental Psychology: General*, *129*(1), 61–83. https://doi.org/10.1037/0096-3445.129.1.61

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The Structure of Working Memory From 4 to 15 Years of Age. *Developmental Psychology*, *40*(2), 177–190. https://doi.org/10.1037/0012-1649.40.2.177

Gathercole, S. E., Willis, C. S., Emslie, H., & Baddeley, A. D. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental Psychology*, *28*(5), 887–898. https://doi.org/10.1037/0012-1649.28.5.887

Gernsbacher, M. A., & Faust, M. E. (1991). The mechanism of suppression: A component of general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(2), 245–262. https://doi.org/10.1037/0278-7393.17.2.245

Gernsbacher, M. A., Varner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(3), 430–445. https://doi.org/10.1037/0278-7393.16.3.430

Giofrè, D., Donolato, E., & Mammarella, I. C. (2018). The differential role of verbal and visuospatial working memory in mathematics and reading. *Trends in Neuroscience and Education*, *12*, 1–6. https://doi.org/10.1016/j.tine.2018.07.001

Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language*, *26*(1), 69–83. https://doi.org/10.1016/0749-596X(87)90063-5

Guérard, K., & Tremblay, S. (2008). Revisiting evidence for modularity and functional equivalence across verbal and spatial domains in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(3), 556–569. https://doi.org/10.1037/0278-7393.34.3.556

Gyselinck, V., Beni, R., Pazzaglia, F., Meneghetti, C., & Mondoloni, A. (2007). Working memory components and imagery instructions in the elaboration of a spatial mental model. *Psychological Research*, *71*(3), 373–382. https://doi.org/10.1007/s00426-006-0091-1

Gyselinck, V., Jamet, E., & Dubois, V. (2008). The role of working memory components in multimedia comprehension. *Applied Cognitive Psychology*, *22*(3), 353–374. https://doi.org/10.1002/acp.1411

Gyselinck, V., Meneghetti, C., De Beni, R., & Pazzaglia, F. (2009). The role of working memory in spatial text processing: What benefit of imagery strategy and visuospatial abilities? *Learning and Individual Differences*, *19*(1), 12–20. https://doi.org/10.1016/j.lindif.2008.08.002

Hackmann, A., Clark, D. M., & McManus, F. (2000). Recurrent images and early memories in social phobia. *Behaviour Research and Therapy*, *38*(6), 601–610. https://doi.org/10.1016/S0005-7967(99)00161-8

Haenggi, D., Kintsch, W., & Gernsbacher, M. A. (1995). Spatial situation models and text comprehension. *Discourse Processes*, *19*(2), 173–199. https://doi.org/10.1080/01638539509544913

Hagmann, C. E., & Cohn, N. (2016). The pieces fit: Constituent structure and global coherence of visual narrative in RSVP. *Acta Psychologica*, *164*, 157–164. https://doi.org/10.1016/j.actpsy.2016.01.011

Halford, G. S., Cowan, N., & Andrews, G. (2007). Separating cognitive capacity from knowledge: A new hypothesis. *Trends in Cognitive Sciences*, *11*(6), 236–242. https://doi.org/10.1016/j.tics.2007.04.001

Hintz, F., Voeten, C. C., Dobó, D., Lukics, K. S., & Lukács, Á. (2024). The role of general cognitive skills in integrating visual and linguistic information during sentence comprehension: Individual differences across the lifespan. *Scientific Reports*, *14*(1), 17797. https://doi.org/10.1038/s41598-024-68674-3

Hitch, G. J., Halliday, S., Schaafstal, A. M., & Schraagen, J. M. C. (1988). Visual working memory in young children. *Memory & Cognition*, *16*(2), 120–132. https://doi.org/10.3758/BF03213479

Hoeben Mannaert, L., & Dijkstra, K. (2021). Situation model updating in young and older adults. *International Journal of Behavioral Development*, *45*(5), 389–396. https://doi.org/10.1177/0165025419874125

Huettig, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience*, *31*(1), 80–93. https://doi.org/10.1080/23273798.2015.1047459

Huff, M., Meitz, T. G. K., & Papenmeier, F. (2014). Changes in situation models modulate processes of event perception in audiovisual narratives. *Journal of Experimental Psychology: Learning Memory and Cognition*, *40*(5), 1377–1388. https://doi.org/10.1037/a0036780

Ikkai, A., McCollough, A. W., & Vogel, E. K. (2010). Contralateral Delay Activity Provides a Neural Measure of the Number of Representations in Visual Working Memory. *Journal of Neurophysiology*, *103*(4), 1963–1968. https://doi.org/10.1152/jn.00978.2009

Izmalkova, A., Barmin, A., Velichkovsky, B. B., Prutko, G., & Chistyakov, I. (2022). Cognitive Resources in Working Memory: Domain-Specific or General? *Behavioral Sciences*, *12*(11), Article 11. https://doi.org/10.3390/bs12110459

Jarrold, C., Tam, H., Baddeley, A. D., & Harvey, C. E. (2011). How does processing affect storage in working memory tasks? Evidence for both domain-general and domain-specific effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(3), 688–705. https://doi.org/10.1037/a0022527

Jarrold, C., & Towse, J. N. (2006). Individual differences in working memory. *Neuroscience*, *139*(1), 39–50. https://doi.org/10.1016/j.neuroscience.2005.07.002

Jarvis, H. L., & Gathercole, S. E. (2003). Verbal and non-verbal working memory and achievements on National Curriculum tests at 11 and 14 years of age. *Educational and Child Psychology*, *20*(3), 123–140. https://doi.org/10.53841/bpsecp.2003.20.3.123

Johnson-Laird, P. N. (1983). *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Harvard University Press.

Jonides, J., Smith, E. E., Koeppe, R. A., Awh, E., Minoshima, S., & Mintun, M. A. (1993). Spatial working memory in humans as revealed by PET. *Nature*, *363*(6430), 623–625. https://doi.org/10.1038/363623a0

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review*, *9*(4), 637–671. https://doi.org/10.3758/BF03196323

Kay, L., Keogh, R., & Pearson, J. (2024). Slower but more accurate mental rotation performance in aphantasia linked to differences in cognitive strategies. *Consciousness and Cognition*, *121*, 103694. https://doi.org/10.1016/j.concog.2024.103694

Keogh, R., & Pearson, J. (2018). The blind mind: No sensory visual imagery in aphantasia. *Cortex*, *105*, 53–60. https://doi.org/10.1016/j.cortex.2017.10.012

Keogh, R., Wicken, M., & Pearson, J. (2021). Visual working memory in aphantasia: Retained accuracy and capacity with a different strategy. *Cortex*, *143*, 237–253. https://doi.org/10.1016/j.cortex.2021.07.012

Kessels, R. P. C., van den Berg, E., Ruis, C., & Brands, A. M. A. (2008). The Backward Span of the Corsi Block-Tapping Task and Its Association With the WAIS-III Digit Span. *Assessment*, *15*(4), 426–434. https://doi.org/10.1177/1073191108315611

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, *95*(2), 163–182. https://doi.org/10.1037/0033-295X.95.2.163

Klauer, K. C., & Zhao, Z. (2004). Double dissociations in visual and spatial short-term memory. *Journal of Experimental Psychology. General*, *133 3*, 355–381. https://doi.org/10.1037/0096-3445.133.3.355

Knauff, M., & May, E. (2006). Mental imagery, reasoning, and blindness. *Quarterly Journal of Experimental Psychology*, *59*(1), 161–177. https://doi.org/10.1080/17470210500149992

Kosslyn, S. M. (1973). Scanning visual images:Some structural implications. *Perception & Psychophysics*, *14*(1), Article 1. https://doi.org/10.3758/BF03198621

Kosslyn, S. M., Behrmann, M., & Jeannerod, M. (1995). The cognitive neuroscience of mental imagery. *Neuropsychologia*, *33*(11), 1335–1344. https://doi.org/10.1016/0028-3932(95)00067-D

Kosslyn, S. M., Ball, T. M., & Reiser, B. J. (1978). Visual images preserve metric spatial information: Evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception and Performance*, *4*(1), 47–60. https://doi.org/10.1037/0096-1523.4.1.47

Kuperberg, G. R., Brothers, T., & Wlotko, E. W. (2020). A Tale of Two Positivities and the N400: Distinct Neural Signatures Are Evoked by Confirmed and Violated Predictions at Different Levels of Representation. *Journal of Cognitive Neuroscience*, *32*(1), 12–35. https://doi.org/10.1162/jocn_a_01465

Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends in Cognitive Sciences*, *12*(2), 72–79. https://doi.org/10.1016/j.tics.2007.11.004

Kurby, C. A., & Zacks, J. M. (2012). Starting from scratch and building brick by brick in comprehension. *Memory & Cognition*, *40*(5), 812–826. https://doi.org/10.3758/s13421-011-0179-8

Kutas, M., & Hillyard, S. A. (1980). Reading Senseless Sentences: Brain Potentials Reflect Semantic Incongruity. *Science*, *207*(4427), 203–205. https://doi.org/10.1126/science.7350657

Kwak, S. G., & Kim, J. H. (2017a). Central limit theorem: The cornerstone of modern statistics. *Korean Journal of Anesthesiology*, *70*(2), 144–156. https://doi.org/10.4097/kjae.2017.70.2.144

Kwak, S. G., & Kim, J. H. (2017b). Central limit theorem: The cornerstone of modern statistics. *Korean Journal of Anesthesiology*, *70*(2), 144–156. https://doi.org/10.4097/kjae.2017.70.2.144

Leon Guerrero, S., Whitford, V., Mesite, L., & Luk, G. (2021). Text Complexity Modulates Cross-Linguistic Sentence Integration in L2 Reading. *Frontiers in Communication*, *6*. https://doi.org/10.3389/fcomm.2021.651769

Levi, O., & Heled, E. (2024). Aging Processes of Working Memory in Different Modalities. *Neurology International*, *16*(5), Article 5. https://doi.org/10.3390/neurolint16050084

Li, X., & Qu, Q. (2024). Verbal working memory capacity modulates semantic and phonological prediction in spoken comprehension. *Psychonomic Bulletin & Review*, *31*(1), 249–258. https://doi.org/10.3758/s13423-023-02348-5

Li, Y., & Geary, D. C. (2013). Developmental Gains in Visuospatial Memory Predict Gains in Mathematics Achievement. *PLOS ONE*, *8*(7), e70160. https://doi.org/10.1371/journal.pone.0070160

Lin, Y., & Matsumi, N. (2022). Visuospatial working memory and the construction of a spatial situation model in listening comprehension: An examination using a spatial tapping task. *Cognitive Processing*, *23*(1), 41–54. https://doi.org/10.1007/s10339-021-01063-0

Logie, R. H. (2011). The Functional Organization and Capacity Limits of Working Memory. *Current Directions in Psychological Science*, *20*(4), 240–245. https://doi.org/10.1177/0963721411415340

Logie, R. H. (2016). Retiring the Central Executive. *Quarterly Journal of Experimental Psychology*, *69*(10), 2093–2109. https://doi.org/10.1080/17470218.2015.1136657

Logie, R. H., Camos, V., & Cowan, N. (Eds.). (2021). *Working memory: State of the science* (First edition). Oxford University Press.

Logie, R., & Pearson, D. (1997). The Inner Eye and the Inner Scribe of Visuo-spatial Working Memory: Evidence from Developmental Fractionation. *European Journal of Cognitive Psychology*, *9*, 241–257. https://doi.org/10.1080/713752559

Loschky, L. C., Larson, A. M., Smith, T. J., & Magliano, J. P. (2020). The Scene Perception & Event Comprehension Theory (SPECT) Applied to Visual Narratives. *Topics in Cognitive Science*, *12*(1), 311–351. https://doi.org/10.1111/tops.12455

Magliano, J. P., Kopp, K., Higgs, K., & Rapp, D. N. (2017). Filling in the Gaps: Memory Implications for Inferring Missing Content in Graphic Narratives. *Discourse Processes*, *54*(8), 569–582. https://doi.org/10.1080/0163853X.2015.1136870

Magliano, J. P., Larson, A. M., Higgs, K., & Loschky, L. C. (2016). The relative roles of visuospatial and linguistic working memory systems in generating inferences during visual narrative comprehension. *Memory & Cognition*, *44*(2), 207–219. https://doi.org/10.3758/s13421-015-0558-7

Mammarella, I. C., Caviola, S., Giofrè, D., & Szűcs, D. (2018). The underlying structure of visuospatial working memory in children with mathematical learning disability. *British Journal of Developmental Psychology*, *36*(2), 220–235. https://doi.org/10.1111/bjdp.12202

Markostamou, I., & Coventry, K. (2021). Memory for route and survey descriptions across the adult lifespan: The role of verbal and visuospatial working memory resources. *Journal of Environmental Psychology*, *78*, 101712. https://doi.org/10.1016/j.jenvp.2021.101712

Marks, D. F. (1973). Visual Imagery Differences in the Recall of Pictures. *British Journal of Psychology*, *64*(1), 17–24. https://doi.org/10.1111/j.2044-8295.1973.tb01322.x

Marre, Q., Huet, N., & Labeye, E. (2021). Embodied mental imagery improves memory. *Quarterly Journal of Experimental Psychology*, *74*(8), 1396–1405. https://doi.org/10.1177/17470218211009227

May, J., Andrade, J., Panabokke, N., & Kavanagh, D. (2004). Images of desire: Cognitive models of craving. *Memory*, *12*(4), 447–461. https://doi.org/10.1080/09658210444000061

Mayer, M. (1967). *A boy, a dog and a frog*. Dial Press.

McCloud, S. (1993). *Understanding Comics: The Invisible Art*. HarperCollins.

McKelvie, S. J. (1994). The Vividness of Visual Imagery Questionnaire as a predictor of facial recognition memory performance. *British Journal of Psychology*, *85*(1), 93–104. https://doi.org/10.1111/j.2044-8295.1994.tb02510.x

McNamara, D. S., & Magliano, J. (2009). Chapter 9 Toward a Comprehensive Model of Comprehension. In *Psychology of Learning and Motivation* (Vol. 51, pp. 297–384). Academic Press. https://doi.org/10.1016/S0079-7421(09)51009-2

Meneghetti, C., Gyselinck, V., Pazzaglia, F., & De Beni, R. (2009). Individual differences in spatial text processing: High spatial ability can compensate for spatial working memory interference. *Learning and Individual Differences*, *19*(4), 577–589. https://doi.org/10.1016/j.lindif.2009.07.007

Miller, E. K., & Cohen, J. D. (2001). An Integrative Theory of Prefrontal Cortex Function. *Annual Review of Neuroscience*, *24*(Volume 24, 2001), 167–202. https://doi.org/10.1146/annurev.neuro.24.1.167

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

Milner, B. (1971). Interhemispheric Differences in the Localization of Psychological Processes in Man. *British Medical Bulletin*, *27*(3), 272–277. https://doi.org/10.1093/oxfordjournals.bmb.a070866

Miyake, A., & Friedman, N. P. (2012). The Nature and Organization of Individual Differences in Executive Functions: Four General Conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14. https://doi.org/10.1177/0963721411429458

Morey, C. C., & Cowan, N. (2004). When visual and verbal memories compete: Evidence of cross-domain limits in working memory. *Psychonomic Bulletin & Review*, *11*, 296–301. https://doi.org/10.3758/BF03196573

Morey, C. C., & Cowan, N. (2005). When do visual and verbal memories conflict? The importance of working-memory load and retrieval. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *31 4*, 703–713. https://doi.org/10.1037/0278-7393.31.4.703

Morrow, D. G., Greenspan, S. L., & Bower, G. H. (1987). Accessibility and situation models in narrative comprehension. *Journal of Memory and Language*, *26*(2), 165–187. https://doi.org/10.1016/0749-596X(87)90122-7

Oberauer, K. (2009). Chapter 2 Design for a Working Memory. In *Psychology of Learning and Motivation* (Vol. 51, pp. 45–100). Elsevier. https://doi.org/10.1016/S0079-7421(09)51002-X

Oberauer, K., Farrell, S., Jarrold, C., & Lewandowsky, S. (2016). What limits working memory capacity? *Psychological Bulletin*, *142*(7), Article 7. https://doi.org/10.1037/bul0000046

Oberauer, K., Süß, H.-M., Wilhelm, O., & Sander, N. (2008). Individual Differences in Working Memory Capacity and Reasoning Ability. In *Variation in working memory* (pp. 49–75). https://doi.org/10.1093/acprof:oso/9780195168648.003.0003

Olsthoorn, N. M., Andringa, S., & Hulstijn, J. H. (2014). Visual and auditory digit-span performance in native and non-native speakers. *International Journal of Bilingualism*, *18*(6), 663–673. https://doi.org/10.1177/1367006912466314

Osaka, N., Otsuka, Y., Hirose, N., Ikeda, T., Mima, T., Fukuyama, H., & Osaka, M. (2007). Transcranial magnetic stimulation (TMS) applied to left dorsolateral prefrontal cortex disrupts verbal working memory performance in humans. *Neuroscience Letters*, *418*(3), 232–235. https://doi.org/10.1016/j.neulet.2007.01.087

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*(6), 785–806. https://doi.org/10.1016/0749-596X(92)90039-Z

Papagno, C., & Vallar, G. (1995). Verbal Short-term Memory and Vocabulary Learning in Polyglots. *The Quarterly Journal of Experimental Psychology Section A*, *48*(1), 98–107. https://doi.org/10.1080/14640749508401378

Papenmeier, F., Boss, A., & Mahlke, A.-K. (2019). Action goal changes caused by agents and patients both induce global updating of event models. *Journal of Experimental Psychology: Learning Memory and Cognition*, *45*(8), 1441–1454. https://doi.org/10.1037/xlm0000651

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, *116*(2), 220–244. https://doi.org/10.1037/0033-2909.116.2.220

Pazzaglia, F., Meneghetti, C., De Beni, R., & Gyselinck, V. (2010). Working memory components in survey and route spatial text processing. *Cognitive Processing*, *11*(4), Article 4. https://doi.org/10.1007/s10339-009-0353-0

Pearson, D. G., Ball, K., & Smith, D. T. (2014). Oculomotor preparation as a rehearsal mechanism in spatial working memory. *Cognition*, *132*(3), 416–428. https://doi.org/10.1016/j.cognition.2014.05.006

Pearson, J., Clifford, C. W. G., & Tong, F. (2008). The Functional Impact of Mental Imagery on Conscious Perception. *Current Biology*, *18*(13), 982–986. https://doi.org/10.1016/j.cub.2008.05.048

Pearson, J., & Kosslyn, S. M. (2015). The heterogeneity of mental representation: Ending the imagery debate. *Proceedings of the National Academy of Sciences*, *112*(33), 10089–10092. https://doi.org/10.1073/pnas.1504933112

Pearson, J., Rademaker, R. L., & Tong, F. (2011). Evaluating the Mind's Eye: The Metacognition of Visual Imagery. *Psychological Science*, *22*(12), 1535–1542. https://doi.org/10.1177/0956797611417134

Pettijohn, K. A., & Radvansky, G. A. (2016). Narrative event boundaries, reading times, and expectation. *Memory and Cognition*, *44*(7), 1064–1075.https://doi.org/10.3758/s13421-016-0619-6

Pickering, S. J. (2001). The development of visuo-spatial working memory. *Memory*, *9*(4–6), 423–432. https://doi.org/10.1080/09658210143000182

Picucci, L., Gyselinck, V., Piolino, P., Nicolas, S., & Bosco, A. (2013). Spatial Mental Models: The Interaction of Presentation Format, Task Requirements and Availability of Working Memory Components. *Applied Cognitive Psychology*, *27*(3), 314–327. https://doi.org/10.1002/acp.2909

Postle, B. R., Idzikowski, C., Sala, S. D., Logie, R. H., & Baddeley, A. D. (2006). The selective disruption of spatial working memory by eye movements. *Quarterly Journal of Experimental Psychology*, *59*(1), 100–120. https://doi.org/10.1080/17470210500151410

Potter, M. C., Wyble, B., Hagmann, C. E., & McCourt, E. S. (2014). Detecting meaning in RSVP at 13 ms per picture. *Attention, Perception, & Psychophysics*, *76*(2), 270–279. https://doi.org/10.3758/s13414-013-0605-z

Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin*, *80*(1), Article 1. https://doi.org/10.1037/h0034650

Pylyshyn, Z. W. (2002). Mental imagery: In search of a theory. *Behavioral and Brain Sciences*, *25*(2), 157–182. https://doi.org/10.1017/S0140525X02000043

R Core Team. (2022). *R: A Language and Environment for Statistical Computing* (Version 4.2.2) [R]. R Foundation for Statistical Computing. https://www.R-project.org/

Radvansky, G. A., & Copeland, D. E. (2001). Working memory and situation model updating. *Memory & Cognition*, *29*(8), Article 8. https://doi.org/10.3758/BF03206375

Radvansky, G. A., & Copeland, D. E. (2006). Memory retrieval and interference: Working memory issues. *Journal of Memory and Language*, *55*(1), 33–46. https://doi.org/10.1016/j.jml.2006.02.001

Radvansky, G. A., & Copeland, D. E. (2010). Reading times and the detection of event shift processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(1), 210–216. https://doi.org/10.1037/a0017258

Radvansky, G. A., & Zacks, J. M. (2011). Event perception. *WIREs Cognitive Science*, *2*(6), 608–620. https://doi.org/10.1002/wcs.133

Radvansky, G., & Copeland, D. (2004). *Working memory span and situation model processing.* https://doi.org/10.2307/4149022

Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences*, *20*(2), 110–122. https://doi.org/10.1016/j.lindif.2009.10.005

Repovš, G., & Baddeley, A. (2006). The multi-component model of working memory: Explorations in experimental cognitive psychology. *Neuroscience*, *139*, 5–21. https://doi.org/10.1016/j.neuroscience.2005.12.061

Reynolds, J. R., Zacks, J. M., & Braver, T. S. (2007). A Computational Model of Event Segmentation From Perceptual Prediction. *Cognitive Science*, *31*(4), 613–643. https://doi.org/10.1080/15326900701399913

Ricker, T. J., Cowan, N., & Morey, C. C. (2010). Visual working memory is disrupted by covert verbal retrieval. *Psychonomic Bulletin & Review*, *17*(4), 516–521. https://doi.org/10.3758/PBR.17.4.516

Rinck, M., Hähnel, A., Bower, G. H., & Glowalla, U. (1997). The metrics of spatial situation models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(3), 622–637. https://doi.org/10.1037/0278-7393.23.3.622

Rinck, M., & Weber, U. (2003a). Who when where: An experimental test of the event-indexing model. *Memory & Cognition*, *31*(8), 1284–1292. https://doi.org/10.3758/BF03195811

Robinson, S. J., & Brewer, G. (2016). Performance on the traditional and the touch screen, tablet versions of the Corsi Block and the Tower of Hanoi tasks. *Computers in Human Behavior*, *60*, 29–34. https://doi.org/10.1016/j.chb.2016.02.047

Sadoski, M., & Paivio, A. (2004). A dual coding theoretical model of reading. In R. R. Ruddell & N. J. Unrau (Eds.), *Theoretical models and processes of reading* (5th ed., pp. 1329–1362). Newark: International Reading Association.

Salway, A. F. S., & Logie, R. H. (1995). Visuospatial working memory, movement control and executive demands. *British Journal of Psychology*, *86*(2), 253–269. https://doi.org/10.1111/j.2044-8295.1995.tb02560.x

Sargent, J. Q., Zacks, J. M., Hambrick, D. Z., Zacks, R. T., Kurby, C. A., Bailey, H. R., Eisenberg, M. L., & Beck, T. M. (2013). Event segmentation ability uniquely predicts event memory. *Cognition*, *129*(2), 241–255. https://doi.org/10.1016/j.cognition.2013.07.002

Schulz, C. (2004). *The complete Peanuts: 1950-1952*. Fantagraphics Books.

Seigneuric, A., Ehrlich, M.-F., Oakhill, J. V., & Yuill, N. M. (2000). Working memory resources and children's reading comprehension. *Reading and Writing*, *13*(1), 81–103. https://doi.org/10.1023/A:1008088230941

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General*, *125*(1), 4–27. https://doi.org/10.1037/0096-3445.125.1.4

Sims, V. K., & Hegarty, M. (1997). Mental animation in the visuospatial sketchpad: Evidence from dual-task studies. *Memory & Cognition*, *25*(3), 321–332. https://doi.org/10.3758/BF03211288

Slotnick, S. D., Thompson, W. L., & Kosslyn, S. M. (2012). Visual memory and visual mental imagery recruit common control and sensory regions of the brain. *Cognitive Neuroscience*, *3*(1), 14–20. https://doi.org/10.1080/17588928.2011.578210

Smith, E. E., & Jonides, J. (1999). Storage and Executive Processes in the Frontal Lobes. *Science*, *283*(5408), 1657–1661. https://doi.org/10.1126/science.283.5408.1657

Smith, E. E., Jonides, J., & Koeppe, R. A. (1996). Dissociating Verbal and Spatial Working Memory Using PET. *Cerebral Cortex*, *6*(1), 11–20. https://doi.org/10.1093/cercor/6.1.11

Speckens, A. E. M., Hackmann, A., Ehlers, A., & Cuthbert, B. (2007). Imagery special issue: Intrusive images and memories of earlier adverse events in patients with obsessive compulsive disorder. *Journal of Behavior Therapy and Experimental Psychiatry*, *38*(4), 411–422. https://doi.org/10.1016/j.jbtep.2007.09.004

Speed, L. J., Eekhof, L. S., & Mak, M. (2024). The role of visual imagery in story reading: Evidence from aphantasia. *Consciousness and Cognition*, *118*, 103645. https://doi.org/10.1016/j.concog.2024.103645

Suggate, S., & Lenhard, W. (2022). Mental imagery skill predicts adults' reading performance. *Learning and Instruction*, *80*, 101633. https://doi.org/10.1016/j.learninstruc.2022.101633

Swallow, K. M., Zacks, J. M., & Abrams, R. A. (2009). Event boundaries in perception affect memory encoding and updating. *Journal of Experimental Psychology: General*, *138*(2), 236–257. https://doi.org/10.1037/a0015631

Swanson, H. L., Orosco, M. J., & Kudo, M. (2017). Does Growth in the Executive System of Working Memory Underlie Growth in Literacy for Bilingual Children With and Without Reading Disabilities? *Journal of Learning Disabilities*, *50*(4), 386–407.https://doi.org/10.1177/0022219415618499

Swets, B., & Kurby, C. A. (2016). Eye Movements Reveal the Influence of Event Structure on Reading Behavior. *Cognitive Science*, *40*(2), 466–480. https://doi.org/10.1111/cogs.12240

Tabachnick, B. G., & Fidell, L. S. (2018). *Using Multivariate Statistics*. Pearson.

Tabi, Y. A., Maio, M. R., Attaallah, B., Dickson, S., Drew, D., Idris, M. I., Kienast, A., Klar, V., Nobis, L., Plant, O., Saleh, Y., Sandhu, T. R., Slavkova, E., Toniolo, S., Zokaei, N., Manohar, S. G., & Husain, M. (2022). Vividness of visual imagery questionnaire scores and their relationship to visual short-term memory performance. *Cortex*, *146*, 186–199. https://doi.org/10.1016/j.cortex.2021.10.011

Therriault, D. J., Rinck, M., & Zwaan, R. A. (2006). Assessing the influence of dimensional focus during situation model construction. *Memory & Cognition*, *34*(1), 78–89. https://doi.org/10.3758/BF03193388

Thompson, A. N., & Radvansky, G. A. (2016). Event boundaries and anaphoric reference. *Psychonomic Bulletin and Review*, *23*(3), 849–856. https://doi.org/10.3758/s13423-015-0961-x

Tong, F. (2013). Imagery and visual working memory: One and the same? *Trends in Cognitive Sciences*, *17*(10), 489–490. https://doi.org/10.1016/j.tics.2013.08.005

Van Dijk, T., & Kintsch, W. (1983). *Strategies of Discourse Comprehension*. Academic Press.

Vandierendonck, A., Kemps, E., Fastame, M. C., & Szmalec, A. (2004). Working memory components of the Corsi blocks task. *British Journal of Psychology*, *95*(1), 57–79. https://doi.org/10.1348/000712604322779460

Vergauwe, E., von Bastian, C. C., Kostova, R., & Morey, C. C. (2022). Storage and processing in working memory: A single, domain-general resource explains multitasking. *Journal of Experimental Psychology: General*, *151*(2), 285–301. https://doi.org/10.1037/xge0000895

Wager, T. D., & Smith, E. E. (2003). Neuroimaging studies of working memory: *Cognitive, Affective, & Behavioral Neuroscience*, *3*(4), 255–274. https://doi.org/10.3758/CABN.3.4.255

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Hecht, S. A., Barker, T. A., Burgess, S. R., Donahue, J., & Garon, T. (1997). Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: A 5-year longitudinal study. *Developmental Psychology*, *33*(3), 468–479. https://doi.org/10.1037/0012-1649.33.3.468

West, W. C., & Holcomb, P. J. (2002). Event-related potentials during discourse-level semantic integration of complex pictures. *Cognitive Brain Research*, *13*(3), 363–375. https://doi.org/10.1016/S0926-6410(01)00129-X

Wicken, M., Keogh, R., & Pearson, J. (2021). The critical role of mental imagery in human emotion: Insights from fear-based imagery and aphantasia. *Proceedings of the Royal Society B: Biological Sciences*, *288*(1946), 20210267. https://doi.org/10.1098/rspb.2021.0267

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*(2), 159–177. https://doi.org/10.1080/14639220210123806

Wilde, N. J., Strauss, E., & Tulsky, D. S. (2004). Memory Span on the Wechsler Scales. *Journal of Clinical and Experimental Neuropsychology*, *26*(4), 539–549. https://doi.org/10.1080/13803390490496605

Williams, R., & Suggate, S. P. (2024). Latent Profiles of Visual Imagery: Aphantasics, Mid-Range Imagers, and Hyperphantasics Experience Reading Differently. *Imagination, Cognition and Personality*, *44*(2), 168–194. https://doi.org/10.1177/02762366241279404

Xavier Alario, F., Segui, J., & Ferrand, L. (2000). Semantic and Associative Priming in Picture Naming. *The Quarterly Journal of Experimental Psychology Section A*, *53*(3), 741–764. https://doi.org/10.1080/713755907

Yates, T. S., Sherman, B. E., & Yousif, S. R. (2023). More than a moment: What does it mean to call something an 'event'? *Psychonomic Bulletin & Review*, *30*(6), 2067–2082. https://doi.org/10.3758/s13423-023-02311-4

Zacks, J. M., Speer, N. K., & Reynolds, J. R. (2009). Segmentation in reading and film comprehension. *Journal of Experimental Psychology: General*, *138*(2), 307–327. https://doi.org/10.1037/a0015305

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind-brain perspective. *Psychological Bulletin*, *133*(2), 273–293. https://doi.org/10.1037/0033-2909.133.2.273

Zacks, J. M., & Swallow, K. M. (2007). Event Segmentation. *Current Directions in Psychological Science*, *16*(2), 80–84. https://doi.org/10.1111/j.1467-8721.2007.00480.x

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, *127*(1), 3–21. https://doi.org/10.1037/0033-2909.127.1.3

Zbären, G. A., Meissner, S. N., Kapur, M., & Wenderoth, N. (2023). Physical inference of falling objects involves simulation of occluded trajectories in early visual areas. *Human Brain Mapping*, *44*(10), 4183–4196. https://doi.org/10.1002/hbm.26338

Zeman, A., Dewar, M., & Della Sala, S. (2015). Lives without imagery—Congenital aphantasia. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, *73*, 378–380. https://doi.org/10.1016/j.cortex.2015.05.019

Zeman, A., Milton, F., Della Sala, S., Dewar, M., Frayling, T., Gaddum, J., Hattersley, A., Heuerman-Williamson, B., Jones, K., MacKisack, M., & Winlove, C. (2020). Phantasia–The psychological significance of lifelong visual imagery vividness extremes. *Cortex*, *130*, 426–440. https://doi.org/10.1016/j.cortex.2020.04.003

Zhao, F., Fan, L., Zhang, J., Liu, Y.-E., Jiang, J., & Bing, T. (2024). The effects of visuospatial working memory on older adults' bridging inference processing in visual narrative comprehension. *Discourse Processes*. *61(8), 404-419.* https://doi.org/10.1080/0163853X.2024.2338047

Živković, M., Pellizzoni, S., Doz, E., Cuder, A., Mammarella, I., & Passolunghi, M. C. (2023). Math self-efficacy or anxiety? The role of emotional and motivational contribution in math performance. *Social Psychology of Education*, *26*(3), 579–601. https://doi.org/10.1007/s11218-023-09760-8

Zwaan, R. A. (2016). Situation models, mental simulations, and abstract concepts in discourse comprehension. *Psychonomic Bulletin & Review*, *23*(4), 1028–1034. https://doi.org/10.3758/s13423-015-0864-x

Zwaan, R. A., & Brown, C. M. (1996). The influence of language proficiency and comprehension skill on situation-model construction. *Discourse Processes*, *21*(3), 289–327. https://doi.org/10.1080/01638539609544960

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The Construction of Situation Models in Narrative Comprehension: An Event-Indexing Model. *Psychological Science*, *6*(5), 292–297. https://doi.org/10.1111/j.1467-9280.1995.tb00513.x

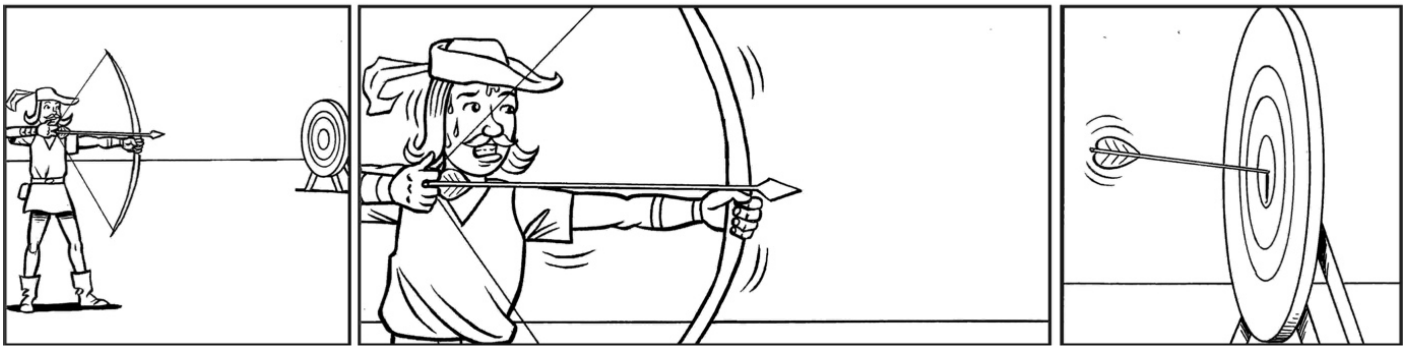
Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The Construction of Situation Models in Narrative Comprehension: An Event-Indexing Model. *Psychological Science*, *6*(5), 292–297. https://doi.org/10.1111/j.1467-9280.1995.tb00513.x

Zwaan, R. A., Radvansky, G. A., Hilliard, A. E., & Curiel, J. M. (1998). Constructing Multidimensional Situation Models During Reading. *Scientific Studies of Reading*, *2*(3), 199–220. https://doi.org/10.1207/s1532799xssr0203_2

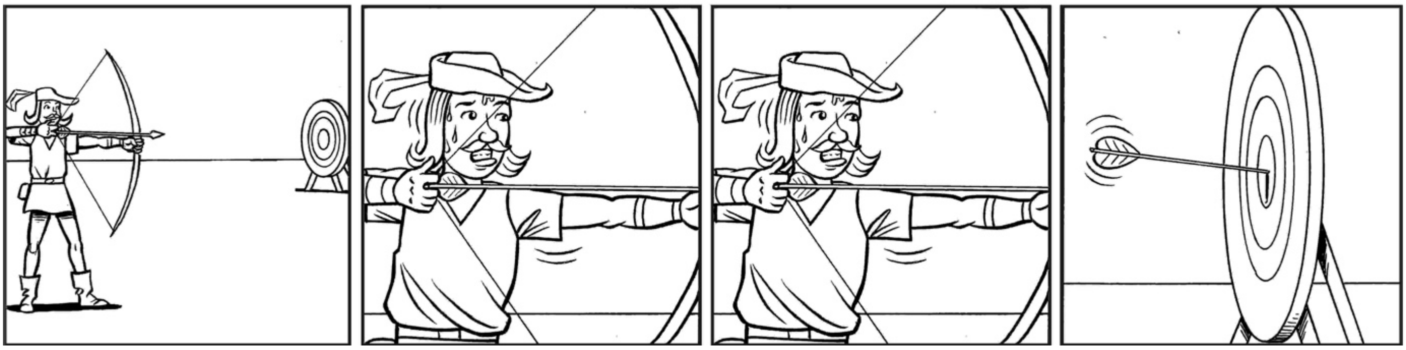
Appendix

Appendix A. Archer Comic and Variants

Appendix A.1



Appendix A.2

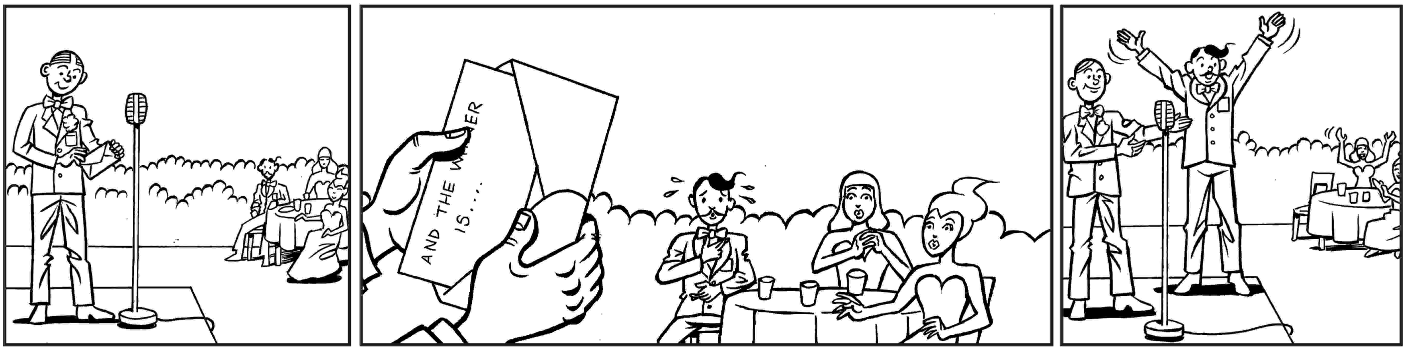


Appendix A.3

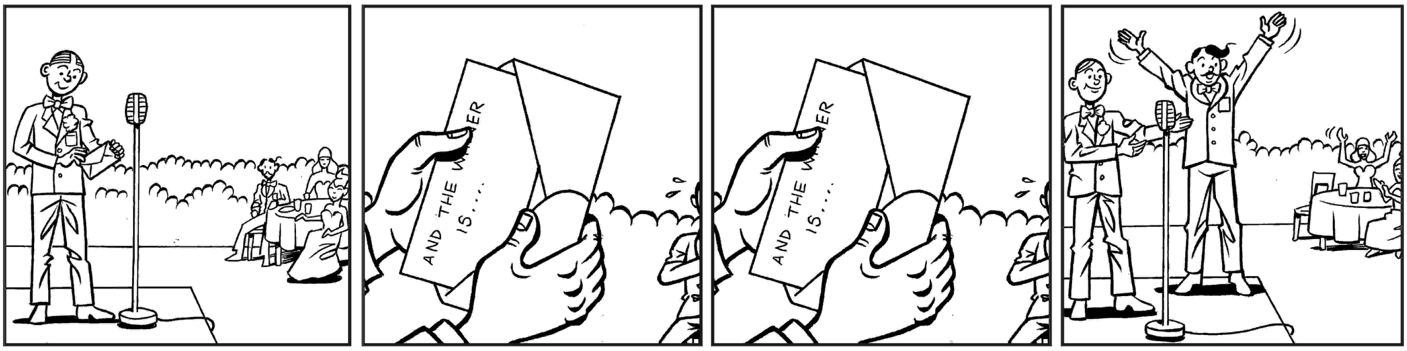
Appendix A.4

Appendix B. Award Comic and Variants

Appendix B.1

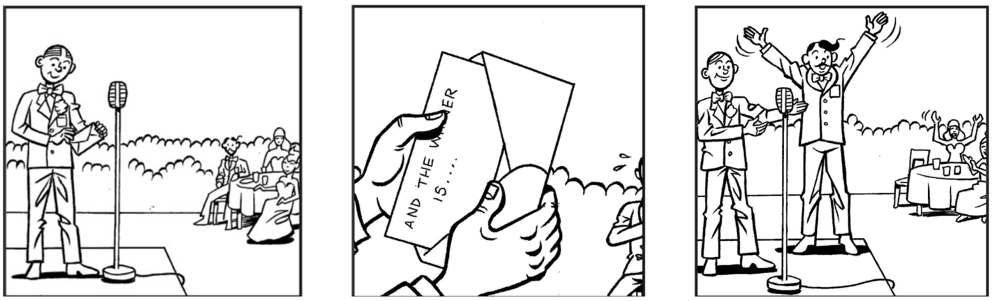


Appendix B.2



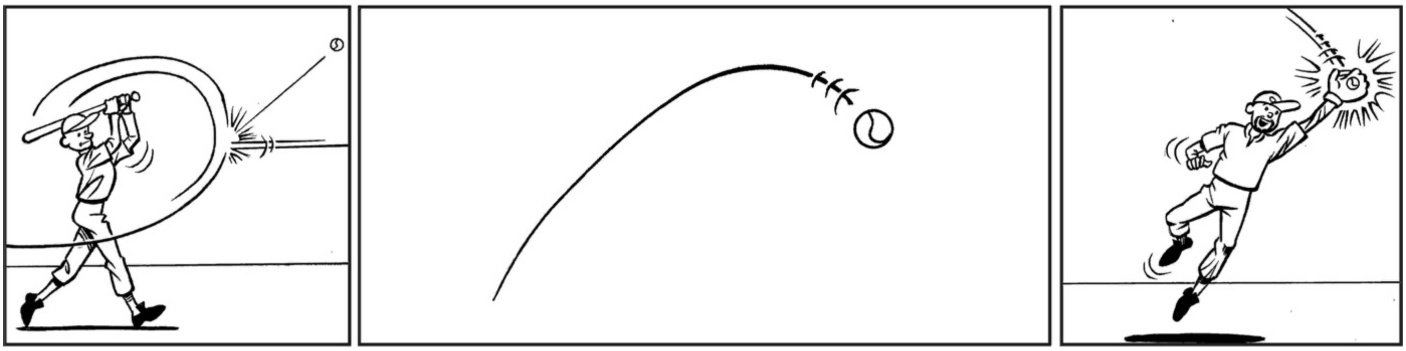
Appendix B.3

Appendix B.4

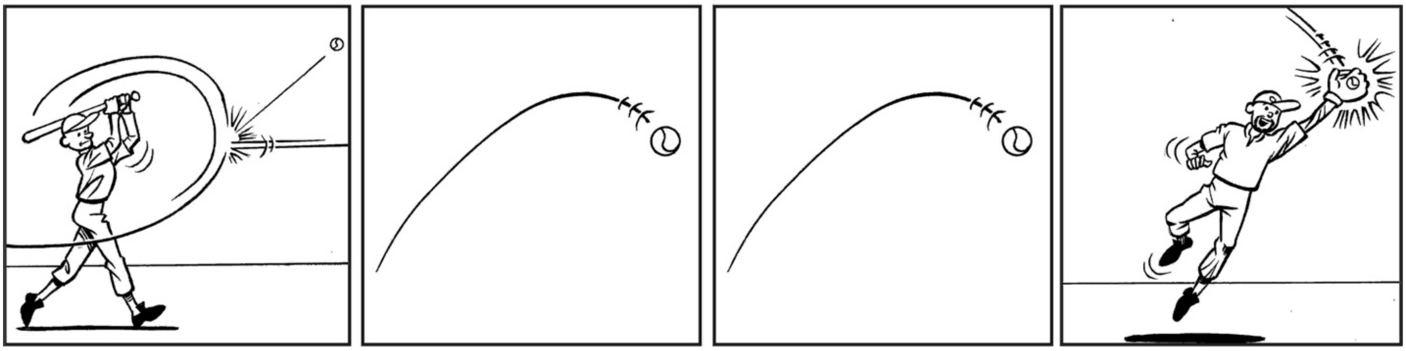


Appendix C. Baseball Comic and Variants

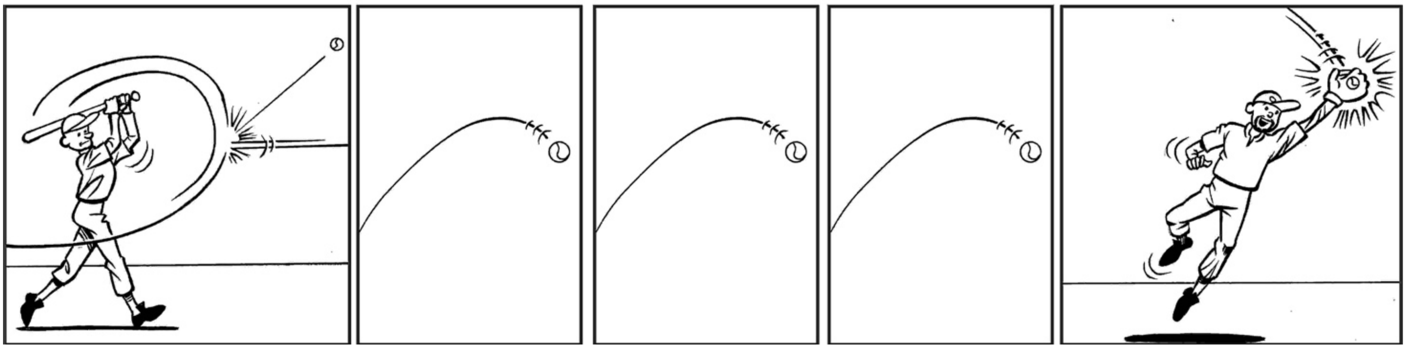
Appendix C.1



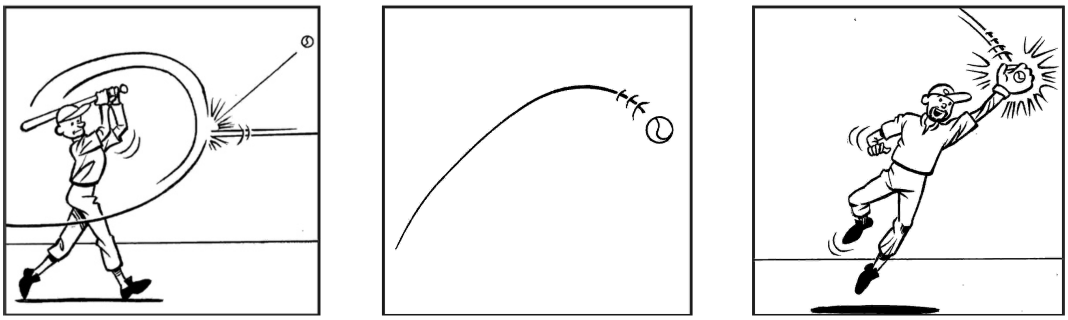
Appendix C.2



Appendix C.3

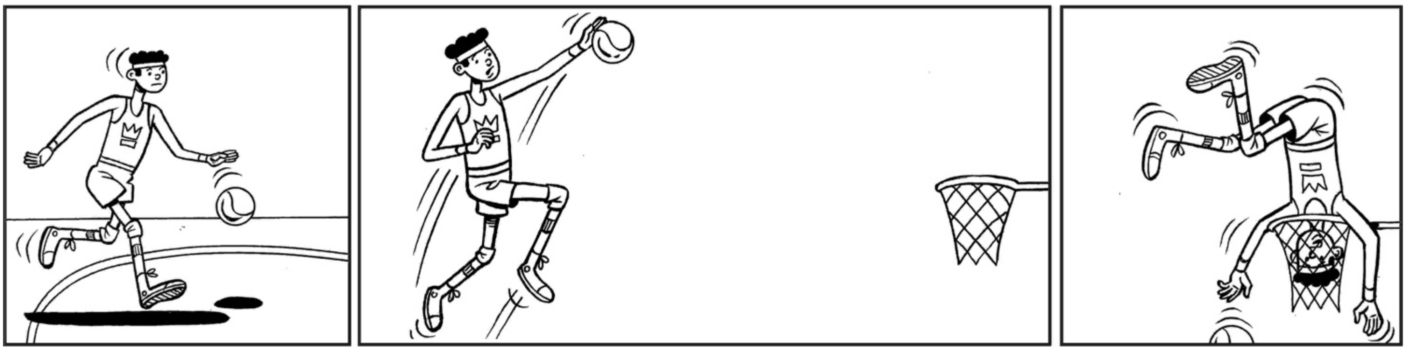


Appendix C.4

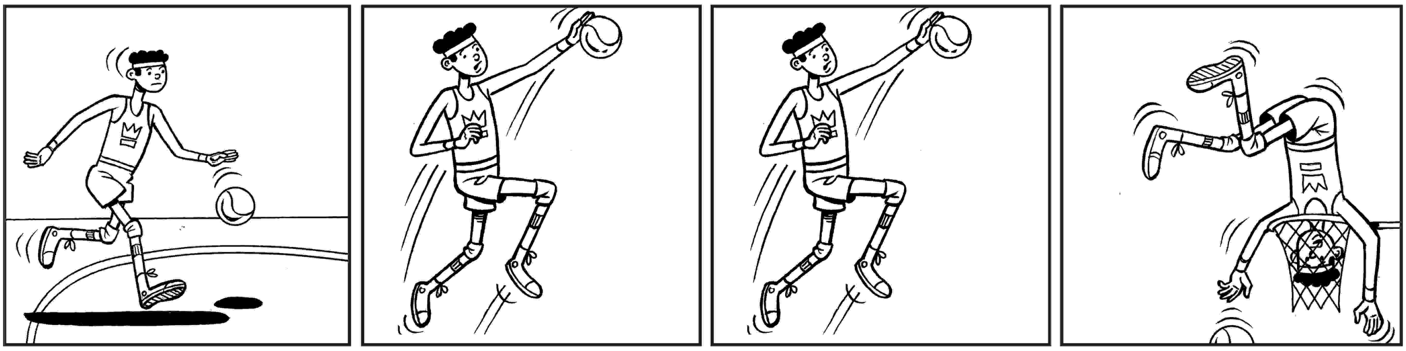


Appendix D. Basketball Comic and Variants

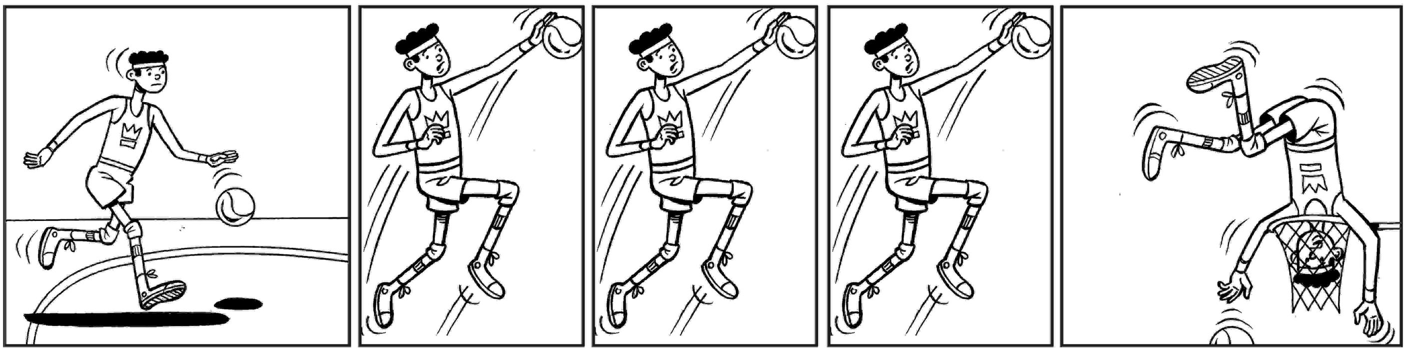
Appendix D.1



Appendix D.2



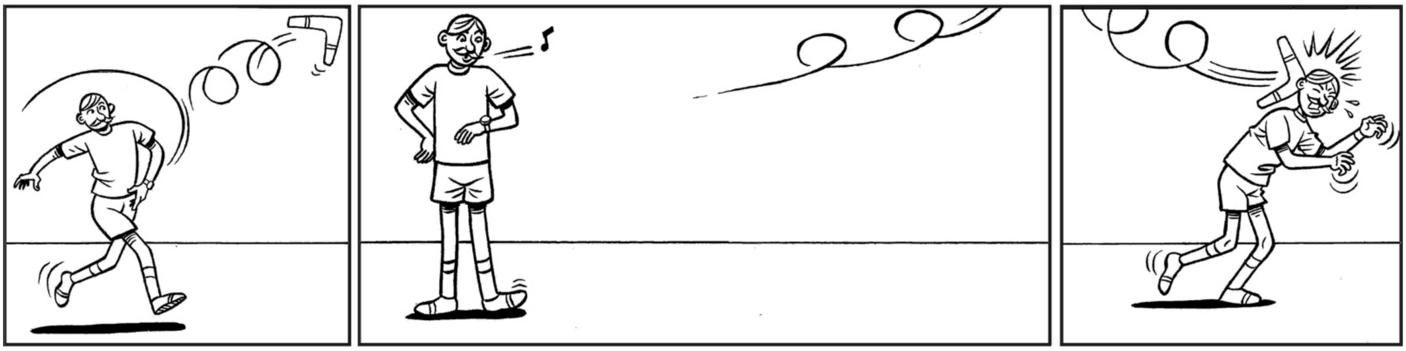
Appendix D.3



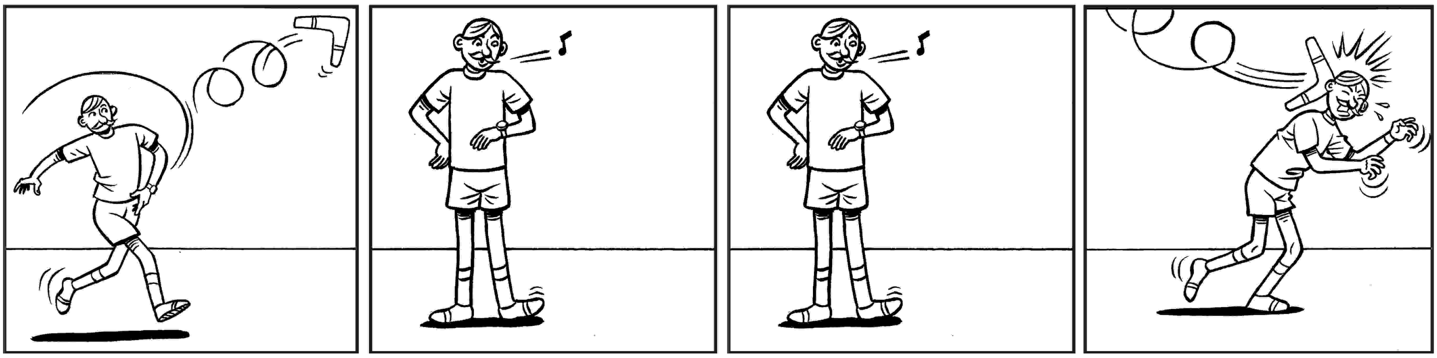
Appendix D.4

Appendix E. Boomerang Comic and Variants

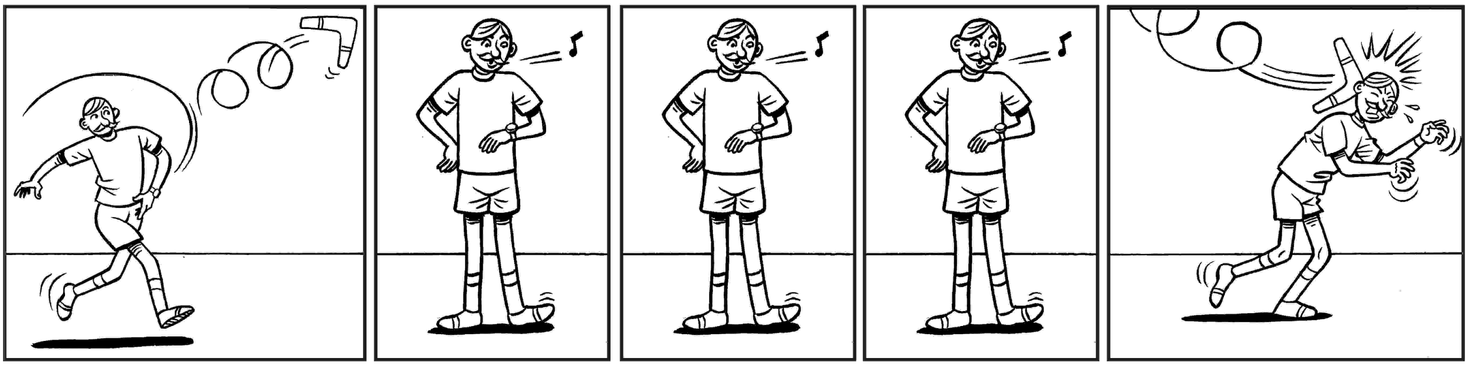
Appendix E.1



Appendix E.2



Appendix E.3



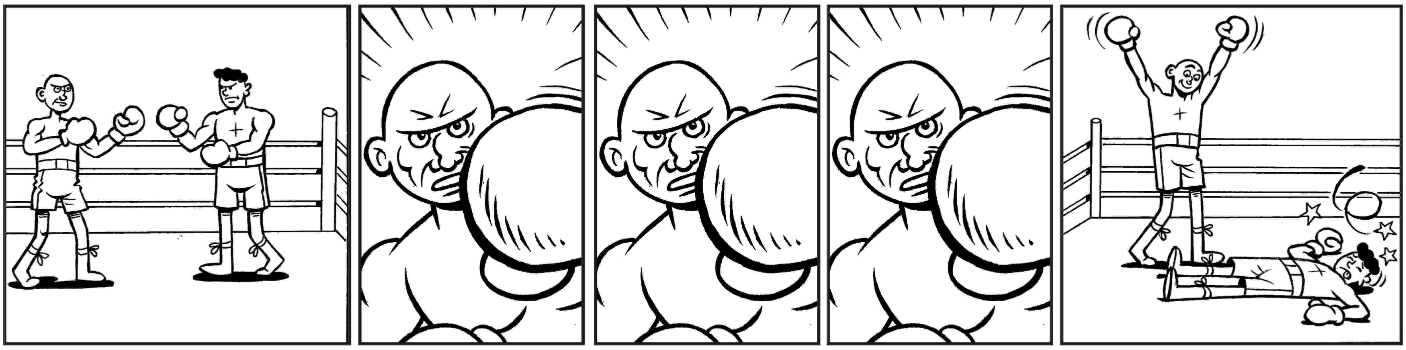
Appendix E.4

Appendix F. Boxing Comic and Variants

Appendix F.1

Appendix F.2

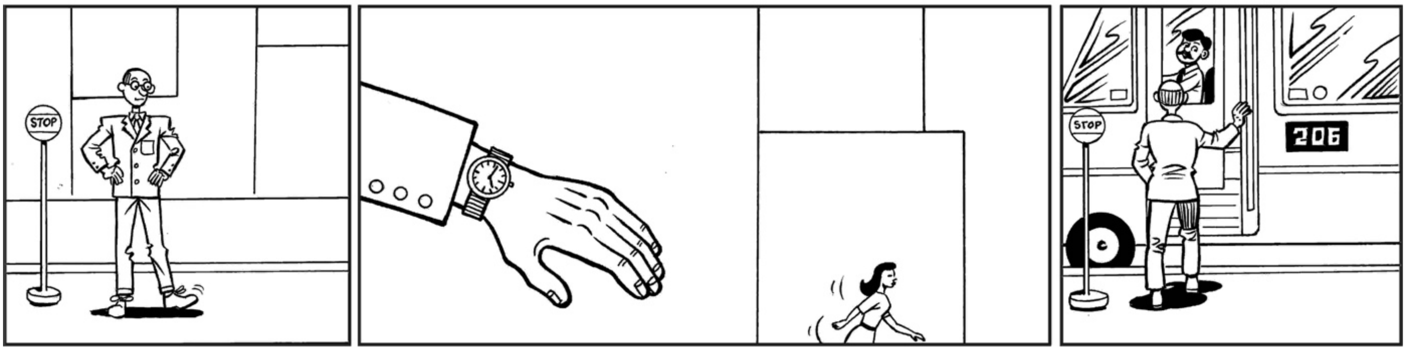
Appendix F.3



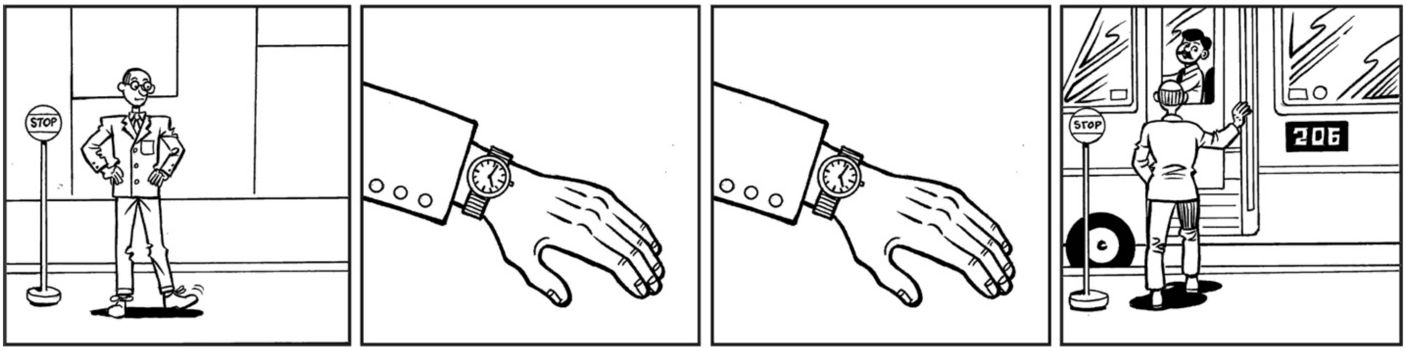
Appendix F.4

Appendix G. Bus Comic and Variants

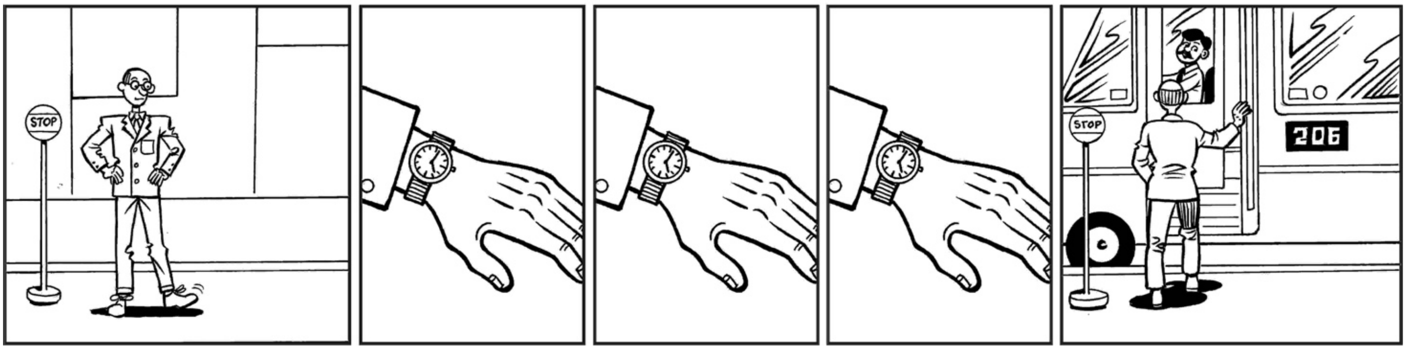
Appendix G.1



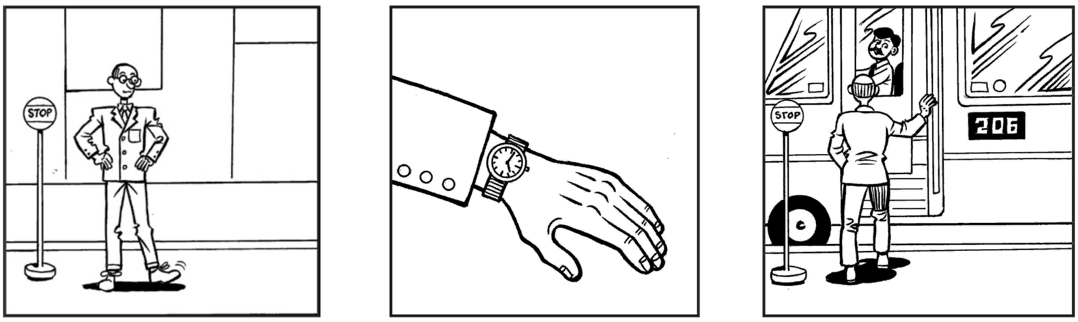
Appendix G.2



Appendix G.3

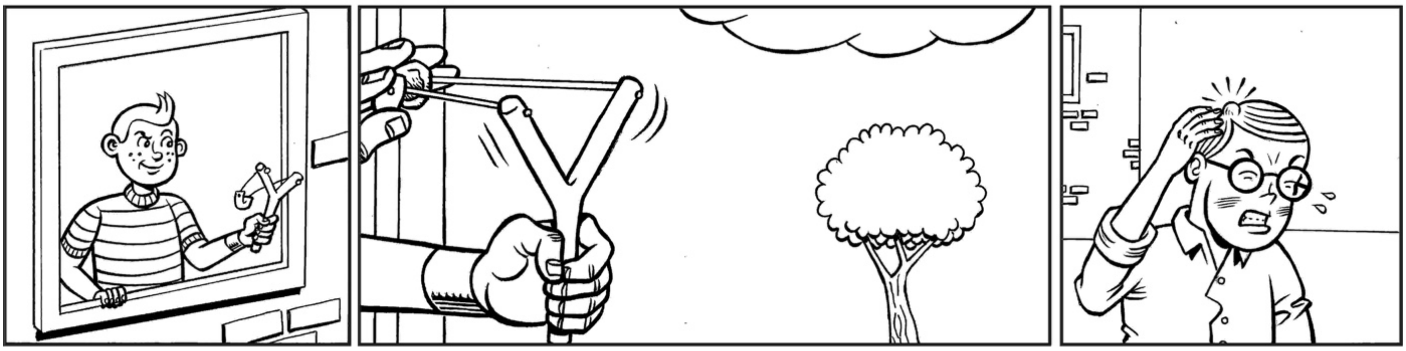


Appendix G.4

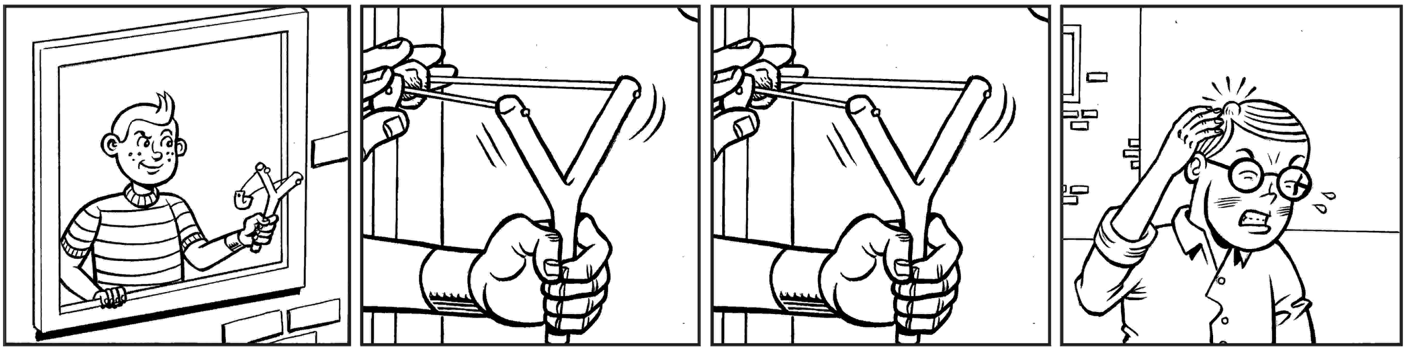


Appendix H. Catapult Comic and Variants

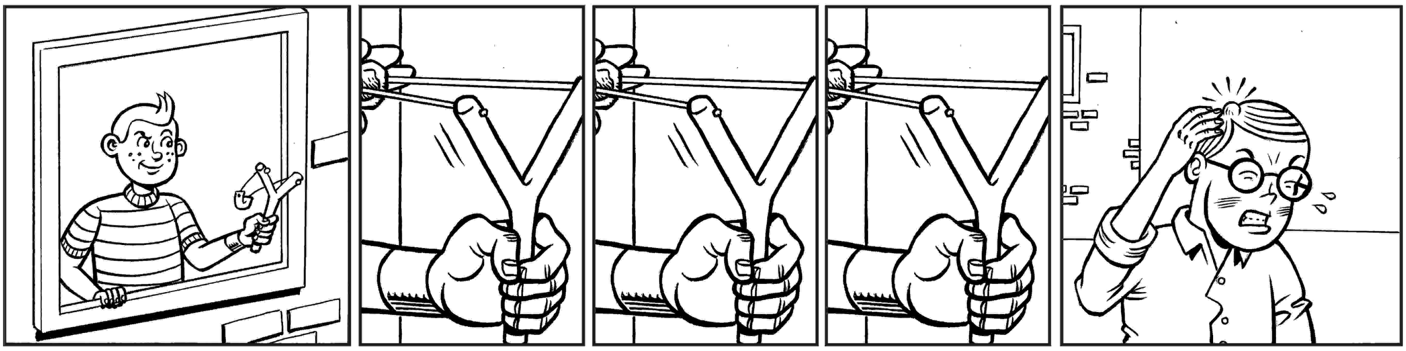
Appendix H.1



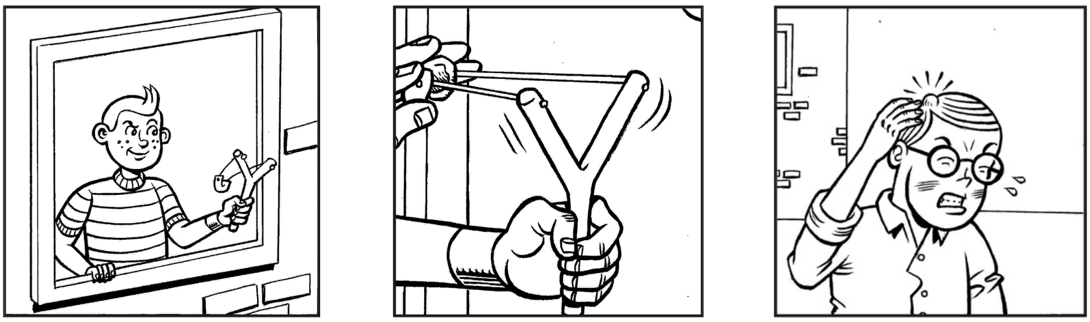
Appendix H.2



Appendix H.3

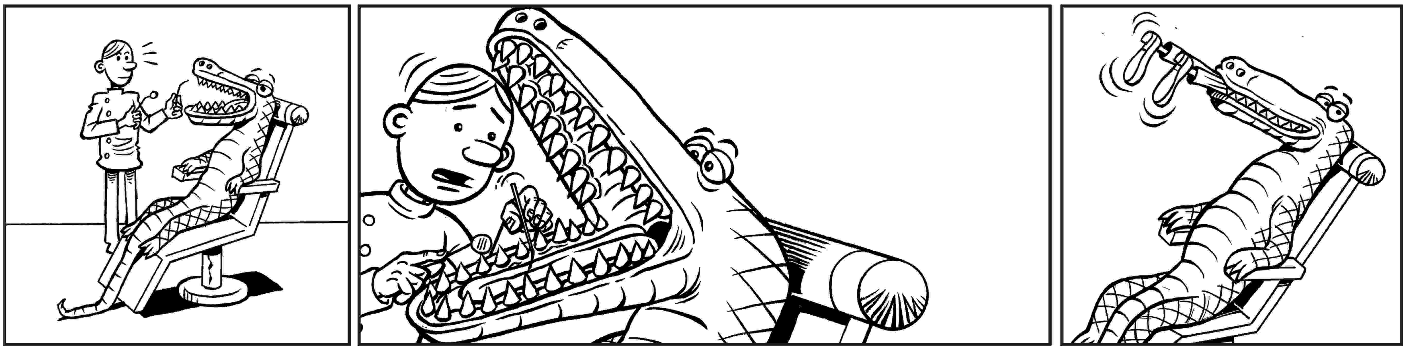


Appendix H.4

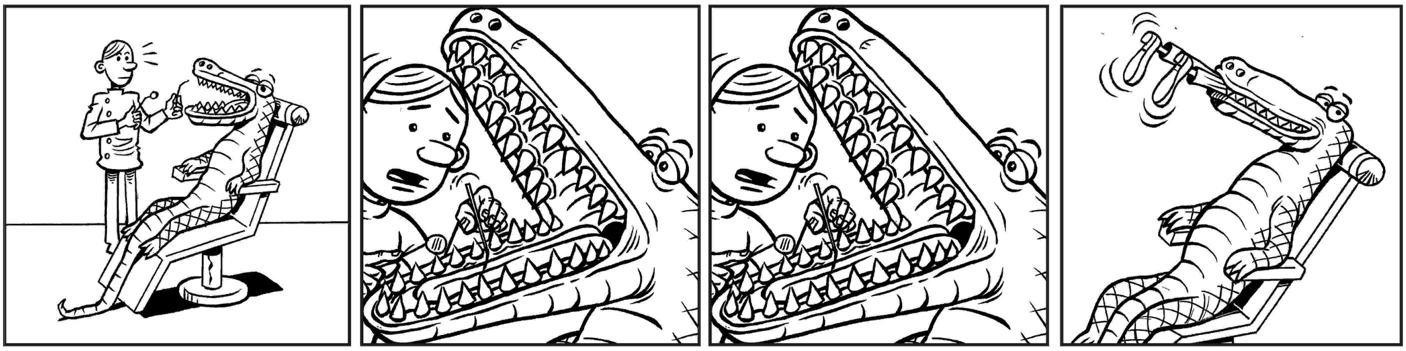


Appendix I. Croc Comic and Variants

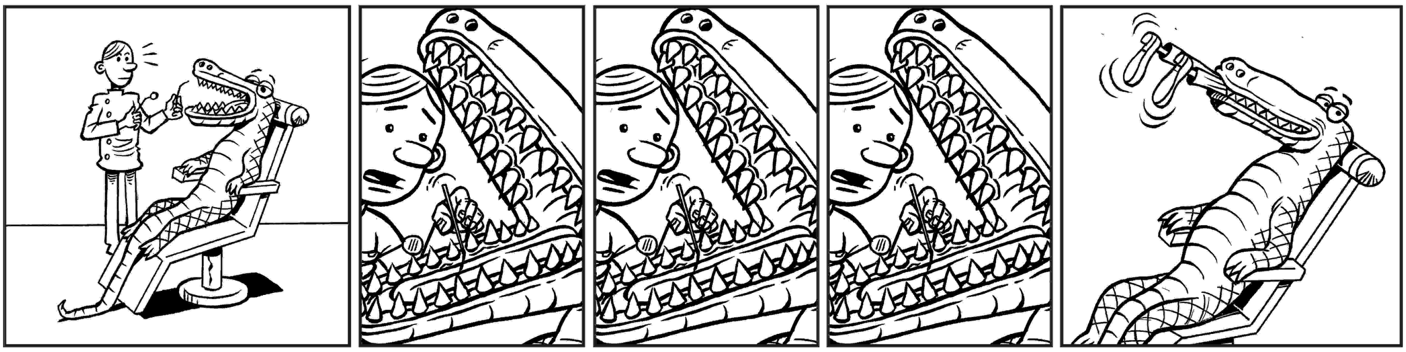
Appendix I.1



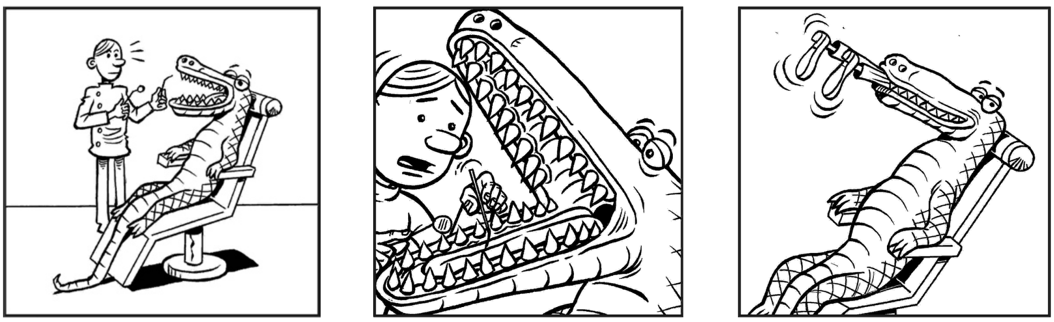
Appendix I.2



Appendix I.3

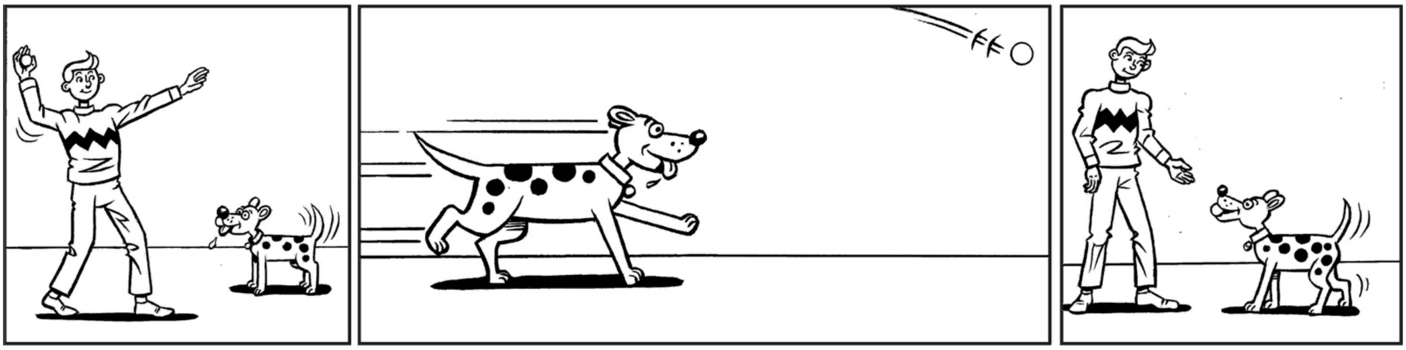


Appendix I.4

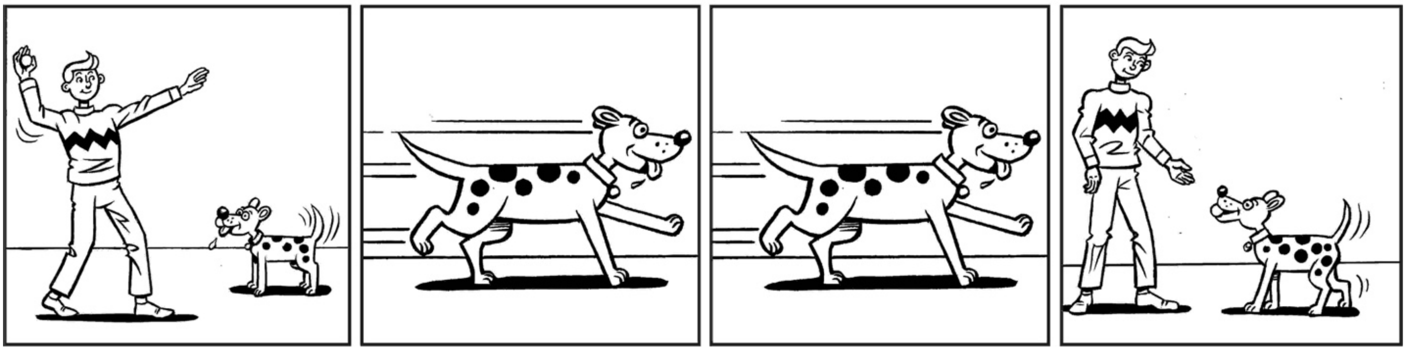


Appendix J. Dog Comic and Variants

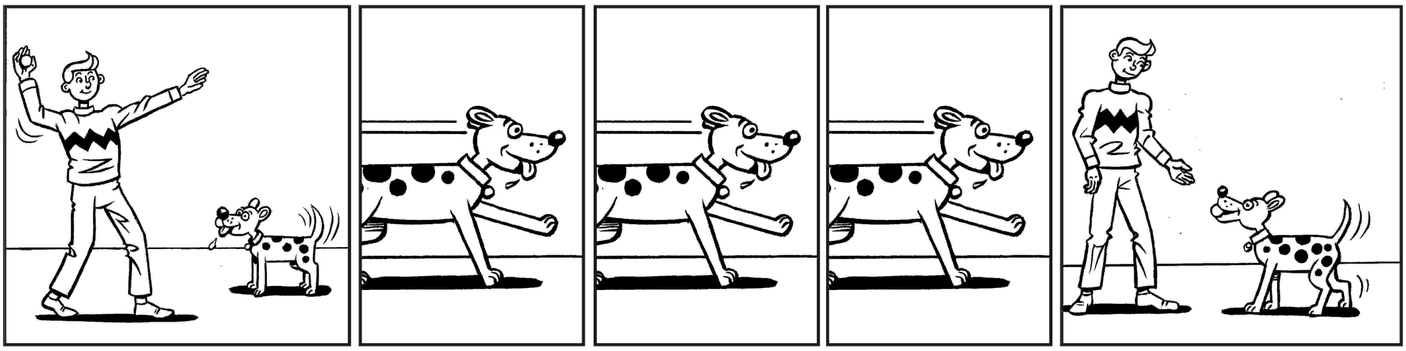
Appendix J.1



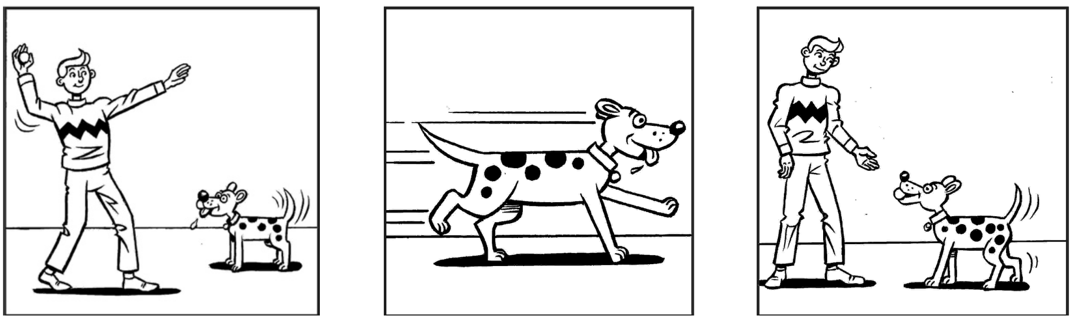
Appendix J.2



Appendix J.3



Appendix J.4

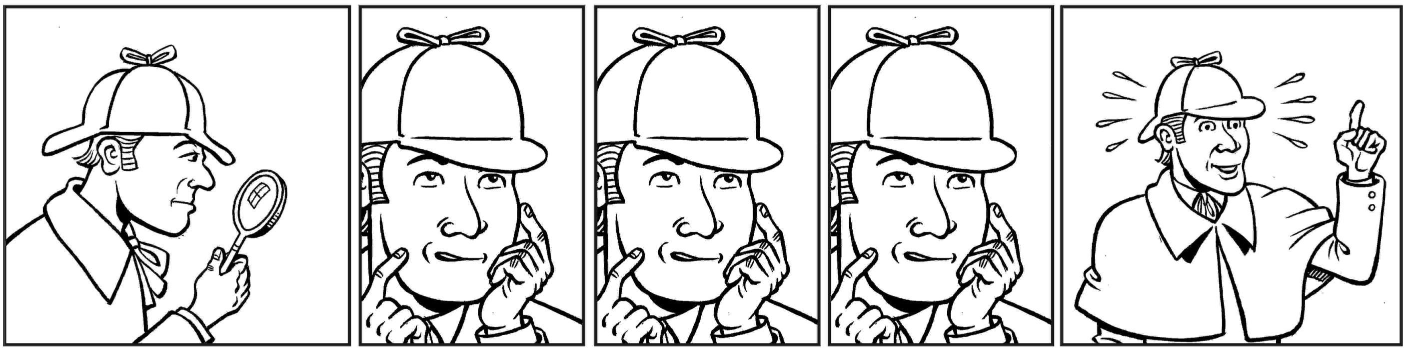


Appendix K. Sherlock Comic and Variants

Appendix K.1

Appendix K.2

Appendix K.3



Appendix K.4

Appendix L. Juggler Comic and Variants

Appendix L.1

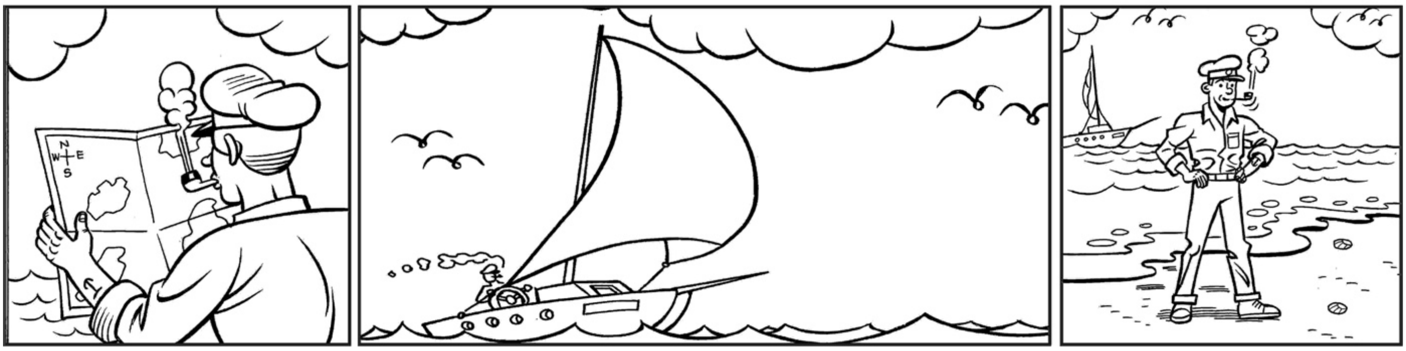
Appendix L.2

Appendix L.3

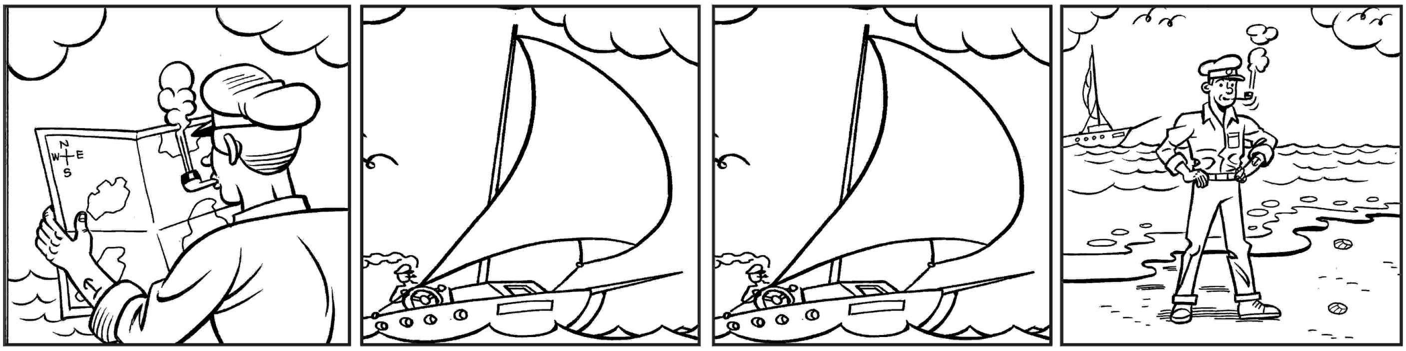
Appendix L.4

Appendix M. Sailor Comic and Variants

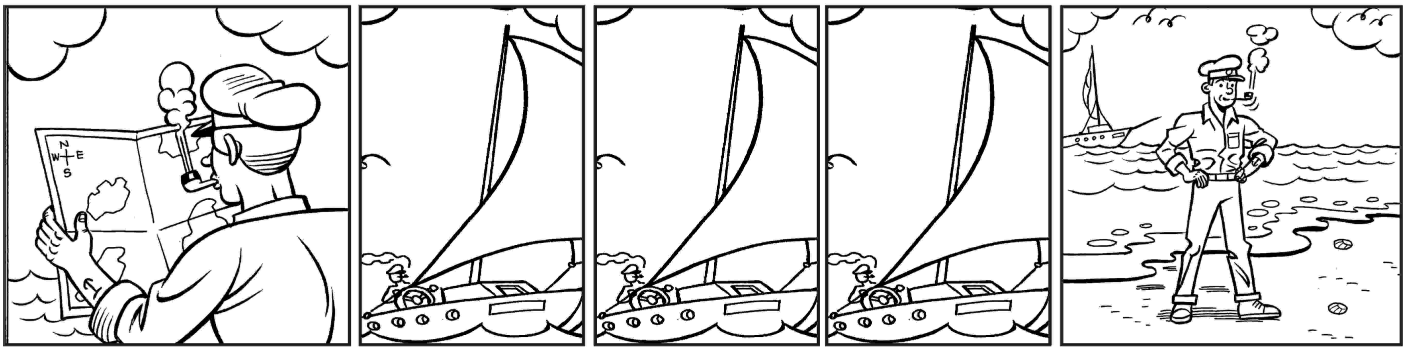
Appendix M.1



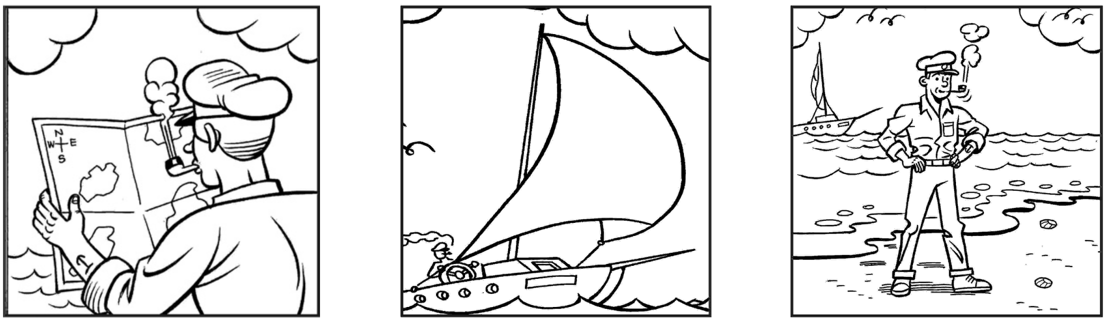
Appendix M.2



Appendix M.3

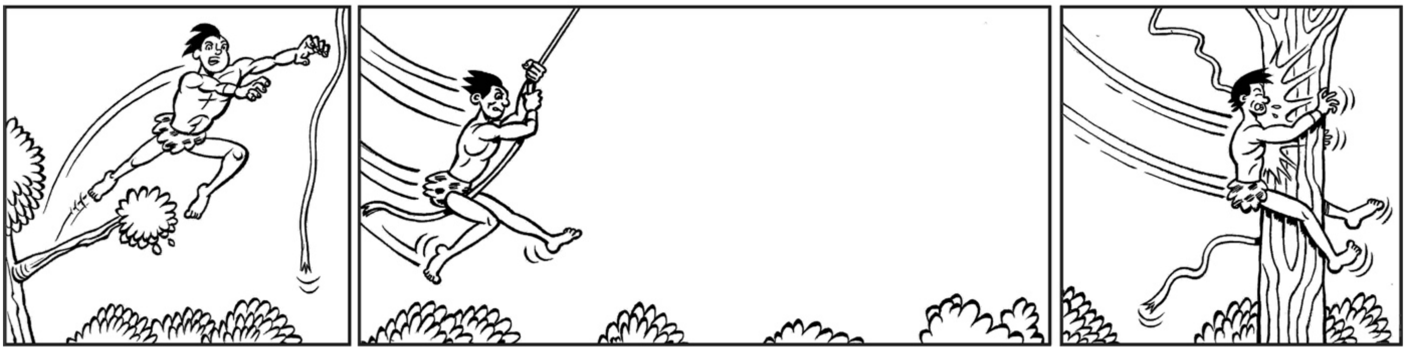


Appendix M.4

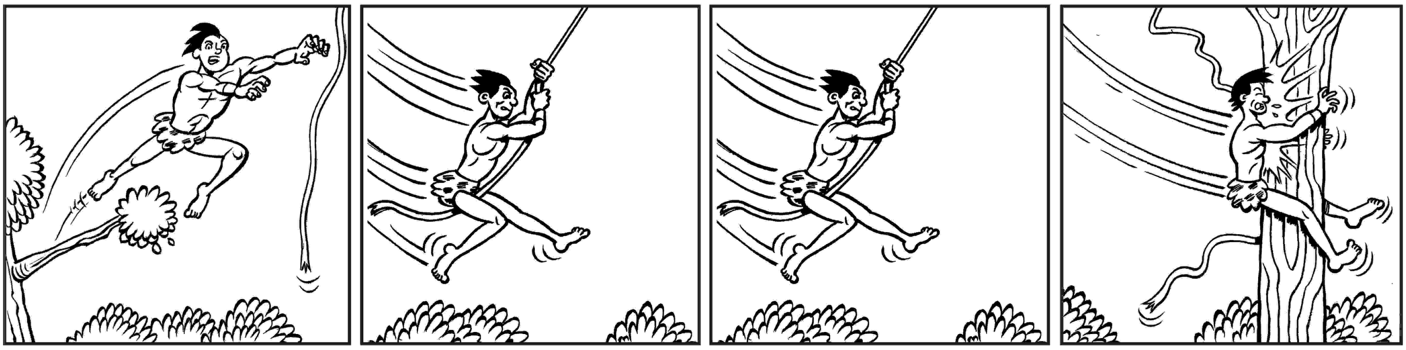


Appendix N. Tarzan Comic and Variants

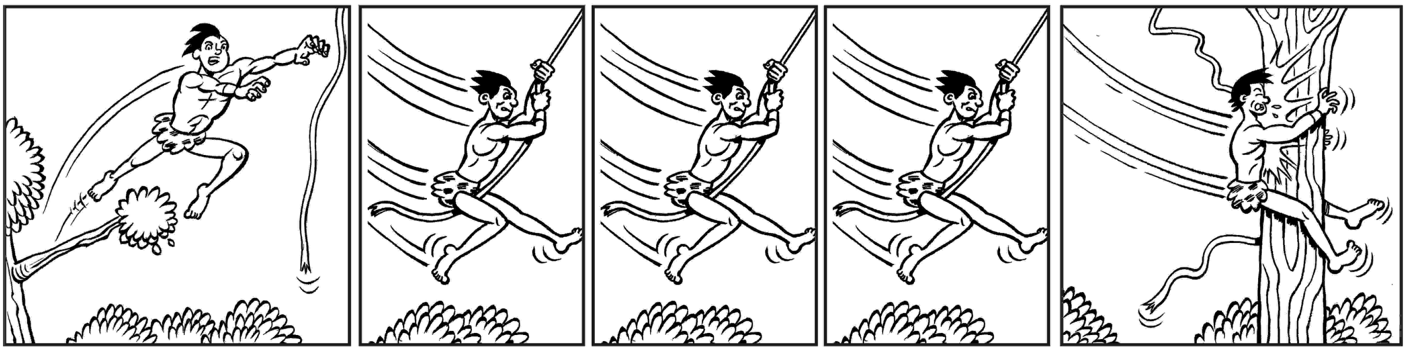
Appendix N.1



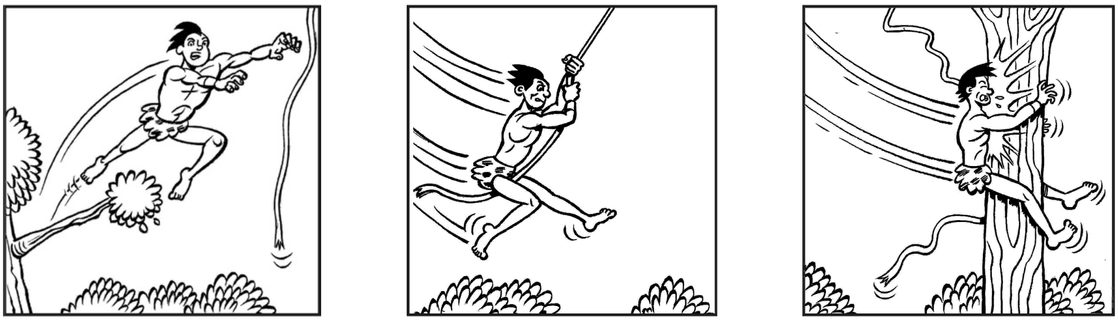
Appendix N.2



Appendix N.3



Appendix N.4



Appendix O. VVIQ Questions – Marks (1973)

In answering items 1 to 4, think of a relative or friend who you see frequently and consider the image that comes your mind's eye.

1. Picture the exact contours of their face, head, shoulders and body.
2. Characteristic poses of their head and body language etc.
3. Their appearance as they move, the way they move, the length of their step, etc.
4. The kind of clothes they wear, the different colours or styles.

In answering item 5 to 8, think of a horizon. Read the following statements and rate the vividness of your mental pictures.

5. The sun is rising above the horizon into a hazy sky.
6. The sky clears and surrounds the sun with blueness.
7. Clouds appear, a storm blows up, lighting flashes.
8. A rainbow appears.

'In answering items 9 to 12, think of the front of a shop which you often go to.

9. Picture the overall appearance of the shop from the opposite side of the road.
10. Picture a window display including colours, shape and details of individual items for sale.
11. You are near the entrance. Picture the colour, shape and details of the door.
12. You enter the shop and go to the counter. The counter assistant serves you. Money changes hands.

In answering items 13 to 16, think of a country scene which involves trees, mountains and a lake.

13. The contours of the landscape.
14. The colour and shape of the trees.
15. The colour and shape of the lake.
16. A strong wind blows on the trees and on the lake causing waves.

Respones scale for each question:

No image at all, 1,
Vague and dim image, 2
Moderately clear and vivid, 3
Reasonably clear and Vivid, 4
Perfectly clear and vivid, 5

Appendix P. VLFI Questions – Cohn et al., (2012)

Rated 1 – 7:

On average, how often per week do you read text-only books for enjoyment?
On average, how often per week do you watch films?
On average, how often per week do you watch cartoons/anime?
On average, how often per week do you read comic books?
On average, how often per week do you read comic strips?
On average, how often per week do you read graphic novels?
On average, how often per week do you read Japanese comics (manga)?
On average, how often per week do you draw comics?

On average, how often per week did you read text-only books for enjoyment while growing up?
On average, how often per week did you watch films while growing up?
On average, how often per week did you watch cartoons/anime while growing up?
On average, how often per week did you read comic books while growing up?
On average, how often per week did you read comic strips while growing up?
On average, how often per week did you read graphic novels while growing up?
On average, how often per week did you read Japanese comics (manga) while growing up?
On average, how often per week did you draw comics while growing up?'],

Rated 1 – 5:

How would you currently rate your expertise with reading comics (of any sort)?
How would you currently rate your drawing ability?
How would you rate your expertise with reading comics (of any sort) while growing up?
How would you rate your drawing ability while growing up?

Free Answer:

Approximately how old were you when you first read a comic?
Approximately how old were you when you first started drawing comics?